



# **Making Impact Evaluation Matter Better Evidence for Effective Policies and Programs**

---

3–5 September 2014 • ADB Headquarters

## **Conference Program**



#IEmatters

# Acknowledgments

## Conference Advisory Committee

Juzhong Zhang, Deputy Chief Economist, Economics and Research Department, Asian Development Bank (ADB)  
Vinod Thomas, Director General, Independent Evaluation Department, ADB  
Gilberto Llanto, President, Philippines Institute for Development Studies (PIDS)

## Conference Steering Committee

Cyn-Young Park, Assistant Chief Economist, Economics and Research Department, ADB  
Sakiko Tanaka, Senior Economist, Economics and Research Department, ADB  
David Raitzer, Economist, Economics and Research Department, ADB  
Binh Nguyen, Senior Evaluation Specialist, Independent Evaluation Department, ADB  
Howard White, Executive Director, International Initiative for Impact Evaluation (3ie)  
Radhika Menon, Senior Communication Officer, 3ie  
Heather Lanthorn, Evaluation Specialist, 3ie  
Marian Adel, Consultant, 3ie  
Sheila Siar, Director, PIDS

## Conference Team

This conference has been the product of collaboration and teamwork of agencies and a large number of individuals. Implementation of conference design and arrangements at ADB headquarters was undertaken by a team within the ADB Economics and Research Department, Economic Analysis and Operations Support Division, under the leadership of Assistant Chief Economist Cyn-Young Park. This team included members of the Steering Committee, as well as Lilibeth Poot, Maria Melissa Gregorio-Dela Paz, Mary Jane Romero and Angeli Magbanua, who assisted with logistical arrangements. Design contributions and support were provided by ADB Impact Evaluation Committee chaired by Hans Carlsson and co-chaired by Christopher Edmonds. Colleagues from the Department of External Relations provided support for media relations and program design. A team at 3ie undertook promoting the conference internationally, managing the call for workshops and presentations, the conference website and App, raising funds for the event from donors, overall program design, logistical arrangements for funded participants, conference registration, producing the videos to support the event and overall coordination. Logistical arrangements have been managed by Bindu Joy, Mithlesh Joshi and Marian Adel. They also managed conference registration with the assistance of Nilesh Chauhan. Promotion of the event was undertaken by Paromita Mukhopadhyay, Kanika Jha and Rajesh Sharma with advice and supervision from Beryl Leach. The review process was facilitated by Heather Lanthorn and Marian Adel. Website design and construction was undertaken by Wishtree with direction and inputs from Charu Kanwar, Radhika Menon, Kanika Jha and Rajesh Sharma, who also worked on the Conference App, and were also part of the video production team which included Pradeep Singh, Drew Cameron, Ami Bhavsar, Kes Stern, and Charlie Navy. 3ie staff also assisted with media engagement and the use of online platforms to promote the event. Reporting to donors on the use of funds will be done by Jatin Juneja and Ditto Joy under the supervision of Hitesh Somani. PIDS research fellows Aniceto Orbeta and Jose Ramon Albert provided expertise in reviewing the submitted proposals. PIDS staff team including Jane Alcantara, Anne Cleofas, Clarissa Lagoras, Jose Ignacio Tenorio, Kristine Carla Oteyza, Felipe Salvosa, Claudette Malana, Rommel Lopez, and Gizelle Manuel provided logistics and documentation support as well as assistance in coordinating with the national media.

## Conference Sponsors

Conference benefited from financial support by ADB, 3ie, PIDS, IDRC, USAID and the William and Flora Hewlett Foundation.

# Overview of Conference Program

## Wednesday, 3 September

- 2:30 p.m.–3:15 p.m. **Opening Ceremony**
- Bindu N. Lohani, Vice-President for Knowledge Management and Sustainable Development, ADB
- Socioeconomic Planning Secretary Arsenio M. Balisacan, Director General of the National Economic and Development Authority and Chairman of the PIDS Board of Trustees
- Chair: Richard Manning, Chair International Initiative for Impact Evaluation Board of Commissioners*
- 3:20 p.m.–4:20 p.m. **Opening Plenary: ‘Challenges in Using Rigorous Evidence to Make Better Policy’**
- Professor Paul Gertler, Professor of Economics, University of California, Berkeley
- Secretary of Social Welfare and Development Corazon Juliano Soliman
- Chair: Gilberto Llanto, President Philippine Institute for Development Studies*
- 4:20 p.m.–4:50 p.m. Break
- 4:50 p.m.–6:20 p.m. Parallel Session 1
- Session 1 RCTs: Rhetoric, Reality, and Practice**  
Auditorium A
- Session 2 Evaluating the impact of roads**  
Auditorium B
- Session 3 Training for Young Asian Workers: Nongovernment vs. Government Perspectives**  
Auditorium C
- Session 4 The Impact of Farmer Training: Evidence from Systematic Reviews**  
Auditorium D1
- Session 5 Health systems reform**  
Auditorium D2
- Session 6 Funding Female Enterprises**  
Briefing Theater 2

4:50 p.m.–5:30 p.m. Special Session on Impact Evaluations in the Philippines  
Briefing Theater 1

6:30 p.m. onwards Cocktail Reception, Courtyard

### Thursday, 4 September

9:00 a.m.–10:30 a.m. Parallel Session 2

**Session 7 Health, Nutrition, and Education: Evidence from Field Experiments in the People’s Republic of China**  
Auditorium A

**Session 8 Impact of Water and Sanitation Interventions**  
Auditorium B

**Session 9 Self-Help Groups, Empowerment and Poverty Reduction in Rural India**  
Auditorium C

**Session 10 REDD-y and Warned for Impact Evaluation: Making the Case for Evaluating Climate Change Initiatives in REDD/REDD+ Projects and Early Warning Systems**  
Auditorium D1

**Session 11 Better Data for Better Impact Evaluations**  
Auditorium D2

**Session 12 Microfinance in South Asia**  
Briefing Theater 2

**Session 13 Tax, Property Rights, and Crime**  
Briefing Theater 1

10:30 a.m.–11:00 a.m. Break

11:30 a.m.–12:30 p.m. Parallel Session 3

**Session 14 Measuring Direct and Indirect Impacts of Policy and Program Interventions**  
Auditorium A

**Session 15 The Design and Implementation of WASH Impact Evaluations**  
Auditorium B

**Session 16 Health Insurance**  
Auditorium C

**Session 17 Investing in Smallholder Agriculture**  
Auditorium D1

**Session 18 Education in Africa**  
Auditorium D2

**Session 19 The Challenge of Evaluating Women’s Empowerment in South Asia**  
Briefing Theater 2

**Session 20 Using Impact Evaluation to Assess Agency Performance**  
Briefing Theater 1

12:30 p.m.–1:15 p.m. Lunch, Executive Dining Room

1:15 p.m.–2:00 p.m. Poster Presentation, Cafeteria Hallway

2:00 p.m.–3:30 p.m. Parallel Session 4

**Session 21 Education in Asia**  
Auditorium A

**Session 22 Impact Evaluations in Infrastructure Planning and Policy in Urban and Rural Contexts—Uses, Challenges, and Opportunities**  
Auditorium B

**Session 23 Improving Health Outcomes through Private Sector Interventions—Experimental Evidence from Ghana, Kenya, Jordan and Madagascar**  
Auditorium C

**Session 24 Designing Agricultural Impact Evaluations**  
Auditorium D1

**Session 25 The Success of Poverty Alleviation Programs**  
Auditorium D2

**Session 26 Complexity, External Validity, and Policy Influence**  
Briefing Theater 2

**Session 27 The role of think tanks in impact evaluation**  
Briefing Theater 1

3:30 p.m.–4:00 p.m. Break

4:00 p.m.–5:30 p.m. Parallel Session 5

**Session 28 Approaches to Infrastructure Impact Evaluations**  
Auditorium A

**Session 29 Evaluating Energy and Mining Investments**  
Auditorium B

**Session 30 Assessing Health Impacts**  
Auditorium C

**Session 31 Impact Evaluations In and On Conflict: Methodological Innovations and Implications from the Field**  
Auditorium D1

**Session 32 Intended and Unintended Consequences of Social Protection**  
Auditorium D2

**Session 33 Generating More and Better Employment**  
Briefing Theater 2

**Session 34 Evaluating Governance Interventions**  
Briefing Theater 1

### Friday, 5 September

9:00 a.m.–10:30 a.m. Parallel Session 6

**Session 35 Community-Driven Development in Asia**  
Auditorium A

**Session 36 ICT and Education**  
Auditorium B

**Session 37 Improving Newborns and Infant Health and Nutrition**  
Auditorium C

**Session 38 Promoting Adoption of New Crops and Techniques in Agriculture**  
Auditorium D1

**Session 39 Information, financial literacy and financial inclusion**  
Auditorium D2

**Session 40 Donor agency experiences with impact evaluation**  
Briefing Theater 2

10:30 a.m.–11:00 a.m. Break

11:00 a.m.–12:30 p.m. Parallel Session 7

**Session 41 Evaluating *Pantawid Pamilyang Pilipino Program* (4Ps)**  
Auditorium A

**Session 42 Infrastructure Matchmaking Clinic**  
Auditorium B

**Session 43 Policies for Better Maternal and Reproductive Health**  
Auditorium C

**Session 44 Keeping Kids in School: Evidence from Field Experiments in the People's Republic of China**

Auditorium D1

**Session 45 Does Community Driven Development Work?**

Auditorium D2

**Session 46 Experiences in Producing, Managing and Using Impact Evaluations**

Briefing Theater 2

12:30 p.m.–1:30 p.m. Lunch, Executive Dining Room

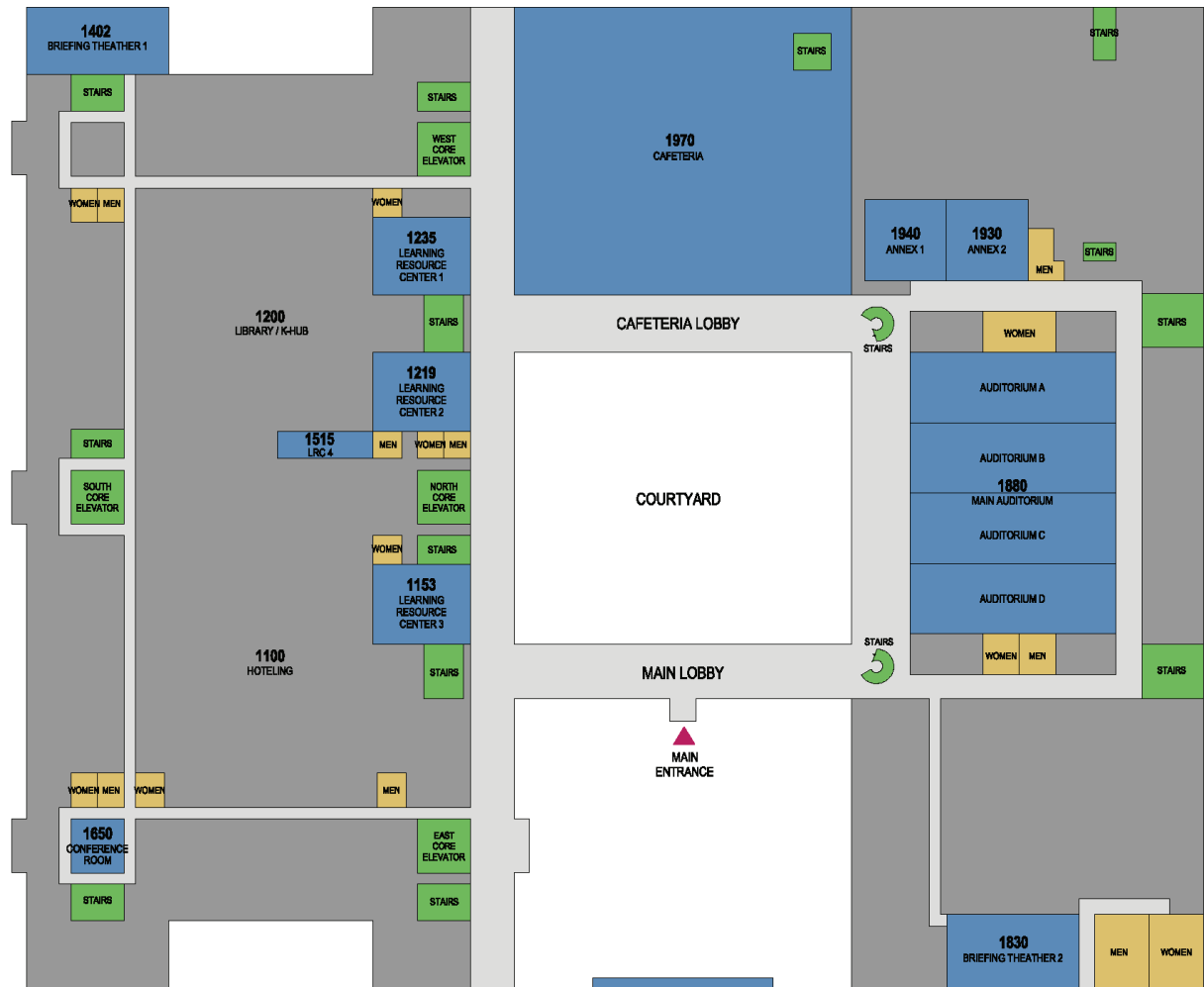
**Closing Plenary**

1:30 p.m.–1:45 p.m. Award ceremony for conference prize winners

1:45 p.m.–3:30 p.m. **Panel Discussion: 'What Have We Learned and Where Do We Go from Here?'**

Panel: Hon. Kabwegyere, Minister of General Duties, Office of the Prime Minister, Uganda; Emmanuel Jimenez, Director, Public Sector Evaluations, IEG, World Bank, incoming 3ie Executive Director; Juan Miranda, ADB Managing Director General; Deputy Director General Rolando G. Tungpalan, National Economic and Development Authority (tbc)  
*Chair: Shang-Jin Wei, Chief Economist, ADB*

# ADB First Floor Map



- LEGEND:**
- ASSEMBLY AREAS
  - ELEVATORS / STAIRS
  - OFFICES
  - RESTROOMS
  - HALLWAY

## LOCATION MAP












# Overview of Parallel Sessions

	Auditorium A	Auditorium B	Auditorium C	Auditorium D1	Auditorium D2	Briefing Theater 2	Briefing Theater 1
<b>Wednesday, 3 September</b>							
4:50 p.m.– 6:20 p.m.	1. RCTs: Rhetoric, Reality, and Practice	2. Evaluating the Impact of Roads	3. Training for Young Asian Workers: Nongovernment vs. Government Perspectives	4. The Impact of Farmer Training: Evidence from Systematic Reviews	5. Health Systems Reform	6. Funding Female Enterprises	Press Conference and Media Session
6:30 p.m. onwards	Cocktails Venue: Courtyard						
<b>Thursday, 4 September</b>							
9:00 a.m.– 10:30 a.m.	7. Health, Nutrition, and Education: Evidence from Field Experiments in People's Republic of China	8. Impact of Water and Sanitation Interventions	9. Self-Help Groups, Empowerment and Poverty Reduction in Rural India	10. Evaluating Climate Change Initiatives in REDD/REDD+ Projects and Early Warning Systems	11. Better data for Better Impact Evaluations	12. Microfinance in South Asia	13. Tax, Property Rights, and Crime
11:00 a.m.– 12:30 p.m.	14. Measuring Direct and Indirect Impacts of Policy and Program Interventions	15. The Design and Implementation of WASH Impact Evaluations	16. Health Insurance	17. Investing in Smallholder Agriculture	18. Education in Africa	19. The Challenge of Evaluating Women's Empowerment in South Asia	20. Using Impact Evaluation to Assess Agency Performance
12:30 p.m.– 1:15 p.m.	Lunch, Executive Dining Room						
1:15 p.m.– 2:00 p.m.	Poster Presentation, Cafeteria Hallway						
2:00 p.m.– 3:30 p.m.	21. Education in Asia	22. Infrastructure Planning and Policy in Urban and Rural Contexts	23. Improving Health Outcomes through Private Sector Interventions	24. Designing Agricultural Impact Evaluations	25. The Success of Poverty Alleviation Programs	26. Complexity, External Validity, and Policy Influence	27. The Role of Think Tanks in Impact Evaluation

	Auditorium A	Auditorium B	Auditorium C	Auditorium D1	Auditorium D2	Briefing Theater 2	Briefing Theater 1
4:00 p.m.– 5:30 p.m.	28. Approaches to Infrastructure Impact Evaluations	29. Evaluating Energy and Mining Investments	30. Assessing Health Impacts	31. Impact Evaluations In and On Conflict: Methodological Innovations and Implications from the Field	32. Intended and Unintended Consequences of Social Protection	33. Generating More and Better Employment	34. Evaluating Governance Interventions
<b>Friday, 5 September</b>							
9:00 a.m.– 10:30 a.m.	35. Community-Driven Development in East Asia	36. ICT and Education	37. Improving Newborns and Infant Health and Nutrition	38. Promoting Adoption of New Crops, Techniques and Technology in Agriculture	39. Information, Financial Literacy and Financial Inclusion	40. Donor Agency Experiences with Impact Evaluation	Special Session: Establishing the Philippine Impact Evaluators Network for Rural and Agricultural Development
11:00 a.m.– 12:30 p.m.	41. Evaluating <i>Pantawid Pamilyang Pilipino Program</i> (4Ps)	42. Infrastructure Matchmaking Clinic	43. Policies for Better Maternal and Reproductive Health	44. Keeping Kids in School: Evidence from Field Experiments in People's Republic of China	45. Does Community Driven Development Work?	46. Experiences in Producing, Managing and Using Impact evaluations	

Code to Parallel Session Themes

	Infrastructure		Institutional and policy reform (including public management and governance)
	Climate change, Environment, Natural resources, Agriculture, Rural development		Institutionalization of IE
	Financial inclusion and governance		IE and SR methods
	Social sector including health, education, TVET, and social protection		

## Session 1

### RCTs: Rhetoric, Reality, and Practice

*Chair: Pam Jagger, University of North Carolina at Chapel Hill*

#### Use of Randomization to Rigorously Evaluate International Development Initiatives: Are the Perceived Obstacles False Alarms?

*Stephen H. Bell, Abt Associates Inc., USA*

When deciding how to allocate limited funds for social programs, policymakers and funders in the United States and the international development community increasingly ask for evidence of effectiveness based on studies that rely on solid methodology, providing credible scientific evidence. Experimental evaluations that use random assignment provide this credible evidence. Despite their transparency and conceptual strength in revealing the causal connection between social initiatives or economic development interventions and the outcomes of their participants, skeptics often question the wisdom or feasibility of conducting social experiments on a variety of grounds. As consideration of the use of experimental design for measuring causal impacts extends around the globe, it is important to know if these objections hold up to careful scrutiny and creative analytic responses. This presentation defines 15 concerns regarding randomized experiments in social, economic and development applications and provides examples from the scholarly literature and applied evaluations in the United States. It then examines the basis for and soundness of each concern in an international development context and contrasts the potential difficulty posed by issues arising over several decades in America when taken to newer venues. Results suggest that a range of issues concerning ethics, scientific integrity, and practical feasibility need not stand in the way of expanded use of social experiments even under circumstances commonly thought of as prohibitive. If vetted and found sound by the community of development program evaluators, these conclusions have the potential to broaden the use of randomized experiments in development policy evaluation in all parts of the world.

#### Slips between the Cup and the Lip: Experiences with Planning and Implementing Impact Evaluations

*Jyotsna Puri and Radhika Menon, International Initiative for Impact Evaluation (3ie)*

Impact evaluations rarely get implemented the way they were planned. Our presentation will draw on 3ie's experiences with supporting more than 150 impact evaluations in over 50 low- and middle-income countries. The experiences have been pulled together from impact evaluation proposals, progress reports, final reports and field visits to project sites.

We focus on the lessons we have learned through the entire lifecycle of impact evaluations. This includes planning, designing, implementing an impact evaluation and ensuring the uptake of findings in informing policies and programs. We will present results from an in-house portfolio review of various delays that impact evaluations face.

The presentation will use examples which illustrate the various ways in which impact evaluations diverge from what is planned on paper, and the challenges encountered by researchers and implementers on the ground. In particular, we will analyze the complexities of making impact evaluations work on the ground in diverse contexts.

In drawing lessons from these experiences, we will highlight the key barriers to and facilitators of successful impact evaluations. High quality impact evaluations require a strong emphasis on formative research, a comprehensive theory of change and the analysis of outcomes across the causal chain. The uptake of impact evaluation findings however requires on-going engagement with policymakers and

implementers. Findings need to also address real policy priorities, be translated for policymaker needs and be timely and appropriate to the context.

Lessons will be drawn for planners, researchers and users of impact evaluations.

## Session 2

### Evaluating the Impact of Roads

*Chair: Jack Molyneaux, Millennium Challenge Corporation*

#### How Investment in Road Quality Boosts Economic Activity and Household Welfare

*Paul Gertler, University of California Berkley, USA*

Maintaining and improving road quality, a major function of government, is often justified as a major public good investment in economic activity and job creation. Unlike previous empirical work on roads that focuses on gains from trade arguments, we posit a model based on roads as a productive input into firm and agricultural productivity and test the predictions of the model using data from Indonesia. We estimate the effect of road network quality on economic activity and welfare in Indonesia using a uniquely detailed nationwide panel dataset of road surface roughness merged into panel household and manufacturing firm surveys. We identify the impact of road quality from temporal variation in national, provincial and districts budgets for road maintenance authorities under the testable assumption of two stage budgeting.

We find that higher road network quality causes improvements in household consumption and income. In terms of mechanisms, we show that roads increase firm profitability and job creation in the manufacturing sector. Second, we show evidence of an occupational shift from agriculture into manufacturing and higher profits for those who stay in agriculture. The gap in average earnings between agriculture and manufacturing employment is reduced with road quality but not eliminated. Because wages in the manufacturing sector do not change with road quality the results are consistent with the existence of dual labor markets. Our results improve upon previous work on the effects of roads by analyzing the effects of comprehensive road quality measures and by using household level outcomes which are not confounded with migration effects.

#### Re-examining Poverty and Road Access in Rural Papua New Guinea

*Christopher Edmonds, Senior Economist, Pacific Department, Asian Development Bank (ADB)*

This evaluation proposes to use newly available Household Income and Expenditure Survey data, as well as new geo-referencing and space imaging (luminosity) techniques to update an earlier study by Gibson and Rozelle (2003) that provided measures of how access to transport infrastructure, specifically rural feeder roads, impacted poverty in rural PNG—with a focus on the Highlands region. The assessment will cover the entire highlands region with a focus on rural areas, where approximately 85% of the populations live. The results of this study will be used to estimate the overall impact of ADB and Government road sector investments on poverty and welfare in the Highlands region, and provide a benchmarking tool for the design of future investments into the region. Analysis of data consolidated from 1996 and 2009–10 HIES will focus on household level analysis using quasi-experimental and econometric methods. The analysis will be primarily quantitative in nature, complemented by qualitative descriptions of household and individual welfare in the region. Poverty estimates will be combined with data on the quality of rural feeder roads during this period as measured from Department of Transport data (including data on

ADB road projects), and supplemented with road quality data obtained from remote sensing of Linear Imaging Self Scanning Sensor (LISS-III) data. Available data set of 'luminosity' in the highlands region will also be employed to more accurately assess the impact of road conditions on poverty and welfare in the highlands region. The research will seek to clarify the empirical relationship between poverty and access to roads, measured as traveling time from a community to the nearest road. The analysis will test the impact of access to transport infrastructure on a household's poverty status between 1996 and 2010. Recognizing the possibility that access to roads is endogenous with other factors that drive household poverty status, impact measures will control for the effect of such endogeneity on estimation results using an instrumental variable approach in estimating the model. The comparatively rich data set will also be used to estimate the impact that improving average years of schooling and literacy has in the marginal impact of improved road quality and poverty reduction.

Reference: Gibson, J. and Rozelle, S. (2003). *Economic Development and Cultural Change*, Vol. 52, No. 1 (October), pp. 159–185.

### **Impact of Rural Road rehabilitation in Sri Lanka**

*Dinuk Jayasuriya, ANU*

The Asian Development Bank disburses a considerable number of loans for infrastructure projects, such as road rehabilitation. However, impact evaluations on these infrastructure projects are rare. Partly in response to this, ADB have commissioned one impact evaluation focusing on the impacts of road rehabilitation in Sri Lanka. In June 2014, approximately 2750 households and 250 enterprises were surveyed in 50 villages, surrounding 16 roads, for the purposes of a baseline survey. Four roads are counterfactual, while the remaining are treatment roads. Preliminary analysis reveals that the households surrounding treatment and counterfactual roads are statistically similar. iPad technology was used to capture survey data, and GPS devices were used to capture household location. We are also exploring the use of cellular technology to investigate travel times and distances of respondents over the intervention period. An endline survey is expected to be undertaken in two to three years' time. Additionally, ADB is investigating undertaking a larger impact evaluation, which would involve assessing the impact of 200 roads across 500 villages in the Southern Province of Sri Lanka. If approved, this study would easily represent the largest impact evaluation of road rehabilitation globally.

The presentation will discuss the road rehabilitation IE currently underway, as well as the associated challenges. It will also touch upon the larger planned road rehabilitation IE for Sri Lanka.

## **Session 3**

### **Training for Young Asian Workers: Nongovernment vs. Government Perspectives**

*Chair: Emmanuel Jimenez, Independent Evaluation Group, World Bank*

Young workers face the challenge of finding new job opportunities in Asian cities. The panel looks at two different experiences of young unemployed workers training for jobs in garment factories in Dhaka or for job opportunities in Mongolia. Across Asia, well-intentioned interventions to training the young are doomed to fail if they do not take into labor market conditions. These two evaluations are in their initial stages. Both teams will share the design and inception phase of two very different interventions: a custom-made training led by a nongovernment organization and a government vocational training. Each team brings a different perspective from the direct users of the evaluation: a Bangladeshi nongovernment organization working with rural youth in Northwestern Bangladesh and the Mongolian Ministry of Labor.

The panel will answer questions such: Do Southern-led evaluations bring a different perspective? Can their evaluations provide lessons to other similar efforts in the developing world? How these evaluations fit the local needs?

**Bangladesh: Training of Young Workers from Gaibandha to Dhaka, Evaluation by South Asian Network for Economic Modelling (SANEM), IZA-DFID Growth and Labor Markets in Low Income Countries, and IDRC**

*Abu Shonchoy, Institute of Developing Economies, Japan and Representative from a Bangladeshi NGO (GUK)*

**Mongolia: Vocational Training for Unemployed Youth in Mongolia, Evaluation by the Partnership for Economic Policy (PEP), DFID, and IDRC**

*Altantsetseg Batchuluun, National University of Mongolia and Tungalag Nyamaa, Government of Mongolia*

**The Impact of Vocational Schooling on Human Capital Development in Developing Countries: Evidence from the People's Republic of China**

*Prashant Loyalka, Stanford University*

The primary aim of this study is to understand whether VET at the high school level in fact contributes to human capital development in developing countries. To fulfill that aim, we conduct two sets of analyses using longitudinal data on more than 10,000 vocational and academic high school students in the People's Republic of China. First, we estimate the causal impacts of attending vocational versus academic high school on dropout, specific skills and general skills. Estimates from matching and instrumental variables analyses show that attending vocational high school (relative to academic high school) substantially reduces general skills and does not improve specific skills. Heterogeneous effect estimates also show that attending vocational high school increases dropout, especially among disadvantaged (low-income or low-ability) students. Second, we use vertically scaled (equated) baseline and follow-up test scores to measure gains in specific and general skills among the students. We find that students who attend vocational high school experience absolute reductions in their general skills. Taken together, our findings indicate that the rapid expansion of vocational schooling as a substitute for academic schooling may in fact be detrimental to building human capital in developing countries, such as the People's Republic of China.

**Evaluating the Promotion of Technical and Vocational Education and Training (TVET) in Viet Nam**

*Jochen Kluge, Humboldt-Universität zu Berlin and RWI*

We present an evaluation of a TVET intervention in Viet Nam, a program jointly implemented by the Germany–Viet Nam development cooperation. It consists of three components: teacher training; curricula reform; technical support (machinery for welding instructions, etc.)

The project is remarkable because it brings together all relevant actors to form a unique collaboration: the local team that implements the project; the responsible line ministry MoLISA; the responsible government authority, General Department for Vocational Training (GDVT); a Vietnamese research institute, National Institute for Vocational Training (NIVT); all TVET schools involved; and the evaluation research team.

Specifically, the collaboration started in 2009, and since then we trained staff from the TVET schools to collect survey data, receiving full support from MoLISA and GDVT. The institute NIVT was integrated to provide continuing advice to schools in collecting data. Because of the unrestricted support by the



government, the study was also extended to TVET schools that did not receive the intervention, to use as control schools in the impact evaluation. Identification is based on a cross-sectional DiD design using variation within schools (supported, non-supported occupations) and between schools (treatment, control). The data cover 9,000 TVET graduates.

Another remarkable aspect is that this set-up generated sustainable structures of monitoring among TVET schools and gathering data at the level of federal government authorities. Since all stakeholders were involved from the beginning, ownership and institutional buy-in has been high throughout the project. Hence, continuous structures were created that no longer depend on support from the German development cooperation.

## Session 4

### The Impact of Farmer Training: Evidence from Systematic Reviews

*Chair: Fedes van Rijn, LEI Wageningen UR, The Netherlands*

#### Farmer Field Schools for Improving Farming Practices and Armer Outcomes in Low- and Middle-Income Countries: A Systematic Review

*Hugh Waddington and Martina Vojtkova, 3ie*

The majority of the population in low-income countries, including some of the poorest and most vulnerable people in the world, rely on agriculture for their livelihoods and survival. Since the late 1980s, support to agriculture has moved from top-down agricultural extension towards more participatory approaches which better suit smallholder farmers. One such approach is the Farmer Field School (FFS), a farmer-centered participatory adult education approach that involves farmer experimentation and development of problem solving capabilities. FFSs have been used to train 12 million farmers in over 90 countries across Asia, Africa and Latin America.

This presentation provides results of a mixed-methods systematic review of existing evidence on the effectiveness of FFS and associated barriers and facilitators. The review synthesizes over 500 impact evaluations, qualitative studies and project documents along the causal chain. The review finds beneficial effects of farmer field schools on farmer outcomes, including adoption of FFS practices, increased yields and revenues and increased perceptions of empowerment. Factors that may affect the effectiveness of FFS include targeting and participant characteristics, the appropriateness and delivery of the curriculum, facilitation skills of FFS trainers and others. However, the evidence suggests that FFS have not been effective when taken to scale and there is not enough convincing evidence of a diffusion effect to non-trained members of the community. The review suggests that FFSs are unlikely to be cost-effective in comparison with other agricultural approaches, and there is insufficient evidence on empowerment impacts to assess whether FFS are a cost-effective way to empower the poor.

#### Is Targeting Effective? The Impact of Targeting on Farmer Field School Programs

*Daniel Phillips, 3ie*

This presentation reports on the findings of a mixed-methods systematic review that draws on content analysis, meta-analysis and meta-regression analysis to explore how a particular type of agricultural program called 'Farmer Field Schools' (FFS) are targeted and how this process affects program participation and impact. Targeting is intended to ensure that limited resources are used efficiently and that a program's impacts on outcomes of interest are maximized. Drawing upon evidence from

FFS programs around the world, the presentation examines how programs are targeted, whether the targeting works (do programs reach their intended beneficiaries?) and whether targeting affects outcomes of interest. Some FFS programs are found to include ‘equity’ criteria and specifically target the poorest farmers. Such programs, however, may include ‘effectiveness’ criteria designed to promote inclusion of farmers with more resources, education and social agency, with the aim of maximizing program effectiveness. While programs typically achieved effectiveness-related inclusion objectives, some failed to fulfill equity-related inclusion goals. There is also evidence that programs with relatively more educated participants may be more effective in improving adoption of farming practices, increasing yields and passing on learning to neighbor farmers living in the same communities.

### **The Effectiveness of Training, Innovation, and New Technology on Smallholder Farmers’ Wealth and Food Security: Findings from a Systematic Review**

*Evans Muchiri, University of Johannesburg*

This presentation outlines the findings from a Campbell Collaboration systematic review commissioned by Canadian Foreign Affairs, Trade and Development agency (DFATD) and International Initiative for Impact Evaluation (3IE) to identify interventions that work to improve agricultural outcomes for smallholder farmers.

The review adopted an innovative three-stage approach where a systematic review-of-reviews, a scoping map, and a traditional Campbell effectiveness review were conducted. Studies were screened, coded and inter-coder reliability scores were calculated for a randomly selected sample using the Cohen’s Kappa Statistic. The review also piloted the Cochrane tool for assessing risk of bias in non-randomized studies. Reviewed studies were appraised against six potential sources of bias and we only included studies with the fewest threats to validity. In this presentation, we reflect on our experiences applying the tool compared to other critical appraisal tools for systematic reviews. We planned to perform a statistical meta-analysis of homogenous studies with sufficient summary data.

The exhaustive search yielded 15,437 citations from 39 sources. Only 19 studies met our inclusion criteria (15 on new technology and 4 on training). Meta-analysis of extracted findings could not proceed because of heterogeneity between studies. Furthermore, summary statistics were not always reported for included studies, making it problematic to standardize effect sizes.

Explanatory evidence from this review points to the value of agricultural interventions to positively influence smallholder farming in improving food security and household incomes. However, this evidence is based on individual studies and not a synthesized result. Lessons for Asian smallholder farmers will be presented.

## **Session 5 Health Systems Reform**

*Chair: Neha Kumra, Institute of Applied Manpower Research, India*

### **The Political Economy of Public Sector Absence: Experimental Evidence from Pakistan**

*Saad Gulzar, New York University and Muhammad Yasir Khan, International Growth Centre*

In many developing countries, public sector absence is both common and resistant to reform. One explanation for this is that politicians provide public jobs with limited work requirements as patronage.



We test this patronage hypothesis in Pakistan using: (i) a randomized controlled evaluation of a novel smartphone absence monitoring technology; (ii) data on election outcomes in the 240 constituencies where the experiment took place; (iii) attendance recorded during unannounced visits and; (iv) surveys of connections between local politicians and health staff. Four results support this view. First, while doctors are present at 42% of clinics in competitive constituencies, they are present at only 13 percent of clinics in uncompetitive constituencies. Second, doctors who know their local parliamentarian personally are present at an average of 0.727 of three unannounced visits, while doctors without this connection are present at 1.3 of the three visits. Third, the effect of the smartphone monitoring technology, which almost doubled inspection rates, is highly localized to competitive constituencies. Last, we find evidence that program impact is in part due to the transmission of information to senior officers. We test this by manipulating the salience of staff absence in data presented to officials using an online dashboard. These effects are also largest in politically competitive constituencies. Our results have implications for the study of bureaucratic incentives in fragile states and are potentially actionable for policymakers trying to build state capacity.

### **Social Accountability through Health Facility Committees: Pilot Experiments in Burundi**

*Jean-Benoît Falisse, DPhil candidate, Oxford Department of International Development; Juvenal Ndayishimiye, MD, MPH, Cordaid Burundi; René Queffelec, Cordaid Burundi*

For decades now, community-elected Health Facility Committees (HFC) has been described as a strategy for ensuring that health facilities adequately respond to the needs and concerns of their users. Yet, the efficiency of this mechanism is rarely evaluated. This paper seeks to answer a double question: “how can HFCs be fruitfully supported?” and “what difference do they make in terms of management, quality, and use of services?”

A randomly controlled pilot project was carried out between 2011 and 2013 in rural Burundi (n = 252 HFCs). It tested four interventions: (1) informing HFC members and the medical staff about the missions of the HFCs; (2) team-building activities between HFC members and medical staff; (3) introduction of reporting cards to monitor performances; and (4) facilitation of meetings with local leaders.

Early assessments do not show an obvious impact of the project on the use of services (number of visits, deliveries, etc.). Team-building activities could not fix enduring intra-HFC and HFC-medical staff conflicts and the idea of introducing joint HFC-local leaders meetings did not survive the progressive withdrawal of project facilitators. However, accountability indicators have improved, especially in the case of interventions 3 and 4, as well as indicators on the quality of services (days of stock-out and human resources).

Overall, the attempt to make HFCs more functional may well have resulted an increased social accountability but more in the sense of responsiveness than answerability. The original interventions have been modified to integrate lessons from the pilot project and they are being scaled up.

### **Measuring the Impact of the Second Urban Primary Health Care Project in Bangladesh**

*Brian Chin, Social Sector Economist, ADB; Manuel Leonard F. Albis, University of the Philippines; and Subrata K. Bhadra, National Institute of Population Research and Training*

The Urban Primary Health Care Project (UPHCP) was developed by the Government of Bangladesh with financial support from the Asian Development Bank (ADB). The goal of the project is to provide health care services to the urban poor living in city corporations and urban municipalities, through public-private partnerships with NGOs. The project started in 1998 and is currently in its third phase,

covering more urban areas than the first two phases. This paper evaluates the impact of the second phase (UPHCP-II) on health outcome indicators, mainly child mortality, child nutrition, immunization, maternal mortality, women's access to antenatal care, and health-related behavioral change. The effect of the project was estimated through propensity score matching (PSM) between designated project and non-project areas. The baseline and endline surveys that were used, namely the 2006 Urban Health Survey (UHS) and the 2012 UPHCP-II Endline Household Survey, had different sampling designs but similar Demographic and Health Survey-type questionnaires. An innovation of this paper is the recalibration of the sampling weights that allows the use of two unrelated surveys in impact evaluation. The results of the paper will help inform implementation of ongoing urban health interventions and design of future ones in Bangladesh.

## Session 6

### Funding Female Enterprises

*Chair: Natalie Chun, Economist, ERD, ADB*

#### Impact Assessment of Loans to Female-Run Start-Ups

*Farah Said, Lahore School of Economics, Pakistan*

The presentation will discuss a Randomized Control Trial involving business loans provided to women in households where no female-run enterprise exists currently. The study is being conducted in selected districts of Punjab, Pakistan in collaboration with one of the largest private sector microfinance providers in the country and is expected to inform policy making by investing in a product that is targeted to draw new clients into their client pool. From a practical point of view, researchers can evaluate the viability of the product as an alternative to existing microfinance models that do not focus on micro-loans targeted towards start-ups.

Problems with randomizing at the individual level has traditionally meant that majority of studies have relied on clustered RCTs comparing outcomes across villages or neighborhoods. The presentation will explain the randomization strategy employed in this study that is expected to provide better balance and one that will eventually provide us with a pool of about 900 treatment and control clients in the selected districts. Baseline survey and disbursement of loans is underway and there will be a follow-up survey after one year of the completion of the baseline, coinciding with the end of the loan cycle. The presentation will then move on to discussing the expected outcomes i.e. improvement in household gender empowerment levels and socioeconomic well-being. Finally, we will conclude by explaining the mechanisms that will be used to control for possible network effects and attrition.

#### Access to Credit and Women Entrepreneurship: Evidence from Bangladesh

*M. Jahangir Alam Chowdhury, University of Dhaka Bangladesh, Bangladesh*

This paper intends to assess the impact of the participation in microcredit programs on women entrepreneurship. The women entrepreneurship has been defined as the ownership of businesses by women. The descriptive statistics and multivariate techniques have been used to achieve the objective of the paper. The study uses Household Income and Expenditure survey (HIES) 2010 data set. HIES 2010 survey covers 12,240 households from all districts in the country. Considering the endogeneity in the microcredit program participation of women, the study uses instrumental variable method (IV method) for assessing the impact of the microcredit program participation on the entrepreneurial status of women. After the adjustment of the endogeneity, the results from multivariate technique indicate that the participation of households in the microcredit programs does not promote women

entrepreneurship. This result highlights that women do not use their microcredit loans for owning businesses. Rather, they pass those loans on to household male members. So, it is evident from these results that the easing of credit constraint does not necessarily promote women entrepreneurship. Therefore, it is important for the policy makers in the government and the microfinance sector in Bangladesh to reconsider the existing methodology of giving microcredit loans to women. The redesigning of the microcredit loan giving methodology is necessary for enabling women to use their microcredit loans for income generating purposes.

### **The Effects of Financial Inclusion beyond Financial Outcomes**

*Carlos Chiapa, El Colegio de Mexico, Mexico*

We investigate the effects, beyond financial outcomes, of offering access to savings account to a sample of poor women. A field experiment in Nepal randomized access to savings accounts among a largely unbanked population, generating random variation in savings behavior. We provide evidence showing that financial access improved households' perceived ability to manage their (financial) lives, and it increased the schooling level of daughters and the educational aspirations and expectations parents have for them. We reject the hypotheses that wealth and women empowerment are behind our findings. Our results appear to be consistent with the hypothesis that offering the savings account is helping households overcome a number of scarcity issues. This suggests that the cognitive resources parents once spent fretting over budgetary needs are now partly freed up, and they can use them to focus on their daughters' current and future education.

## Session 7

### **Health, Nutrition, and Education: Evidence from Field Experiments in the People's Republic of China**

*Chair: Linxiu Zhang, Associate Director, Center for Chinese Agricultural Policy (CCAP) at the Chinese Academy of Sciences; co-Director, Rural Education Action Program (REAP)*

Though rural students in the People's Republic of China face many educational disadvantages due to a lack of resources, an equally serious obstacle in their academic performance may not be as immediately apparent: their health. Surveys conducted by the Rural Education Action Program (REAP) have shown that a large fraction of schoolchildren in the People's Republic of China are affected by basic health problems for which simple, cost-effective solutions are available. For example, between 25% and 40% of elementary school students in rural People's Republic of China suffer from nutritional deficiencies, over 10% have uncorrected vision and nearly 40% are infected with intestinal worms. All of these conditions can have significant negative impacts on children's learning, putting them at a disadvantage compared to their richer, urban counterparts. These health problems are all easily treated, but are often overlooked in resource-constrained communities. REAP aims to identify health care delivery solutions that can help give these children a better future.

All four papers included in this session present evidence on high prevalence of treatable health conditions among rural Chinese children, and three include evaluations of interventions designed to reduce prevalence and improve cognitive performance. The first paper analyzes a randomized controlled trial conducted in rural Guizhou province, looking at the impact on school performance and cognitive function of interventions designed to reduce infection with intestinal worms. The second paper evaluates the impact of an incentive scheme introduced to school principals to encourage them to improve student nutrition in their schools. The third paper presents baseline findings on the prevalence of under nutrition among Chinese babies living in rural parts of Shaanxi province, focusing especially on the behavioral correlates of iron-deficiency anemia. The fourth and final paper presents results from a

randomized controlled trial conducted in rural Shaanxi province that aimed to measure the impact on cognitive development of a nutritional intervention targeting rural babies and their caregivers.

### **Breaking the Cycle of Infection: An Impact Evaluation of Strategies to Control Intestinal Worms and Improve Human Capital in Rural People's Republic of China**

*Chengfang Liu, Chinese Academy of Sciences, People's Republic of China*

Empirical evidence suggests that the prevalence of soil-transmitted helminth (STH) infections in remote and poor rural areas is still high among children, the most vulnerable to infection. There is concern that STH infections may detrimentally affect children's healthy development, including their cognitive ability, nutritional status, and school performance. In this paper, we examine the causal relationship between STH infections and developmental outcomes in 2,180 school-aged children in seven nationally-designated poverty counties in rural People's Republic of China. We conducted a large-scale randomized controlled trial in Guizhou province in southwest People's Republic of China between May, 2013 and May, 2014. We use data collected as part of this study to evaluate the impact of a school-based deworming and health education program on children's academic performance, executive function (short-term memory and processing speed), and several health indicators (linear growth and hemoglobin levels). We find that overall adherence to the deworming program is low, indicating a need for better institutional support of STH eradication goals. In our intent to treat (ITT) analysis, we find no significant differences between the control and intervention groups in terms of academic performance, executive function, or either of the two health indicators; however, the effect of the treatment on the treated is positive and significant for both executive function and linear growth, indicating a causal link between STH infection and these two child development outcomes.

### **Breastfeeding and Iron Deficiency Anemia: Correlates of Under Nutrition from Rural People's Republic of China**

*Ai Yue, Shaanxi Normal University, People's Republic of China*

Quality nutrition in the first 1000 days of life has been shown to have large and significant effects on health and development throughout childhood and even into adulthood. Despite the widespread understanding of this critical window, and the long-term consequences of leaving nutritional deficiencies unaddressed, little is known about the status of infant nutrition in rural People's Republic of China, or about the feeding practices that might contribute to this status. In this paper, we present data from a survey of 1800 rural Chinese households with babies aged 6–12 months. We show that anemia rates are high, at 48.8%. Rates of breastfeeding are high, and are positively and significantly correlated with anemia status. Formula feeding, by contrast, is negatively and significantly correlated with anemia status.

### **Nourishing the Future: Targeting Infants and Their Caregivers to Reduce under Nutrition in Rural People's Republic of China**

*Renfu Luo, Chinese Academy of Sciences, People's Republic of China*

Research increasingly indicates the importance of the nutritional programming that occurs in the first two to three years of life. Quality nutrition during this brief window has been shown to have large and significant effects on health and development throughout childhood and even into adulthood. Despite the widespread understanding of this critical window, and the long-term consequences of leaving nutritional deficiencies unaddressed, little is known about the status of infant nutrition in rural People's Republic of China, or about the relationship between infant nutrition and cognitive development in rural People's Republic of China. Between April, 2013 and April, 2014, we conducted a cluster randomized

controlled trial designed to measure the impact on child health and development of a nutritional intervention that provided caregivers with children aged 6–12 months at the start of the study with one-on-one nutritional training as well as a 12-month supply of micronutrient supplement packets. We find that although caregivers who received the intervention were irregular in their administration of the packets to their children, overall anemia rates still experienced a significant decrease relative to the control group. By contrast, we find that children’s age-appropriate cognitive development did not improve over the course of the study, suggesting that there are other factors besides nutrition that are slowing rural children’s cognitive development.

## Session 8

### Impact of Water and Sanitation Interventions

*Chair: Amy Leung, Director, Urban Development and Water Division, SERD, ADB and Chair of Water CoP*

#### **Water Quality and Education in a Brawn-Based Economy: The Rural Drinking Water program in the People’s Republic of China**

*Lixin Colin Xu, World Bank and Jing Zhang, Renmin University of the People’s Republic of China*

While previous research has demonstrated health benefits of safe drinking water programs, we know relatively little about the long-term benefits such as its effect on education. In this paper we examine the educational benefits on rural youth in the People’s Republic of China of a major drinking water program started in the 1980’s, one of the largest of such programs in the world, costing more than \$8.8 billion and covering around 300 million people. By employing a longitudinal dataset with more than 10,000 individuals between 16 and 25 years old, we find that this health program has improved their education substantially: increasing the grades of education completed by one years and the probabilities of graduation from a middle and high school by around 22% and 94%, respectively. The qualitative results hold when we control for local educational policies and resources (by including county-year fixed effects), the village dummies, the distance of villages to schools, and by instrumenting the water treatment dummy with its topographic feature, among others. Moreover, the female benefit much more than the male in terms of schooling attainment, consistent with the brawn-based theory of gender division of labor in Pitt et al. (2012). The program can completely account for the gender gap in educational attainment in rural People’s Republic of China in the sample period. Our estimates suggest that this program is highly cost-effective.

#### **Reducing Early Childhood Diarrhea: Enhanced Own-Learning and Technology Adoption**

*Agha Ali Akram, Yale University*

Childhood diarrhea is a major global health crisis and claims the lives of close to a million children every year. Although effective mechanisms such as chlorine tablets (to chlorinate drinking water) exist to reduce childhood diarrhea, there has been a consistent problem in health experiments of maintaining participation amongst poor households. Even when the technology is provided free, households stop using it soon after commencement. We test the hypothesis that participants are not able to learn about the efficacy of chlorine tablets because they do not get clear signals on its effect. To this end, we conducted a small field experiment in a poor urban setting (Karachi, Pakistan). The experiment had a simple structure with a control and treatment arm, where both arms received freely delivered chlorine. The major difference between the two was that the treatment arm received a visual tool (Info-Tool)



which allowed them to track their diarrhea levels in reference to a norm. They used this tool before receiving tablets and a few months into receiving tablets, which allowed them to better detect the efficacy of the tablets. Participation rates were significantly and persistently (very long term) higher in the group that received Info-Tool. The results suggest that providing households with clearer signals on the effect of the tablets made the intervention far more successful. This provides a basis for better structured social policy and interventions that look to provide recipients clearer signals on how they benefit so that their decision to participate is more rational.

## Session 9

### Self-Help Groups (SHG), Empowerment and Poverty Reduction in Rural India

*Chair: Leena Sushant, Breakthrough, India*

#### Impacts of the Andhra Pradesh Rural Poverty Reduction Program

*G. Prensushi and A. Gupta, World Bank*

The paper explores whether one of the largest programs in the world for women's empowerment and rural livelihoods, the Indira Kranti Patham in Andhra Pradesh, India, has had an impact on the economic and social well-being of households that participate in the program. The analysis uses panel data for 4,250 households from two rounds of a survey conducted in 2004 and 2008 in five districts. Propensity score matching was used to construct control groups and outcomes are compared with differences-in-differences. There are two major impacts. First, the Indira Kranti Patham program increased participants' access to loans, which allowed them to accumulate some assets (livestock and durables for the poorest and nonfarm assets for the poor), invest in education, and increase total expenditures (for the poorest and poor). Women who participated in the program had more freedom to go places and were less afraid to disagree with their husbands; the women participated more in village meetings and their children were slightly more likely to attend school. Consistent with the emphasis of the program on the poor, the impacts were stronger across the board for the poorest and poor participants and were more pronounced for long-term Scheduled Tribe participants. No significant differences are found between participants and nonparticipants in some maternal and child health indicators. Second, program participants were significantly more likely to benefit from various targeted government programs, most important the National Rural Employment Guarantee Scheme, but also midday meals in schools, hostels, and housing programs. This was an important way in which the program contributed to the improved well-being of program participants. The effects captured by the analysis accrue to program participants over and above those that may accrue to all households in program villages.

#### The Effects of Economic Self-Help Group Programs on Women's Empowerment: A Systematic Review

*Thomas de Hoop, American Institutes for Research, Washington D.C., USA*

This presentation discusses the effects of self-help groups (SHGs) on women's empowerment using an integrated mixed-methods systematic review to benefit from data generated through quantitative and qualitative research. SHGs are programs where female participants come together and receive a collective finance, enterprise and/or livelihoods group intervention. Nowadays, SHGs are the most popular program to stimulate women's empowerment in South Asia. The theory of change underlying SHGs indicates that SHGs influence empowerment by exposing women to group support and by enabling women to gain access to savings, credit, and training.

For our quantitative synthesis we found 23 experimental and quasi-experimental studies that matched our inclusion criteria. We statistically combined the effect sizes and associated standard errors of 16 studies in a meta-analysis and found small but positive effects ranging from 0.13–0.41 standard deviations on women’s economic, social, psychological, and political empowerment. We found these effects after excluding studies with a high risk of selection-bias, which we determined based on a comprehensive risk of bias assessment. Our meta-analysis suggests that studies with a high risk of selection-bias overestimate the impact of SHGs.

For our qualitative synthesis we found 8 studies that matched our inclusion criteria. The studies emphasize women’s perspectives on empowerment trajectories in SHGs. The evidence indicates that SHGs are more effective when the number of members is between 10 and 20, when the group is formed with assistance from the government or a community-based organization, when there is a strong training component, and when there is strong group leadership.

**Impact Assessment of SHG Bank Linkage Program: A Case of Indian Economy**

*Sudipta De, Assistant Professor Sivanath Sastri College under University of Calcutta Kolkata West Bengal, India*

The impact of participation in a micro- credit program has featured prominently in recent research studies and has become an increasingly important aspect of development, as agencies and particularly aid donors have sought to ensure that funds are well spent. Microfinance has drawn millions of women into commercial economic activities through self-help groups (SHGs) to enable them to earn and control their own income and thus women gain decision making power and better status. As to Indian context, women under the rural industrial set-up were not able to actively participate in income generating economic activities owing to historical and socio-cultural reasons. This study seeks to measure the level of empowerment of women participating in the micro credit program through SHGs and to examine the impact of empowerment of women on their children’s nutritional status and protein intake of their households.

The study is based on the data obtained from field survey in West Bengal. The methodologies that have been used in the study are propensity score matching, simple index, regression and anthropometric measures etc. The study suggests that women participating in the microcredit program through SHGs have enjoyed higher level of empowerment than that of control group. Moreover, there is a better impact on the children’s health and nutritional status as well as on the protein intake of households of women hold higher level of empowerment.

Session 10

**REDD-y and Warned for Impact Evaluation: Making the Case for Evaluating Climate Change Initiatives in REDD/REDD+ Projects and Early Warning Systems**

*Chair: Hanna Uusimaa, Climate Change Specialist, PARD, ADB*

Pamela Jagger, Assistant Professor, University of North Carolina, Chapel Hill  
 Jesse Ribot, Department of Geography, University of Illinois  
 Dennis Eucker, GIZ/CIM Integrated Expert for Adaptation to Climate Change, Mozambique  
 Gunther Bensch, Researcher, Department of Environment and Resources, RWI  
 Discussant: Jyotsna Puri, 3ie

This panel will feature four presentations that discuss impact evaluations in climate change and disaster risk reduction. Two presentations will focus on REDD+ interventions and two others on early warning systems in Africa. Research team members will present the intervention, study design, and anticipated challenges with evaluating impact for these types of programs with inputs from implementing agency staff. The teams will also report on lessons learnt so far through preparatory research, including political, logistical, and evaluative challenges that will need to be addressed in constructing robust impact evaluations. The presentation from each country case study will last 15 minutes. These will be followed by reactions from two discussants and a moderated discussion.

## Session 11

### Better Data for Better Impact Evaluations

*Chair: Bernard Woods, Principal Results Management Specialist, SPD, ADB*

#### **LISA 2020: Building Statistical Collaboration Capacity in Developing Countries**

*Eric Vance, Department of Statistics, Virginia Tech, USA*

Impact evaluations can play a significant role in evidence-based decision making, however, around the world, and especially in developing countries, many institutions lack the statistical capacity to design, implement, and analyze quality evaluations. LISA—The Laboratory for Interdisciplinary Statistical Analysis at Virginia Tech—aims to meet this challenge by educating and training statisticians from developing countries to become interdisciplinary collaborators. Since 2009, LISA has collaborated on several impact evaluations in Africa that laid the foundation for “LISA 2020”—an ambitious program to partner with universities, institutions, and individuals around the world to create a network of 20 new statistical collaboration laboratories in developing countries by 2020. LISA 2020 features training in the United States coupled with ongoing support to the host institution to create a sustainable statistical collaboration laboratory that works with local researchers and policy makers and trains the next generation of collaborative statisticians. The training program helps statisticians move from theory to practice and collaborate with researchers and policy makers to design studies and experiments; collect, analyze, and plot data; run statistical software; interpret results; and communicate statistical concepts and results to non-statisticians. So far, statistical collaboration laboratories in the LISA 2020 network are being created in Nigeria, Tanzania, and Brazil. The LISA approach unlocks the collaborative potential of statisticians to work with their collaborators and teach other statisticians to do the same. Opening these collaborations to the power of statistical thinking will be key to improving human welfare worldwide.

#### **Introducing the Field Statistician: How Integrating a Statistics Expert into the Evaluation Process Improves Results**

*Marcos Carzolio, Virginia Tech*

Dataset cleaning for household surveys in impact evaluations is typically done after fieldwork is completed, when suspicious outliers cannot be verified. This often leads to unreliable data and misleading conclusions. Complex surveys performed in the multi-lingual setting of developing countries often have magnified enumerator and data entry errors. Our approach to meeting the challenge of collecting high-quality data to support impact evaluations and other research is to introduce a field statistician who is involved in the research process from beginning to end. For the research, this ensures dependable data and sound analyses. This talk will cover the challenges encountered and lessons learned while working on-site for an impact evaluation in Mozambique, which seeks to quantify the socioeconomic impacts



of installing hand pumps in rural Mozambican villages. Implementation of this survey includes having a statistician personally train enumerators and clean data while in the field. Involving a statistician at every step of the process— from research question formulation and study design to enumerator training, data collection, cleaning, analysis, and report writing— reduces the chances of having an inconclusive or erroneous study due to experimental error or poor modeling. Ultimately, we find that embedding a statistician in the field during data collection produces a clean and reliable dataset for better analyses and policy-making.

**The Challenge of Defining and Measuring Program Participation: Experience from Multipronged, Multi-Partner HIV Prevention Evaluation in Central America**

*Rebecca Firestone, PSI; Susana Lungo, PASMO PSI; Jorge Rivas, PSI; Jennifer Wheeler, Webster Family Medicine; Alejandra Cabrero and Lung Vu, Research & Metrics, PSI*

The Pan-American Social Marketing Organization (PASMO) and partners are implementing a 5-year HIV combination prevention program across seven countries in Central America. Combination prevention requires linking behavioral, biomedical and structural interventions to address multiple channels by which individuals may acquire HIV and the social environments that increase vulnerability. In this program, PASMO has created an integrated network of public, private and NGO service providers and outreach teams to serve key populations with health promotion, medical services, and social and legal support. Clients enter the network through program outreach and are referred to various service providers according to their needs, using referral protocols designed to ensure that clients receive all components of combination prevention.

Several calls have been made for rigorous evaluations of combination prevention (Kurth et al. 2011; Padian et al. 2011) to determine the most effective combinations of services, appropriate measures of impact, and valid counterfactuals. However, a core challenge is to measure beneficiaries’ receipt of individual and combined program components.

We discuss how we designed a passive client monitoring system based on unique identifier codes to track how clients enter the service network and the services they receive. We review how we linked this monitoring system to survey data to assess 1) program reach into the target population, 2) variations in program participation according to data source, and 3) the effectiveness of combination prevention for reducing HIV risk behaviors. We discuss the challenges of using this approach and methods for strengthening future evaluations that improve practice.

References

Kurth AE, Celum C, Baeten JM, Vermund SH, Wasserheit JN. Combination HIV prevention: significance, challenges, and opportunities. *Current HIV/AIDS Reports*. 2011;8(1):62–72.  
 Padian NS, McCoy SI, Manian S, Wilson D, Schwartlander B, Bertozzi SM. Evaluation of large-scale combination HIV prevention programs: essential issues. *JAIDS*. 2011. p. E23–E8.

**Role of Technology in Impact Evaluation**

*Manju Bajjiya, Shruti Ahuja, Shashank Batra*

- I. Impact evaluation framework/overview
  - a. Reporting on inputs & outputs, monitoring of processes, evaluation of outcomes and impact, and decision making all depends on the availability of reliable and timely data
- II. OpASHA’s rigorous use of data & technology
  - a. eCompliance: mobile biometric technology that identifies patients by their unique fingerprints and compiles patient data

## MAKING IMPACT EVALUATION MATTER

- i. Introduction
- ii. Use in reporting/MIS
- iii. Use in program monitoring and process evaluation
- iv. Use in impact evaluation
- b. Role and benefits of technology in M&E
  - i. Collection of real time data from the point of collection
  - ii. Built-in logical flow and validation checks that improve data quality
  - iii. Ensures data reliability, validity, integrity, accuracy, timeliness, security
  - iv. Electronic data aggregation and analysis
  - v. Improved transparency and accountability/allows for better monitoring, sharing, and application of data
  - vi. Cost-effectiveness
  - vii. Ability to collect new types of data, i.e. location (GIS)
  - viii. Reduction of intermediate levels of data transmission
  - ix. Assists in assessing performance and making evidence-based decisions
- c. Integration of data into program/operational decision making processes
  - i. Example: use of eCompliance data in designing payment structures
- d. Challenges with mobile data collection
- III. Impact evaluations with J-PAL
  - a. Study #1: Evaluation of effectiveness of incentives on efficiency of healthcare providers
  - b. Study #2: Evaluation of effectiveness of eCompliance in improving adherence to TB treatment
  - c. Findings
  - d. Impact of RCTs on policy

## Session 12

### Microfinance in South Asia

*Chair: Noritaka Akamatsu, OREI Deputy Head and Chair of ADB's Finance Sector CoP*

#### Evaluating the Returns to Rural Banking: Village and Household Evidence from Southern India

*Isabelle Cohen, Researcher, Center for Micro Finance*

The argument that the provision and expansion of financial services to rural and underserved populations can reduce poverty is rooted in the belief that financial services improve the ability of poor households to develop income-generating activities and cope with economic shocks. Whether and how this occurs is a long-relevant question for policymakers, which has nonetheless grown in importance over the last decades.

This presentation will draw from on-going work by Professors Rohini Pande and Erica Field in Southern India, which rigorously evaluates the impact of the provision of targeted formal banking services for rural households. The study uses a randomized control trial framework to study the impact of new village bank branches, which provide a broad complement of formal banking services in a rural, financially marginalized setting. I will present preliminary results from the first stage of the study, and focus specifically on how formal financial inclusion impacts informal financial arrangements. Understanding the holistic impact of formal financial services on a household's entire financial services portfolio is

crucial to second-stage impacts in areas like agriculture and health. Our results, based on the provision of financial services to a subsample of our full study for a period of three years, suggest that access to banking services may transform both the formal and informal financial landscapes for these households.

### **Impact of Seasonality-Adjusted Flexible Microcredit on Repayment and Food Consumption: Experimental Evidence from Rural Bangladesh**

*Abu S. Shonchoy, Institute of Developing Economies (IDE-JETRO) and University of Tokyo.*

The mismatch between credit repayments and income seasonality poses a challenge for microfinance institutions (MFIs) working in developing countries. For instance, in northern Bangladesh, income and consumption downfalls during the lean season after the transplanting of major paddy crops are a serious threat to a household's economy. Poor landless agricultural wage laborers suffer the most owing to this seasonality as they face difficulties in smoothing their consumption. However, in designing microcredit products, MFIs do not usually provide flexibility or seasonal adjustment during the lean season. This is mainly because MFIs are afraid that such flexibility might break the repayment discipline of borrowers, resulting in higher default rates. We thus conducted a randomized controlled trial in 2011–12 in northern Bangladesh to empirically test whether seasonality-adjusted flexible microcredit leads to an increase in repayment problems for MFIs as well as whether it can increase and stabilize consumption of borrower households. Our results suggest no statistically discernible difference among the treatment arms in case of default, overdue amount, or repayment frequency. On the other hand, we find no positive impact of repayment flexibility on immediate food consumption during the period of seasonality, except for in-kind full moratorium treatment group. After a year of initial intervention, however, we see positive changes in food intake during the lean season. Thus, our preliminary results are in favor of seasonality-adjusted flexible microcredit.

### **Banking the Poor via Savings Accounts: Evidence from a Field Experiment**

*Silvia Prina, Case Western Reserve University, Ohio, US*

The majority of the poor lack access to bank accounts and have to use costly informal savings mechanisms. Using a field experiment, I randomly gave access to bank accounts with no fees at local bank branches to a large sample of female household heads in Nepal. Results show that there is untapped demand for savings accounts and that the poor do save, despite the lack of commitments. Access to the savings accounts increased monetary assets and total assets without crowding out other kinds of assets or savings institutions. Finally, financial access increased households' investments in education and reduced income volatility.

## Session 13

### **Tax, Property Rights, and Crime**

*Chair: Thabisile Zuma, DPME, the Presidency, South Africa*

### **The Impact of Property Rights on Income and Economic Activity: Evidence from a Field Experiment in Mongolia**

*Daniel Rubenson, Ryerson University, Toronto, Canada*

This paper reports early findings from an impact evaluation of the Millennium Challenge Corporation funded Peri-Urban Rangeland Project (PURP) in Mongolia. The PURP consisted of randomly assigning property rights—in the form exclusive use rights to large plots of grazing land—to herder groups in peri-

urban areas of three Mongolian provinces. The aim of the project was to reduce rangeland degradation by changing herder practices and, ultimately, to increase incomes through improvements to land and animal quality and productivity. Baseline data was collected in winter/spring 2012 from roughly 1000 households in treatment and control. A first follow-up was conducted in spring 2014. We report on project impacts on a range of outcomes including income, household level economic activity, land tenure security, and herding practices. We also study spillovers of project effects by analyzing data from a sample of 400 households neighboring treatment and control households. Our findings are discussed in light of the literature on property rights and micro economic activity at the household level as well as in terms of the project’s relevance for understanding how large scale institutional change (in this case land reform) can affect poverty.

We embedded a series of behavioral economics games in a large scale randomized impact evaluation of a Millennium Challenge Corporation project in which the treatment consists of randomly assigning property rights—in the form exclusive use rights to large plots of grazing land—to herder groups in peri-urban areas of three Mongolian provinces. The paper tests some of the central causal claims in the literature on the development of pro social behavior. According to this literature, social preferences and the cooperative behaviors they produce reflect long term, transmitted cultural norms that have coevolved with economic institutions. Given this, these preferences should not be susceptible to short-term changes in economic institutions, even large-scale changes. However, as we demonstrate, there is some evidence that they are in fact affected even the short term as a result of changes in property rights institutions. We also find support for our theory on trust. Generalized trust is higher among those in the treatment group. Herders who were exogenously placed in a condition of property rights have odds of being trusting some 30% higher than those who are in the status quo common use land condition. We discuss these findings in light of the literature on institutions and norms in development as well as in light of the policy and practice of large scale development projects more generally.

### **Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors**

*Adnan Q. Khan, LSE; Asim I. Khwaja, Harvard; and Benjamin A. Olken, MIT*

Despite the importance of performance-based payment systems, there is little rigorous evidence on their impact in a civil service. In tax, high-powered incentives for tax collectors could increase revenues, but come at a high political cost if collectors exert excessive pressure on taxpayers. We report the results of the first large-scale, randomized field experiment designed to investigate these issues. Working with the entire property tax department in Punjab, Pakistan, we experimentally allocated collectors in the entire 482 property tax units into one of three performance-pay schemes or a control group. We find that incentivized units experience an average of 9 percentage points higher revenue growth than controls, and the effects persist over the two years of the experiment. Yet these units do not report greater taxpayer discontent.

The performance pay scheme that rewarded purely on revenue collection does better, increasing revenue growth by 14 percentage points, with relatively little penalty for customer satisfaction and assessment accuracy compared to the two other schemes that explicitly rewarded these dimensions. Digging deeper into our findings reveals an interesting heterogeneity: most taxpayers in incentivized areas do not get reassessed or pay higher taxes, but do report higher level of bribes and instances of corruption. In contrast, conditional on their property being reassessed, taxpayers in incentivized areas pay higher tax and report less corruption. These results are consistent with a simple model of collusion between collector and taxpayer, in which taxpayers in incentivized areas either have to pay higher bribes to avoid being reassessed, or pay substantially higher taxes if the bargaining process breaks down. Our results suggest that while incentives for tax collectors can substantially boost revenue collected and lead to little overall discontent, they do empower the tax collector in a manner that can both lead to increased revenue and rents at the expense of the taxpayer.

## Nexus of Education, Social Protection and Crime: A Systematic Review

*Paul Omondi, ATI-Network*

Crime reduction has increasingly received much attention on the policy agenda and academic literature. Much discussion of insecurity is dominated by arguments that crime is generally influenced by economic conditions. In this context, poverty is more broadly associated with the likelihood of involvement in crime. Nonetheless, this way of considering possible links between poverty and crime can be ambiguous. Accordingly, one main theory that has been put forward in an attempt to explain the link between income inequality and crime is that of “relative deprivation.” The important argument is that inequality breeds social tensions as the less well-off feel dispossessed when compared with wealthier people. The main aim of this paper is to demonstrate the nexus of education, social protection and crime using various data sources from Kenya. The connection between education and social protection contributes to a richer understanding of the dimensions and perceptions of inequality. The core assumption in this paper is that education generates expectations, which if they are not met influences feelings of deprivation and discontent. Vulnerability is a reflection of the absence of social protection policies which can provide some degree of protection against economic shocks. A systematic review is used in examination of the results of prior research and evaluation studies. New data base will be developed from the existing data sources and used to generate estimates of the causality. The main question in this review is basically whether existing evidence suggests that education increases crime in the absence of social protection.

## Session 14

### Measuring Direct and Indirect Impacts of Policy and Program Interventions

*Chair: Jyotsna Puri, 3ie*

There will be four consecutive papers presented in this session to discuss different methodological approaches for impact evaluation. Policy and program interventions often result in direct impacts on the affected beneficiaries (e.g. due to changes in productivity, income and consumption) but also indirect effects transmitted through changes in prices and wages which may amplify or offset the direct impacts. Many empirical impact studies often attempt to measure the direct effects and fail to provide any estimates on the indirect effect and the net impacts of policy or program interventions. The indirect impacts result from general equilibrium and economy-wide impacts of interventions that affect inter-related sectors in the local economies. In this case, commodity and factor prices are endogenous and can be affected by the program, which could in turn affect income and consumption patterns. The proposed workshop will therefore present a pluralistic methodological approach which combines econometric analysis using panel data as well as partial and general equilibrium modeling using CGE and micro-simulation approaches to measure these effects. It also features innovative approaches for measuring multi-dimensional poverty outcomes using local welfare monitoring systems which can be used to measure the impacts of policy changes and interventions in multiple dimensions using overlapping deprivation analysis.

### Evaluating the Impact of Technological Change on Farm Household Welfare and Poverty: Econometric Approach

*Bekele Shiferaw, PEP-Global office, Nairobi, Kenya*

This paper will present various econometric methods used for measuring the impacts of technological change and impacts on poverty, food security and asset creation. The first involves econometric

estimation using panel data to assess the impacts of technology adoption in Malawian agriculture. It uses a nationally representative panel data and a control function approach and IV models to account for possible endogeneity of input subsidy receipt and improved maize planting decision. The study finds that while access to subsidized input did not affect the likelihood of modern maize planting, it has significant influence on the amount of improved maize planted. Adoption of modern maize technologies is positively correlated with consumption, income and asset holdings of households. A 1% increase in the area planted to modern varieties is associated with a 0.48% increase in income, a 0.34% increase in the maize available for consumption and a 0.24% increase in asset wealth.

The second case study applies a quasi-experimental approach based on Propensity Score Matching (PSM) to evaluate the impact of agricultural technologies (groundnut varieties) on income and poverty in Uganda. The impact of technology must be isolated from that of observed and unobserved factors that simultaneously determine crop income and poverty status. Using the PSM methods and survey data from Uganda, the study finds that adoption of new agricultural technologies can significantly increase crop income and reduce rural poverty. Improved technologies have the potential to increase crop income in the range of \$183 to \$295 and to decrease the poverty incidence by 6%–11%.

The third case study presents an endogenous treatment effects model using an endogenous switching regression treatment effects to evaluate the impact of agricultural technologies (wheat) on food security and welfare in Ethiopia. It employs a large and unique dataset of over 2000 farm households to estimate the model complemented with PSM methodology to test robustness and reduced selection bias stemming from both observed and unobserved characteristics. The analysis expands this further with the generalized propensity score (GPS) approach to evaluate the effects of continuous treatment on multidimensional outcome variables. The study finds consistent results across models indicating that technological change increases food security and welfare. The actual effect that adopters experience is about ETB976 and 2.7% higher for food consumption expenditure and food security outcome variables, respectively. This provides evidence that policies that enhance access to modern agricultural innovations and efforts to improve access to modern technologies and services can contribute significantly to poverty reduction and food insecurity.

### **Evaluating the Poverty and Welfare Impacts of Agricultural Investment Policies in Ethiopia—Economy-Wide Modeling Approach**

*Lulit Mitik Beyene, PEP*

Agriculture remains the key sector for sustained growth and pro-poor development in Ethiopia. However, the sector under utilizes its irrigation capacities as well as its abundant human resources. This paper aims at measuring the impact of public investment in small-scale irrigation and training for farmers on growth and agriculture-led development, on food security, and on poverty in Ethiopia. It is line with the current five year development strategy of the government and will give insights on the effect of selected targeted indicators. We use a dynamic Computable General Equilibrium (CGE) model to capture the outcomes of agricultural investment shocks. Public investment in agriculture increases the supply of skilled agricultural labor and irrigated land by transforming unskilled labor and non-irrigated land. Two types of technologies are utilized in agriculture to produce the same crop: a more productive technology that is intensive in skilled labor and irrigated land and a less productive technology that is intensive in unskilled labor and non-irrigated land. Households have the ability to increase their endowments in labor and land. Hence, the increase in skilled labor due to investment in the form of short term training enables households to increase the share of skilled labor they detain while reducing the share of unskilled labor. The same applies for land. Finally, the model has a poverty module using a top-down approach where changes in the CGE model are imported in the household data. The CGE model is a PEP type model and is calibrated to a SAM of Ethiopia for the fiscal year 2005/06. The poverty module uses the most recent Household Income and Expenditure Survey.



This study supported by PEP showed that the Ethiopian government policy strategy regarding agriculture sector development has a great potential for reducing poverty and food insecurity. Simulation results show that investing in training and irrigation contributes to the effort towards achieving the MDGs. Exports expand and in particular export of high-value crops increasing income at household and national levels. The results also show that an agriculture-led development is less likely to occur because of weak forward and backward production linkages between agriculture and manufacturing sectors where a great deal of manufacturing inputs are imported. The increment in public investment has a crowding-out effect indicating that financing such investment plans may require an alternative allocation of public resources or even a different financing mechanism.

### **Impacts of Cambodia's Tariff Reduction on Households' Welfare and Labor Market: A CGE Analysis**

*Senh Senghor, Cambodia Economic Association*

This study supported by PEP analyzes the welfare and labor market impacts on Cambodia's provincial households from unilateral tariff removal using Computable General Equilibrium (CGE) calibrated to the Cambodia's Social Accounting Matrix (SAM) built for year 2011. Following tariff removal with indirect tax compensation, certain manufacturing industries are a beneficiary relative to agriculture and service sectors. Those manufacturing sectors include textiles, basic metal, fabricated metal, machinery, office and computing machinery. Cambodia's textile industry remains buoyant with the increase of both domestic output and export. This industry will continue to be a backbone for the growth and job employment for a short and medium term.

In terms of effect on labor market, high-skilled and median-skilled workers are relatively beneficial from the tariff elimination. At the household group level, the magnitude impacts on their income and consumption are significantly almost the same with Phnom Penh household seems to be negatively less affected by the indirect tax increase. Generally, welfare gains for the whole country and most household groups are positively small. Households expecting to have welfare losses are Kampot, Kratie, Preah Vihear, Rattanakiri, and Stung Treng. On-job training program, vocational program, and other employment assistance programs are the safety net for the low-skilled workers to mitigate the negative effect of trade liberalization measures.

### **Local Welfare Monitoring Systems for Measuring Multidimensional Poverty Impacts of Programs and Policy Changes: CBMS Approach**

*Celia Reyes, PEP-Asia office, Manila, Philippines*

Based on a local multi-dimensional welfare monitoring using selected indicators, CBMS generates data on household and individual socioeconomic characteristics, demography, access to infrastructure and social programs, income sources, agriculture, migration and remittances, and disaster preparedness. CBMS monitors a core set of indicators of multi-dimensional nature of poverty that includes the areas of income and employment, basic education, health and nutrition, water and sanitation, housing, and peace and order. Data from CBMS can be disaggregated across geopolitical levels (e.g. village, city/municipality, and province) and population sub-groups (by age, sex, income-class, or urban/rural). CBMS is being used as a tool for local planning and budgeting, for poverty diagnosis and analysis, for program design and targeting, and for impact-analysis.

Recently, CBMS has developed the Accelerated Poverty Profiling (APP) as one of its innovations to generate relevant data and provide real time information on poverty profiles and program impacts. The APP entails the organized, systematic and efficient combined use of latest information and communication technology tools e.g. tablets, and the standard CBMS instruments for data collection

(core household profile and village questionnaire), processing (STATSIM), poverty mapping and data management instruments. Data collected through CBMS APP is transmitted to and stored in a data portal where registered users are given access depending on levels of authorization for use. The CBMS APP also helps researchers to respond better to the growing demand from various users, particularly local government units, to fast track the generation of data for use in the preparation of development plans and budgets, and use of data for various development programs.

With wider access to and more affordable ICT tools and infrastructures in developing countries, the CBMS APP provided researchers and development planners the option to fast track the generation and access to data which can inform evidence based decision making at local levels. As of date, the CBMS APP is being adopted in the Philippines and Argentina, but ready for piloting across countries.

## Session 15

### The Design and Implementation of WASH Impact Evaluations

*Chair: Ganesh Rauniyar, Principal Evaluation Specialist, IED, ADB*

#### How Do Health Promotional Nudges Work? Evidence from the Dirty Business of Eliminating Open Defecation in Four Countries

*Paul Gertler, University of California Berkley, USA*

Much of health status is driven by individual behavioral choices such as diet, exercise and hygiene. However, successfully promoting healthy behavior has been notoriously difficult to achieve. Despite vast investment in information campaigns, large numbers of people who know better continue to behave in ways that are detrimental to their long-term health such as smoking, open defecation, and not using condoms in risky situations. Investing in health products such as piped water inside the house and toilets can reduce the marginal time and hassle cost of healthy behaviors.

There are two broad categories of interventions designed to improve healthy behavior. The first category subsidizes the purchase of health products that lower the time and hassle cost of healthy behavior (Cohen and Dupas 2010, Cohen et al 2014, Dupas 2014). The second category is health promotion campaigns aimed at behavioral change (Ashraf et al., 2006, Thaler and Sunstein, 2008, Giné et al. 2010, Karlan et al, 2012, Dupas and Robinson 2013). Health promotion campaigns work through nudging increased good behavior conditional on existing health products and convincing households to invest in health products that lower the marginal time and hassle cost of good behavior.

In this paper we report on the effects and mechanisms of health promotion campaigns designed to eliminate open defecation in at-scale randomized field experiments in 4 countries: India, Indonesia, Mali, and Tanzania. Open defecation is thought to be a major cause of the persistent worldwide burden of diarrhea and enteric parasite infection among children < 5 years old (Mara et al., 2010). Open defecation can lead to the fecal contamination of water and food supplies facilitated primarily by flies and to the transmission of soil borne helminthes (Chavasse et al., 1999). Reducing open defecation requires access and use of improved sanitation facilities, which are defined as facilities that prevent human feces from re-entering the environment; in 2010, an estimated 47% of the world population did not have access to onsite improved sanitation facilities (WHO-UNICEF, 2012). All 4 interventions are Community-Led Total Sanitation (CLTS) campaigns that inform, shame, and help individuals and communities to develop plans and commitments to become 100% open defecation free. India additionally provided subsidies for sanitation (toilet or latrine) construction and Tanzania provided supply side assistance such as the availability of trained masons who knew how to construct latrines. The combination of



at-scale randomized field experiments in 4 countries provides not only strong internal validity but also a degree of external validity not seen in most studies.

### **Impact Evaluation of the MCC-Funded Rural Water Supply Program in Nampula, Mozambique**

*Ralph P. Hall, Virginia Tech*

This paper presents the results from an impact evaluation of the Millennium Challenge Corporation (MCC)-funded rural water supply program in the province of Nampula, Mozambique. This program installed over 350 Afridev hand pumps in six districts across the province. In order to assess the impacts of the installed hand pumps on households receiving the intervention, the research design employed a panel survey in treatment and comparison groups. To collect the panel data, a pre-intervention baseline survey was undertaken in 2011 (in 54 communities—27 treatment and 27 comparison) and a post-intervention follow-up study was undertaken in the same communities in 2013. The analysis employed a mixed-effects model to account for correlations within households and communities. Additionally, fixed effects were introduced to correct for within-district trends and trends over time (i.e., from baseline to follow-up). The results show that the rural water supply program had a significant impact on household water consumption, the time spent collecting water, and household perceptions of the impact of water fetching on school attendance. No significant impacts were found on key variables such as household income, expenditure, or number of children with diarrhea, which challenges the stated purpose of the program to increase beneficiary productivity and income. The policy implications of these findings are discussed.

### **Time Matters: Short Term vs. Medium Term Impacts**

*Soumya Balasubramanya, International Water Management Institute (IWMI-CGIAR), India*

Most evaluations occur two years from the end of a program. These evaluations determine whether the program was successful in achieving its goals. Programs influence changes in the decision-making of beneficiaries, either through the introduction of new information or new technology, the uptake of which takes time; consequently impacts also take time to occur. Thus, evaluations in the short run provide a limited understanding of impacts.

To demonstrate this point, we evaluated the impact of a donor-funded countrywide well testing program in Bangladesh, which was implemented to motivate rural households to switch, for drinking water, from high arsenic wells to neighboring low arsenic wells. We studied the well-switching dynamics over time by re-interviewing, in 2008, a randomly selected subset of households in the Arahazar region who had been interviewed in 2005. We found that:<sup>1</sup>

- For households that recalled arsenic information through 2008; compared to two years after the end of the program, twice the number of households had switched after five years.
- However, for households who did not recall test results by 2008; the loss of arsenic knowledge led to staying at unsafe wells and switching from safe wells.
- These results support ongoing testing for arsenic to reinforce beneficial information.

Thus, a longer term evaluation of this program improved understanding of the dynamics of impacts over time, as well as the differences in impacts across subgroups of beneficiaries. From a policy perspective, such analysis is vital for guiding evidence based-policy making, and for gauging success and failure of development projects.

<sup>1</sup> Balasubramanya, S., Pfaff, A., Benneer, L., Tarozzi, A., Ahmed, K.M., van Geen, A. (2014). Evolution of households' responses to the groundwater arsenic crisis in Bangladesh: information on environmental health risks can have increasing behavioral impact over time. *Environment and Development Economics*, Available on CJO 2013 doi:10.1017/S1355770X13000612.

## Session 16

### Health Insurance

*Chair: Karl Hughes, Senior Evaluation Specialist, IED, ADB*

#### **An Impact Evaluation of the Voluntary Student Health Insurance Scheme in Viet Nam**

*Marian Theresia Valera, UPEcon Health Policy Development Program, Philippines*

Viet Nam implemented a nationwide Voluntary Student Health Insurance (VSHI) scheme in 1998 with the aim of strengthening the school health system and reducing the level of health inequity among school-aged children. After more than a decade of its implementation, challenges remain in terms of high out-of-pocket health spending and low access to health services, especially among rural and poor children.

A panel dataset was constructed from a cohort of children from the Young Lives (YL) study in Viet Nam. Demographic and health information were captured using two household survey rounds conducted in 2006 and in 2009. A difference-in-differences (DID) method was used to exhibit that the VSHI scheme is associated with differential improvements in health outcomes among primary school students in Viet Nam.

The scheme was estimated to have contributed to improvements in health-seeking behavior, with an observed 10%–10.8% increase in the probability of a health facility visit. However, the increase in health care utilization translated to an 11.8%–14.2% rise in the probability of incurring OOP expenditures for treatment. Finally, the results suggest a positive association between health insurance coverage and nutrition status, with documented improvements in the nutrition outcomes of VSHI-enrolled students. The results suggest that improving access to a school-based social health insurance program has a positive impact on health indicators among primary school children in Viet Nam. Important policy implications include ensuring continuous coverage and integrating lessons from the implementation of the scheme as the government embarks in its goal to attain universal coverage in Viet Nam.

#### **Out-of-Pocket Payments and Expenditure Reallocation: Experimental Evidence from the Philippines about Health Insurance versus a Provider-Based Performance Scheme**

*Stella Quimbo, UPEcon, Philippines*

Using data from a randomized policy experiment in the Philippines, we find that expanded insurance coverage and improved hospital quality similarly reduce out-of-pocket payments and increase preventive health care spending. The experiment was implemented in 30 district hospitals in the central regions of the Philippines from 2003 to 2008 and consisted of two randomly assigned interventions plus a control group. One intervention offered full insurance support for children under 5 years old, while the other intervention made bonus payments to hospital staff when predetermined quality standards were met. The experiment provides a unique opportunity to contrast a health insurance intervention with a provider-based intervention. Using linear fixed effects and Poisson models, we find evidence of reduced out-of-pocket expenses and total hospital charges in both intervention sites. We found a similar magnitude of financial risk protection provided by expanded health insurance and performance-based provider payments indicating that effective health reforms do not exclusively have to rely on the expansion of insurance. Furthermore, we find that freed-up resources at the household level are used for personal hygiene, water and sanitation.

## Effects of Interventions to Raise Voluntary Enrollment in a Social Health Insurance Scheme: A Cluster Randomized Trial

*Joseph Capuno, University of the Philippines School of Economics, Manila*

A cluster randomized controlled trial was undertaken, testing two sets of interventions to encourage enrollment in the Philippines' Individual Payer Program, a membership program for the self-employed and those in informal sector under the National Health Insurance Program. Of 243 municipalities, 179 were randomly assigned as intervention sites and 64 as controls. In early 2011, 2,950 families were interviewed; unenrolled Individual Payer Program-eligible families in intervention sites were given an information kit and a 50% premium subsidy until the end of 2011. In February 2012, the "non-compliers" had their voucher extended, were re-sent the enrollment kit, and received Short Message Service (SMS) reminders. Half were told that in the upcoming end-line interview the enumerator could help complete the enrollment form, deliver it to the insurer, and have their identification cards mailed to them. The control and intervention sites were balanced at baseline. In the control sites, 9.9% (32/323) of eligible individuals had enrolled by January 2012, compared with 14.9% (119/801) in intervention sites. In the sub-experiment, enrollment was 3.4% (10/290) among eligible non-compliers and who did not receive assistance but 39.7% (124/312) among those who did. A premium subsidy combined with information can increase voluntary enrollment in a social health insurance program, but less than an intervention that reduces the enrollment burden; even that leaves enrollment below 50%.

## Impact of Social Protection Program on Access to Health Care for Informal Sector Workers from the Socially Excluded Groups in India

*Nidhi S. Sabharwal, Vinod Kumar Mishra, and Ajaya Kumar Naik, Indian Institute of Dalit Studies, New Delhi*

It is now being increasingly recognized that out-of-pocket health expenditure is key driver of poverty in India. Millions of households have been pushed into, or deeper into, poverty as a result of catastrophic health expenditure. This often has long-term implications especially when poor households, with few available coping strategies, resort to selling assets or borrowing with high interest rates to meet inpatient health expenses. Weaknesses in the health system, including variable quality of care in both the public and private sectors and a lack of accountability, including corruption, are key factors which have a bearing on equitable access to health services. Healthcare access and outcomes are also highly unequal in India, with marginalized and socially excluded households experiencing poorer health outcomes and specific barriers to access to healthcare. The recent government initiative, RSBY, aims to overcome some of these challenges by providing social health insurance at a subsidized rate to below poverty line households. In this evaluation, we used a mixed methods research approach to assess the effect of RSBY, with a particular focus on Scheduled Caste (SC), Muslim households and Women. We have used a social exclusion analytical framework to understand whether and how RSBY has contributed to increased health access and utilization and reduced out-of-pocket inpatient health expenditure, as well as looking at the indirect effects on the poor in terms of household well-being, economic activities and livelihoods and perceptions about corruption and state-society relations. Propensity Score Matching (PSM) analysis has been done for quasi-experimental impact evaluation which included a treatment and a control group. The data source of the present paper is based on the primary data collected from the field survey in India. The study highlights not only the direct and indirect effects of social health insurance on poor households, but also take a disaggregated examination and analysis of the effects on the poorest and most marginalized/excluded households (in this case focusing on Scheduled Caste and Muslims), and generates a better understanding of the effects of program design and delivery on social and political spheres—in particular community interactions and state-society relations.

## Session 17

### Investing in Smallholder Agriculture

*Chair: David Raitzer, Economist, ERD, ADB*

#### **Public Insurance and Household Investment—Evidence from the Farmer Income Guarantee Program in Thailand**

*Juliane Zenker, University of Göttingen, Germany*

Outline: The Government of Thailand implemented an income guarantee program for rice, tapioca and maize farmers from 2009 to 2011, replacing the previously existing rice pledging scheme. Both programs aimed to insure farmers against crop price shocks.

In the old program the government bought the rice crops from farmers at a previously announced high price (relative to the market price) and resold the produce in the market. In the new program, the government would also agree with the farmers over a certain standard price. However, instead of buying the crops from the farmer and reselling it, the government paid the difference between agreed price and market price directly to the farmer for the entire produce. This way the new program ensures that poorer small farmers who mostly produce for own consumption and do only sell little of their produce, fully profit from the advantages of the price insurance.

We evaluate the impact of the scheme on these small-scale farmers. This is not trivial because people self-select by registration into the insurance scheme. Comparing registered and non-registered farmers would be substantially biased due to self-selection. We calculated the propensity score of registration on a extensive set of covariates and use various matching and weighting techniques to eliminate this bias and obtain credible estimates of the impact of the insurance on the farmers. Our results show, that reducing the risk of low prices through the insurance decreases farmers risk aversion. This reduced risk aversion translates into higher investment in agricultural assets and land area, an important prerequisite for long-term productivity increase in the region.

#### **Impact of Pro-Poor Intervention Programs on Rural Household Welfare in Africa: The Case of Ghana's Productivity Improvement and Employment Generation Fund**

*Victor Owusu and Issahaku Saani, Department of Agricultural Economics, Agribusiness and Extension, Kwame Nkrumah University of Science and Technology, Kumasi Ghana*

Using a cross-sectional data collected on 300 farm households in four rural communities of West Mamprusi Districts in Northern Ghana, this paper examines the impact of the Productivity Improvement and Employment Generation Fund (PIEGF) program initiated by the Government of Ghana on household welfare and food security status. Self-selection bias as a result of non-random assignment of participants into the poverty reduction program is addressed with the Propensity Score Matching (PSM) approach. The empirical results from the study generally indicate robust and positive impacts of the PIEGF program on both household welfare and food security status of the rural households. The estimated average treatment effects (ATT) of household welfare and food security status for females were found to be statistically higher than that of males, indicating gender heterogeneous effects regarding pro-poor intervention programs toward poverty alleviation in Africa.

## Session 18

### Education in Africa

*Chair: Awuor Ponge, African Policy Centre (APC), Kenya*

#### Community-Based Literacy Interventions in Low- and Middle-Income Countries

*Elizabeth Spier, Ph.D.*

American Institutes for Research, Yale University and Search Institute conducted a systematic research review of programs designed to improve children's literacy outcomes in low- and middle-income countries, with a focus on interventions that take place outside of formal education systems. This review was conducted under 3ie, with funding from USAID.

To be included in our review, a study needed to (1) include children ages 3 to 12 years, (2) assess the effectiveness of an intervention delivered through families or community members, (3) include a comparison group, and (4) include an outcome measure of children's pre-literacy or literacy skills. Our initial search yielded 10,430 journal abstracts and other reports. After screening, 53 studies met our criteria for full text review, and of those, just 13 fully met our criteria for inclusion.

Tutoring (by peers or adults), educational television, and parent education programs were effective in improving children's literacy, but effects varied widely based on context. None of the 13 studies included children over age seven, none were from the Americas, and none addressed children with any risk beyond poverty (such as out-of-school children). There was a lack of rigorous studies regarding the effectiveness of widely-used community-based interventions such as libraries, provision of print materials, provision of e-readers or laptops that children use outside of school, reading campaigns, reading clubs, or literacy supported through religious instruction. The identification of such gaps should inform investment in the development of an evidence base, thereby improving the ability of the field to provide effective interventions.

#### Do Community-Managed Schools Facilitate Social Capital Accumulation? RCT-Based Evaluation of COGES Project in Burkina Faso

*Yasuyuki Sawada, University of Tokyo and JICA Research Institute; Takeshi Aida, Graduate Research Institute for Policy Studies, GRIPS; and Eiji Kozuka, Hiroshima University and Japan International Cooperation Agency*

The role of community participation has attracted a significant amount of attention in recent research on basic education in developing countries. However, existing empirical results on the role of community or school based management (SBM) in improving the quality of education are mixed. To overcome remaining issues of the existing studies, in this paper, we investigate the role of SBM in facilitating social capital among parents, community members and teachers. We employ unique data from Burkina Faso where the School Management Committee (COGES) project was introduced recently. Implementing carefully-designed randomized controlled trials in the phase-in timing of the COGES project, we evaluate its policy impact on social capital accumulation rigorously. To elicit the individual level of social capital, we conduct the public goods games, one of the standard artefactual field experiments, with monetary rewards. We found that the COGES project increases the level of social capital in general. The policy effects can be found especially among community members and between parents and teachers. We believe that our study will make contributions in the following two aspects. First, it is very important to understand the process of social capital accumulation in a country like Burkina Faso where market underdevelopment is serious obstacle of economic development. Second, it is practically important to identify the key factors in promoting democratization process in Burkina Faso, a country with the lowest level of political rights and civil liberty, or "weak governance."

## The Mid-Term Impacts of “Girl-friendly” Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso

*Ali Protik, Mathematica Policy Research*

In this presentation, we present the mid-term term impacts of the Burkinabé Response to Improve Girls’ Chances to Succeed (BRIGHT) program, financed by the Millennium Challenge Corporation. BRIGHT sought to increase girls’ primary school enrollment and completion rates in rural Burkina Faso. The program was implemented in 132 villages from the 10 provinces with the lowest rates of girls’ primary school enrollment. Implementation began in 2005 and consisted of the construction of 132 primary schools housing three classrooms for grades 1–3 and developing a set of complementary interventions. The program was extended in 2008 through the construction of three additional classrooms for grades 4–6 in the same 132 schools and the continuation of the complementary interventions. A short-term impact of the BRIGHT programs using data from a 2008 follow up survey of households and schools, after 3 years of exposure, found positive impacts on school enrollment and test scores for both boys and girls.

We conducted a second follow survey in 2011 and analyzed the survey data using a regression discontinuity design to determine the mid-term impacts of the program. Seven years after the start of BRIGHT, the program increased enrollment by 15.5 percentage points and increased test scores by 0.29 standard deviations. Also, BRIGHT schools are effective at getting children start school on time and at keeping children in school longer. Overall, students in BRIGHT schools progress farther through the grades, by about one grade, compared to students in non-BRIGHT schools.

## Impact Evaluation of Niger’s IMAGINE Program: Long-Term Follow-Up

*Emilie Bagby, Mathematica Policy Research*

This presentation will report the long-term impact of an education project called IMAGINE (IMprove the educAtion of Grls in NigEr) in Niger, financed by the Government of the United States (through the Millennium Challenge Corporation). The IMAGINE project was completed in 2009 and included the building of 68 “girl friendly” schools in rural areas of Niger, with amenities such as separate latrines for boys and girls, a water source, and housing for female teachers.

Mathematica completed an impact assessment of these activities in 2011 and found small positive impacts on school enrollment but no impacts on attendance or test scores. The positive impacts are driven entirely by effects of the program on girls. These findings are smaller than expected given that an evaluation of a similar intervention in neighboring Burkina Faso found large impacts. Several factors might help to understand the results from the initial study, including the presence of schools in nearly all recipient villages prior to the program, selection of villages by the central ministry to receive schools without an application process, incomplete implementation of some intervention activities, and measurement of outcomes only one to two years after completion of the project.

This presentation focuses on a follow-up evaluation of the IMAGINE program that will address key questions regarding the sustainability of educational development programs that focus on school infrastructure, four years after the program was implemented. Preliminary results suggest that longer-term impacts are larger than previously observed and remain larger for girls than for boys.



## Session 19

# The Challenge of Evaluating Women's Empowerment in South Asia

*Chair: Thomas de Hoop, AIR, USA*

The solution to achieve gender equality is to empower women. How to do it? The panel will discuss different interventions to promote women's empowerment in South Asia, by

- Providing affordable day-care services and freeing women's time for market activities.
- Organizing women into groups under an existing large government empowerment program; or
- Launching campaigns against early marriage and improving their chances of a better life.

Based on their own experiences, the panel of scholars and practitioners will focus on two key set of questions (i) how to gauge progress in the empowerment of women, especially their economic empowerment; and (ii) how evaluation can help interventions work better to empower women.

### Women's empowerment program

- Jyotsna Jha, Centre for Budget and Public Studies, CBPS—Bangalore

### Day Care program

- Parul Agarwal, Institute for Financial Management Research IFMR, Bangalore

### Nation against Early Marriage Campaign in India

- Leena Sushant, Director—Research, Monitoring and Evaluation, Breakthrough, Delhi, India

### Early marriage in Bangladesh

- Prof. Sarah White, Sociology professor—University of Bath, United Kingdom

### Commentators

- Kavitha D.L., representative of Mahila Samakhya in Karnataka
- Laxmi Thakur, Seva Mandir—Rajasthan
- Shannon Maloney, Institute for Financial Management Research IFMR
- Mallika Samaranayake, Community of Evaluators in South Asia, Colombo

## Affordable Day Care in India: The Case of Rajasthan

The project will track utilization of a day-care program to estimate the effect on mothers and children in rural Rajasthan, a state that performs poorly in terms of women's empowerment in India. A local nongovernment organization (Seva Mandir) will oversee the day-care centers, which provide child-care, nutritional, and educational services. The primary outcomes of the impact evaluation will be how women themselves perceive their empowerment (self-esteem, role in decision-making, freedom of mobility, and control of material resources). Other outcomes include labor market attachment (employment status, occupation, hours worked), economic status (personal and household income), and mental health (depression measured using a diagnostic instrument). Additionally, the project will measure nutritional status (stunting, wasting and underweight) among the children. The costs of the intervention, together with indicators of inputs, outputs, outcomes, and impact will provide both validity and comparability with other relevant research, and facilitate estimation of the cost-effectiveness of our intervention in achieving key outcomes. The project aims to address an important research gap in women's economic empowerment and has the potential to inform policies for improving the day-care system among poor households in India in ways that promote inclusive economic growth.

### **Together We Can: The Impact of Education and Collective Action among Rural Women in India (CBPS, Bangalore, India)**

The research undertakes a randomized experiment to evaluate the impact of program on the labor market and other economic empowerment related outcomes of women in rural areas in three states of India. The project sites include backward areas in the states of Bihar and Karnataka, with conditions similar to those of women residing in the other low-income areas of South Asia and Sub-Saharan Africa. In these areas, the project will randomly assign half of 144 villages to receive the program intervention (treatment) and half to receive it later (control), representing more than 4,000 households. The project will collect household-level data in three different waves. The data will include: (1) labor market outcomes (such as weekly hours worked, employment status, type of employment, etc.); (2) economic status (household income, personal income, wages, savings, assets, debt); (3) gender labor inequalities (such as wage, type of employment, sector of work among others); (4) potential for future labor market outcomes among adult women and among adolescents (school attendance, highest level of education achieved, age of first marriage, and others); and (5) awareness and use of entitlement programs. The proposed research aims to understand to what extent the program has been successful, if at all, in reducing barriers to women's economic empowerment and influence policy solutions. The project will reach government policy planners, program functionaries, as well as academics, media and civil society at large through a wide-range of activities such as workshops with program functionaries at different levels; film screenings at participating districts and other spaces; mass media, social media and blogs; academic papers and conferences.

### **Nation against Early Marriage Campaign (Breakthrough—Delhi)**

Breakthrough's campaign aims to reduce early marriage by using mass- and multimedia interventions and community engagement to challenge prevailing gender norms; change attitudes and practices; and promote the rights and value of girls. The program is being implemented on the ground in three districts in Bihar and Jharkhand and consists of: (1) mass media intervention; (2) training at block level with NGOs, CBOs, self-help groups (SHGs), and youth groups; and (3) community mobilizations. The purpose of the baseline report is to establish the current status of the impact and outcome indicators so that comparisons can be drawn in the future to assess the impact of the program. Most households report that women generally get married between the age of 18–21 years, with the most frequent response 15–17 years. The actual incidence of early marriage was much higher. This indicates that people underestimate the prevalence of early marriage. The outcome indicators related to gender norms show that there are disparities between the two genders in terms of division of household responsibility and access to education and inheritance. Women and girls were primarily responsible for household chores and men and boys were seen as responsible for financial sustainability. Parents wished for their male children to be more educated than their girl children. The largest disparity existed in the case of inheritance where in 95% of the households only sons inherit familial property. 95% of respondents are aware of the negative consequences of early marriage, but awareness of these consequences has not served to reduce the practice. Most households report being aware of negative consequences on health. There is also awareness of domestic violence and disruption in education as negative consequences.

### **Affordable Day Care in India: The Case of Rajasthan (McGill University, Montreal, Canada)**

The project will track utilization of a day-care program to estimate the effect on mothers and children in rural Rajasthan, a state that performs poorly in terms of women's empowerment in India. A local nongovernment organization (Seva Mandir) will oversee the day-care centers, which provide child-care, nutritional, and educational services. The primary outcomes of the impact evaluation will be



how women themselves perceive their empowerment (self-esteem, role in decision-making, freedom of mobility, and control of material resources). Other outcomes include labor market attachment (employment status, occupation, hours worked), economic status (personal and household income), and mental health (depression measured using a diagnostic instrument). Additionally, the project will measure nutritional status (stunting, wasting and underweight) among the children. The costs of the intervention, together with indicators of inputs, outputs, outcomes, and impact will provide both validity and comparability with other relevant research, and facilitate estimation of the cost-effectiveness of our intervention in achieving key outcomes. The project aims to address an important research gap in women's economic empowerment and has the potential to inform policies for improving the day-care system among poor households in India in ways that promote inclusive economic growth.

### **Together We Can: The Impact of Education and Collective Action among Rural Women in India (CBPS, Bangalore, India)**

The research undertakes a randomized experiment to evaluate the impact of program on the labor market and other economic empowerment related outcomes of women in rural areas in three states of India. The project sites include backward areas in the states of Bihar and Karnataka, with conditions similar to those of women residing in the other low-income areas of South Asia and Sub-Saharan Africa. In these areas, the project will randomly assign half of 144 villages to receive the program intervention (treatment) and half to receive it later (control), representing more than 4,000 households. The project will collect household-level data in three different waves. The data will include: (1) labor market outcomes (such as weekly hours worked, employment status, type of employment, etc.); (2) economic status (household income, personal income, wages, savings, assets, debt); (3) gender labor inequalities (such as wage, type of employment, sector of work among others); (4) potential for future labor market outcomes among adult women and among adolescents (school attendance, highest level of education achieved, age of first marriage, and others); and (5) awareness and use of entitlement programs. The proposed research aims to understand to what extent the program has been successful, if at all, in reducing barriers to women's economic empowerment and influence policy solutions. The project will reach government policy planners, program functionaries, as well as academics, media and civil society at large through a wide-range of activities such as workshops with program functionaries at different levels; film screenings at participating districts and other spaces; mass media, social media and blogs; academic papers and conferences.

### **Evaluating Empowerment: An Example from Bangladesh**

*Sarah White, University of Bath*

The central question of this paper is the challenge of evaluating empowerment, which is an essentially qualitative and to some degree subjective concept, in ways that can be recognized as robust within international development programmatic contexts. The paper reflects briefly on the origins of thinking on empowerment within the women's movement, as something that had to be achieved by women organizing themselves, and where this fits with current practice in large international agencies which set themselves to 'empower women'. The paper emphasizes in particular the importance of qualitative approaches to evaluation, and explores the range of methodological options that these offer and the kinds of insights they can give us, and will also discuss the close inter-relationship between the ways we theorize empowerment, the methods of data collection and analysis, and the data that we produce. This discussion will be grounded empirically in discussion of a major evaluation being undertaken at present of a program to empower adolescent girls in Bangladesh. IWhereas there are some clear areas of complementarity between qualitative and quantitative methods in evaluating empowerment, there will remain some underlying differences in worldview, the bases of which are ultimately political, rather than technical.

## Session 20

### Using Impact Evaluation to Assess Agency Performance

*Chair: Hans Carlsson, Advisor and Head, Portfolio, Results and Quality Control Unit, SARD, ADB*

#### Agency-Wide Performance Measurement Systems: The Role of Impact Evaluation

*Howard White, 3ie*

The results agenda has been with us for close to two decades. Agencies are asked to demonstrate that they are achieving their ‘results’: lifting people out of poverty, empowering women, getting children into school and learning when they are there and so on. Many agencies have developed ‘results systems’ based on monitoring outputs and outcomes. But if a result is to mean what difference the agency has made then impact evaluation should play a role. To date it has largely not done so. This paper draws on contributions from a number of government agencies and NGOs which assess their Agency-Wide Performance Measurement System (AWPMS) to develop principles for the use of impact evaluation in assessing agency performance, and for AWPMS more generally.

#### To Feel Good or to Do Good? Why We Need Strong Institutions to Ensure a Result Focus in (Norwegian) Development Assistance

*Vibecke Dixon and Ida Lindkvist, Norad*

The Norwegian aid administration has been criticized of late as recent evaluations have pointed to its limitations in reporting sufficiently on results at the level of outcomes and impact. When results are reported, focus seems to be on documenting changes, rather than demonstrating effects. This tendency may overestimate the contribution of Norwegian development assistance as many development indicators have improved over the last decades. Using the agency model to explain the dynamics of aid we show that lacking accountability to beneficiaries on the ground, coupled with a conflict of interest between beneficiaries and donors can explain these results. We argue that donors face a real trade-off between being perceived as charitable and actually improving the situation of the poor, while the poor only care about results. In other words, donors may care both about feeling good and doing good, and a clear result focus may conflict with the former. This does not mean that results documentation for example by conducting impact evaluations is not possible, only that donor agencies require strong institutions to ensure that results are documented and current evaluations and studies suggests that this is not the case. Interestingly a stronger results focus from the political leadership may not improve matters unless strong institutions are introduced to ensure the evaluability of aid. As long as agents care about looking good, then a stronger results focus may simply shift aid to where it can most easily be measured rather than where potential to do good is greatest.

#### IFAD’s 2013–2015 Impact Evaluation Initiative

*Alessandra Garbero, Statistics and Studies for Development Division, IFAD*

In 2013 IFAD committed itself to carry out 30 rigorous impact evaluations—3 to 6 RCTs and 24 to 27 ex-post evaluations or deep-dives. These impact evaluations are to be the basis for understanding impact pathways of IFAD supported programs, with the rigor required to attribute these impacts to the intervention. They also constitute a basis for statistical inference of portfolio-wide impact, beyond the selected ex-post evaluations. While, from a learning perspective, it is preferable to select purposively strategically-significant programs, such statistical inference to the portfolio requires random selection of projects for ex-post evaluation; as well as the utilization of RIMS data and qualitative assessments

(shallow-dives) available for projects that were not selected for deep-dives; and meta-studies or systematic reviews. The statistical inference will thus complement the learning-cum-attribution of the deep-dives, with a statement on contributions of the entire portfolio, as a matter of accountability. Besides serving learning and accountability requirements, the initiative also seeks to mainstream in IFAD an approach to impact evaluation that optimizes state-of-the-art rigorous methodologies with cost-effectiveness. The impact being measured consists of “people moved out of poverty” (in money metric terms), economic mobility, empowerment and resilience; and indicators used are household consumption, household assets, and anthropometric measures of child nutrition status. The paper describes: IFAD’s approach for this initiative; the project and methodology selection process; challenges faced due to the specific nature of IFAD supported programs (long term projects, seasonality and shocks related to agriculture, multiple treatment, targeting and randomization, ...); IFAD’s organizational set-up and its partnership approach; lessons learnt and recommendations.

## Session 21

### Education in Asia

*Chair: Usman Mustafa, PIDE, Pakistan*

#### **Economic Success and Awards: Impact Evaluation of the Metrobank Foundation Award for Outstanding Teacher**

*Majah-Leah V. Ravago, Ph.D., School of Economics University of the Philippines and Dennis S. Mapa, Ph.D., School of Statistics University of the Philippines*

Metrobank Foundation Search for Outstanding Teachers (SOT) was launched in 1985 to promote a culture of excellence in education by recognizing the country’s best mentors. Now running on its 28th year, more than 300 exceptional elementary, high school, and college teachers from all over the country have received the highly coveted award. The awardees are evaluated according to the following criteria: personal qualities and character, instructional competence and teaching effectiveness, and professional and community involvement

The basic research question of the study is: What is the impact of the Outstanding Teacher Award on the economic earnings of the recipient? In view of the above the following are the objectives of the study: (1) To investigate the impact of the annual award on the economic success of the teachers and to what extent it influences their success, and (2) To examine whether the impact extends to the school and to the immediate community in general.

In assessing the impact of the Outstanding Teacher Award, this study utilizes Regression Discontinuity Design (RDD), which is a quasi-experimental technique where the conferment of the award is based on an observed score. The basic intuition behind the RDD technique is that individuals “just around” the cut-off score presumably hold similar characteristics. In the absence of the award, their earnings trajectory would be the same. The individuals that are “just around” the cut-off score are the 20 national finalists. Anyone of them can be a winner except that there are only 10 awards to be given.

#### **Managerial Incentives in Public Service Delivery: Evidence from School-Based Nutrition Programs in Rural People’s Republic of China**

*Sean Sylvia, University of Maryland*

A large literature examines performance pay for managers in the private sector, but less is known about the use of performance pay for their public sector counterparts. Improving public service delivery may

nonetheless depend heavily on aligning the incentives of public sector managers with social objectives. In this paper we study performance incentives for school administrators and how their responses to incentives vary with the amount of resources under their control. Our specific focus is the implementation of new, school-based programs to reduce childhood anemia in rural People’s Republic of China. We randomly assigned 170 schools to three levels of performance pay for reductions in student anemia across which we orthogonally assign two levels of block grants. We emphasize three key findings. First, with a smaller block grant, large incentives were effective, but smaller incentives (10% of the size) were ineffective in reducing anemia. Second, absent explicit anemia-based incentives, increasing the size of block grants under the control of school administrators lead to sizeable reductions (but was nearly twice as costly as incentives alone). Third, we find that incentives crowd out the effect of additional resources (or vice-versa). Our evidence suggests that this crowding-out result is attributable, at least in part, to risk avoidance and the nature of existing ‘bureaucratic incentives’ facing school administrators.

**Impact Evaluation of BRAC Community-Based Education for Marginalized Girls in Afghanistan: Lessons from the Baseline Study**

*Mohammed Mahbulul Kabir and Hasib Reza, BRAC*

Long-standing political instability, war, insecurity, and unfriendly socio-cultural norms and practices have been causing girls’ lower access to education in Afghanistan. BRAC has been undertaking Community Based Education for Marginalized Girls in Afghanistan since 2013 to address girls’ education barriers in the country. For evaluation of the project, a randomized control trial (RCT) based experiment design was developed. The study design allows us to capture immediate impact of the project on the beneficiary groups, the marginalized girls in and out of school. Learning, the main outcome of the project success, was measured by girls’ performance in Early Grade Reading Assessment (EGRA) and Early Grade Math Assessment (EGMA) between the control and treatment community schools. First, our presentation will focus on the various practical issues that we considered in planning a rigorous evaluation design in a conflict region. Secondly, we will portray the challenges (and our steps to overcome them) in a large scale baseline survey (conducted in December 2013) where data were collected from 200 communities and 150 government schools across 10 provinces, many of which were highly unsecured. Finally, we will reflect on the data analysis where we had to deal with the ceiling effect of the EGRA and EGMA test scores.

**Session 22**

**Impact Evaluations in Infrastructure Planning and Policy in Urban and Rural Contexts—Uses, Challenges and Opportunities**

*Chair: Tyrrell Duncan, Director, East Asia Transport Division, EARD, ADB*

**Road Improvements in Republic of Georgia**

*John Felkner, Assistant Professor at Department of Urban and Regional Planning, Florida State University*

I will present the impact evaluation of the MCC Compact Road improvements in Republic of Georgia. The Samtskhe-Javakheti Roads Activity was a primary component of the MCC Compact designed to address chronic infrastructure challenges in one of the poorest regions of Georgia. The project undertook rehabilitation of approximately 220 km of roads in a predominantly agricultural area of South Georgia in

order to reduce transportation costs, improve access to markets and employment, and spur economic activity. Construction began in the spring of 2008, and was completed by December 2010. Using three estimation methods: difference-in-difference, continuous treatment approach, matched diff-in-diff, it is found that the roads improvements increased the volume of traffic by an average of 44.2 vehicles per day (4.2%), while the average speed increased by 13.6 km/h (24.4%). The project also+ led to a 26.9% increase in the number of industrial facilities (i.e. canneries, factories, agricultural processing facilities, and similar enterprises) near the project roads. We do not observe changes in cropping patterns or land use at the community level. The project had impacts on the prices of a number of food products on local markets, but in a complex way. No strong evidence is observed on income, consumption, asset ownership, employment, or utilization of health and education services.

### **Evaluating Infrastructure: The Perspective from AfDB**

*Elsa de Morais Sarmento, Principal Evaluation Officer, Independent Development Evaluation, African Development Bank*

The African Development Bank invests almost 80% of its portfolio in infrastructure. I will discuss the development of a monitoring and evaluation framework for the organization and how impact evaluations may fit into this overall framework. I will also discuss what impact evaluations can additionally incorporate into them so that they become more useful to planners and program officers. Finally I will highlight some of the institutional and ground-level challenges of undertaking impact evaluations in the field of infrastructure.

### **Impact Evaluation of Roads in Guatemala**

*Jyotsna Puri, 3ie*

The objective of this study is to examine the impact of road-related investments in Guatemala. Between 1998 and 2005 the country received more than \$200 million in loans for road related projects, accounting for more than 50% of the total foreign investment in the country. These investments have presumed a poverty reduction impact on the people of Guatemala but little analytical work was been done to support the assumption.

Using household and primary sampling unit level data I study if road related investments during the period 1995–2000 led to increases in welfare for 4,844 households in Guatemala. The study uses the propensity score method with nearest neighbor matching to assess the impacts that investments on roads have had on selected rural areas of Guatemala. I estimate the effects of better roads, caused by road rehabilitation and road repair, on access and costs of merit services such as health and education.

The main conclusions from this study are that road investments in Guatemala have had a significant impact on road quality and public transport services. These are important impacts. In a country where more than half the population is poor and almost three-fourths of the rural population is poor, dependence on publicly provided transport infrastructure is high. Guatemala also has few cheap and easily accessible alternatives to road transport: Rail transport is almost absent in Guatemala and air transport is too expensive to be used by the average person. Road investments and their impacts on transportation services have economy-wide repercussions.

### **Impact Evaluation of Urban Infrastructure: Possibilities in the Age of Big Data**

*Somik Lall, World Bank*

Impact evaluations are especially difficult when being planned for large urban infrastructure projects. The presenter will discuss the case of such infrastructure project.

## Session 23

### Improving Health Outcomes through Private Sector Interventions—Experimental Evidence from Ghana, Kenya, Jordan, and Madagascar

*Chair: Patricia Moser, Lead Health Specialist, RSDD, Chair of Health CoP, ADB*

#### Can Mobile Phone Messages to Licensed Chemical Sellers Improve Diarrhea Treatment?—A Randomized Controlled Trial in Ghana

*Minki Chatterji, Abt Associates*

Private-sector providers are a commonly used source of treatment for pediatric diarrhea in low income countries, but often fail to dispense WHO-recommended treatments. This study evaluated an SMS-based campaign to encourage appropriate diarrhea treatment among private-sector licensed chemical sellers (LCS) in Ghana. This study addresses the following research questions: To what extent did the SMS intervention lead to a change in (i) knowledge of the appropriate diarrhea-treatment and (ii) provision of the appropriate diarrhea-treatment.

This study used an experimental study design. The treatment group (455 LCS) received an eight week SMS message intervention to encourage the use of WHO-recommended treatments (oral rehydration solution and); the control group (455 LCS) did not. The study used an innovative combination of data collection methods. A face-to-face provider interview measured self-reported practices. Mystery client visits were used to assess actual behavior of the LCS.

The SMS messages led to an increase in the proportion of LCS who reported providing appropriate treatments. Yet, the intervention did not lead to improvements in actual provider practices. There was a sizable gap between reported and actual practices.

The study highlights the value of SMS message programs/interventions to improve knowledge among LCS, and also showcases barriers to translating knowledge into practice. SHOPS is currently using these findings to implement an innovative system of real-time, mobile device-based program that is providing supportive supervision to more than 2000 Ghanaian LCS.

#### What Is the Impact of Evidence-Based Medicine (EBM) on Family Planning Providers' Knowledge, Attitudes, and Practice Behaviors? A Randomized-Controlled Experiment in Jordan

*Minki Chatterji, Abt Associates*

Private providers' biases and misconceptions can adversely affect the delivery of reproductive health services, as is the case with hormonal injectables in Jordan. The SHOPS Project aims to address those biases against hormonal contraceptive methods and expand their use and availability working with the private sector through the evidence based medicine (EBM) approach. This intervention involves providing health care providers with access to the best research evidence on clinical topics and encourages them to integrate it with their clinical experience and patient preferences.

In this study we evaluated the impact of an EBM intervention in Jordan on providers' knowledge, attitudes, and reported practices related to DMPA, a three-month injectable. Using a randomized controlled trial design, 135 providers (treatment group) were invited to attend a roundtable discussion on evidence-based health benefits and side effects of DMPA and received two detailing visits to reinforce these



messages at their clinics. Providers from the control group (132 providers) were not invited to attend the seminar and received no information on DMPA. Baseline and follow-up surveys were administered.

We found evidence of a positive impact on providers' confidence in prescribing DMPA. However, we found no impact on providers' overall knowledge about DMPA's side effects and on reported clinical practices. Participation in the DMPA EBM roundtables was also very low. Our findings suggest that an EBM intervention alone may not be an effective mechanism to improve providers' knowledge, attitude and reported practices when there is strong provider and consumer bias against a contraceptive method.

### **Do Free Pregnancy Test Kits Supplied by Community Health Workers Increase Their Family Planning Clients? Evidence from a Randomized Controlled Study in Madagascar**

*Randall Juras, Abt Associates*

Many countries in Sub-Saharan Africa use community health workers (CHWs) to increase access to contraceptives. In Madagascar, CHWs use a checklist to assess whether a woman is pregnant or not before providing hormonal contraceptives. In other countries, CHWs are allowed to provide such contraceptives only at time of menstruation. Many women categorized by the checklist as possibly pregnant are not. In some cases, CHWs do not trust the checklist and deny contraceptives to non-menstruating women. This results in missed opportunities to provide women with effective contraceptive methods.

We conducted an experimental study in Madagascar to assess the impact of a program offering free pregnancy test kits to CHWs for distribution to clients, along with training on how to use them. Eligible CHWs had been trained to distribute both oral and injectable contraceptives as part of a social marketing program. We randomly assigned 311 CHWs into a treatment group that received the pregnancy test kits and training, and 311 CHWs into a control group. Four months later we found that CHWs in the treatment group reported 24% more new clients for hormonal contraceptives per month (compared to the control group). Now, Madagascar's Ministry of Health is currently considering including pregnancy tests in the CHWs' basket of products. However, there is concern that scaling up this intervention will result in medicalization of family planning practice which may lead to new requirements that women take pregnancy tests to obtain contraceptives. Translating our findings into practice is bringing interesting and new challenges.

### **Impact Evaluation of a Family Planning Health Service in Kenya**

*Douglas Johnson, Abt Associates*

We estimate the impact of m4RH, an mHealth service providing information on family planning methods in Kenya, on knowledge and use of contraception through a randomized controlled trial. Despite relatively high awareness of contraceptive methods, misconceptions regarding contraception are common among Kenyan m4RH was designed to increase knowledge of family planning and reduce misconceptions by providing free on-demand information on various methods and nearby clinic locations through an SMS-based platform. During the study period, all users were randomly assigned to either the treatment group, which received full access to the m4RH service, or a control group, which only received general motivational messages. Data were collected from users via questions sent directly over SMS. Response rates to the SMS surveys were relatively high (approximately 51% for the first survey) and very similar across treatment and control groups.

We find that having access to the full set of m4RH services led to significant gains in family-planning knowledge: m4RH caused a 13% increase in the number of knowledge questions users answered



correctly. We did not find a corresponding impact on contraceptive use. Our findings show that SMS is an effective means of providing access to family planning information but this increase in knowledge may not necessarily translate into an increase in use. In addition to estimating the impact of m4RH, we utilize the rich dataset of all incoming requests to the m4RH service along with results from qualitative interviews to analyze why and how users used the m4RH service.

## Session 24

### Designing Agricultural Impact Evaluations

*Chair: Victor Owusu, Kwame Nkrumah University of Science and Technology, Ghana*

#### Designing Rigorous Impact Evaluations of Agricultural “Pull” Mechanisms

*Stephen H. Bell and Tulika Narayan, Abt Associates Inc.*

Traditional “push” mechanisms for improving agricultural outcomes in developing countries—program grants to NGOs, government agency actions, informational campaigns—have not produced sustainable results at scale. In response, donor organizations are exploring the use of “pull” mechanisms to achieve development objectives by incentivizing market activities that yield profits to the private sector while advancing social objectives such as greater farm family incomes and healthier foods for urban consumers. These initiatives create incentives by awarding prizes to private sector actors for delivering services similar to what push mechanisms might deliver such as increasing awareness of farmers to produce improved grain for the market, inducing farmers to adopt better technologies, or developing and/or expanding adoption of effective livestock vaccines. Measuring the impact of these initiatives on the agricultural value chain and consumer outcomes poses special challenges, particularly concerning the counterfactual –i.e., what farmers, aggregators, and input suppliers would have done in the absence of the enhanced business payoffs created by the “pull” prizes. Three impact evaluations respond to these challenges to support the “AgResults” multi-donor learning initiative focused on “pull” mechanisms in Sub-Saharan Africa and South and Southeast Asia. These designs use a range of causal inference strategies adapted to in-country circumstances such as random assignment of villages to early versus late intervention implementation, comparative interrupted time series analyses, and specification tests based on “untreated outcomes.” The proposed presentation describes these designs, previews the learning about the use of pull mechanisms for agricultural development they are expected to produce over the next few years, and considers broader lessons for rigorous impact evaluation of agricultural development initiatives.

#### Evaluating the Impacts of Rice Genetic Improvement in the Post Green Revolution Period

*David Raitzer, Economist, ERD, ADB*

Modern varieties of rice have had exceptionally high documented benefits during the Green Revolution, but the degree to which this impact has been maintained by subsequent investments in research by ADB (under the Asian Biotechnology Research Network RETAs) and other donors is uncertain. The evidence generated to date on this issue relies principally on methods that do not address selection bias, and which have unresolved threats to internal validity. Using more rigorous approaches than applied previously, this evaluation assesses the degree to which newer (post 1989) rice varieties have led to increased agricultural productivity, economic surplus, welfare for the poor, food security and environmental benefits in Bangladesh, Indonesia and the Philippines. It does so first by characterizing outputs, the traits embodied in new varieties, including attainable yield and disease and pest resistance,

while accounting for resistance breakdown as varieties age. Subsequently, nationally representative data are used to estimate subnational patterns of varietal and trait adoption over time. Panel and quasi panel national data are used in econometric fixed effects production functions and damage abatement models to identify on farm yield effects attributable to trait adoption as proximate impacts. This is accompanied by bio-economic estimation of disease resistance yield effects in a spatial framework, drawing on on-farm experimental designs that control for selection and placement bias.

Econometric and bio-economic yield effect estimates and adoption parameters are used to estimate supply shifts, which are applied in a subnationally disaggregated partial equilibrium model to quantify annual welfare effects for consumers, producers, hired labor, populations under poverty lines, the environment and the food insecure. Under the main set of modeling assumptions, PPP\$25 billion of benefits are generated from 1990 to 2010, of which PPP\$9 billion are attributable to international collaboration. Approximately 45% of benefits are captured by those under the PPP\$2.0 per day poverty line, while reduced caloric insufficiency is estimate to save nearly 1.5 million Disability Adjusted Life Years (DALYs).

### **Income Dimension of Impact Evaluation**

*Murad Mukhtarov, M&E Consultant Azerbaijan, Azerbaijan*

The paper describes the approach of determining of the income dimension in the impact evaluation of the agricultural projects. The present investigation goes beyond the indicators using modern economical and financial categories as structural components of the model.

In spite of application of the approach to impact evaluation within the project cycle, it is time-invariant and can be applied in the long-term perspective after project completion.

The study approach has been applied to various World Bank Projects in Azerbaijan and contributed to State Program for Poverty Reduction and Sustainable Development (SPPRSD) including “*Supporting sustainable and balanced growth of the non-oil economy*”.

## **Session 25**

### **The Success of Poverty Alleviation Program**

*Chair: Gerry Britan, USAID*

#### **Impact of Social Protection Programs on Social Welfare: The Case of the Philippines**

*Hyun H. Son, Principal Evaluation Specialist, ADB and Nanak Kakwani*

The proposed study will evaluate social protection programs currently implemented in the Philippines. These programs include (i) public and private pension programs; (ii) survivor’s benefit; (iii) Pantawid Pamilyang Pilipino Program; (iv) National Health Insurance Program; (v) Training for Work Scholarship Program; (vi) Agrarian Reform Community Development; (vii) Community-based Employment Program; (viii) disability benefit; and (ix) scholarship benefits and student financial assistance. The evaluation will focus on how well the programs target lower-income groups of the population.

The main objective of social protection programs is to maximize social welfare. This study will develop a new methodology to measure the impact of the mentioned programs on social welfare. This methodology can be used to evaluate any social program by calculating the social rate of return—the higher is the social rate of return, the greater will be the effectiveness of the program in maximizing

social welfare given a government’s budget constraint. This evaluation considers both transfers going to the beneficiaries and the administrative costs of the program. A social protection program, even if it is well targeted, cannot be deemed efficient if it incurs heavy administrative costs. This methodology will weigh a social protection program’s targeting efficiency against its administrative costs because a trade-off may exist between the two. This methodology provides very useful contribution in designing social protection programs to help achieve maximum social welfare given minimum costs.

The study will use the 2011 Annual Poverty Indicator Survey (APIS) conducted by the National Statistical Office. The APIS is a nationwide survey designed to provide poverty indicators at the national and provincial levels. Social protection programs covered by the 2011 APIS include: (i) public and private pension programs; (ii) survivor’s benefit; (iii) Pantawid Pamilyang Pilipino Program; (iv) National Health Insurance Program; (v) Training for Work Scholarship Program; (vi) Agrarian Reform Community Development; (vii) Community-based Employment Program; (viii) Disability benefit; and (ix) Scholarship Benefits and Student Financial Assistance.

### **How Far Does a Big Push Really Push? Evidence from an Extreme-Poverty Eradication Program in Bangladesh**

*Wameq A Raza, BRAC*

Despite numerous development interventions implemented, poverty remains endemic with the extreme poor being circumvented by many development programs. Recognizing this BRAC Bangladesh implemented *Challenging the Frontiers of Poverty Reduction: Targeting the Ultra Poor* (CFPR) in 2002 to mitigate ultra-poverty using a multipronged approach. The program provides support in the form of asset transfer (worth USD\$100), education, healthcare, social and legal support for a period of 2 years. While earlier studies have demonstrated the program’s impact on the short and medium term, this paper is the first to investigate the long term effects of such a holistic program.

The paper utilizes a four-round panel data over a period of 9 years (2002–2005–2008–2011) collected from the three poorest districts. We assess the impact of the program on income, asset holdings, employment and food security using difference-in-differences regression with inverse propensity weighting. Heterogeneity of effects is assessed across time, baseline employment and gender of the household head.

Results confirm marked effects of CFPR participation in the short (2002–2005) and medium term (2005–2008) on all outcomes. In the long term (2008–2011) however, effects are much smaller, mostly driven by a catch-up of the control group. The program’s long term impact on per capita income is larger for entrepreneurs than for beggars/maids or day laborers. Female headed households benefit more than their male counterparts.

Holistic anti-poverty programs such as CFPR can have drastic impacts on households’ socioeconomic status in the short term. Long term effects are more limited, especially for the most vulnerable employment groups.

### **How Does Contract Design Affect the Uptake of Microcredit among the Ultra Poor? Experimental Evidence from River Islands of Northern Bangladesh**

*Abu Shonchoy, Institute of Developing Economies (IDE-JETRO) and University of Tokyo, Japan*

Despite the professed claims of microcredit for poverty alleviation, little is known about the type of suitable credit contracts that are appropriate for extremely poor households or the ultra poor. To fulfill this knowledge gap, we initiated a field experiment in river islands of Northern Bangladesh,

where a substantial portion of dwellers could be categorized as ultra poor who are repeatedly affected by cyclic floods and erosion. We randomly offered four types of loans to such dwellers: typical small cash loans with one-year maturity, large cash loans with three-year maturity with and without a year of grace period, and in-kind livestock loans with three-year maturity with a year of grace period. Analyzing the rates and determinants of uptake, we found that the uptake rate is lowest in the regular microcredit contract followed by the in-kind loan. Contrary to prior belief, we also found that the microcredit demand of the ultra poor is not necessarily small and in particular, the ultra poor are significantly more likely to accept the contract if a grace period coupled with longer maturity is attached in a relatively large credit product, irrespective of the credit type (cash or in-kind). This paper provides evidence that the typical “rigid” microcredit of yearly maturity without a grace period is not particularly attractive for the ultra poor. Microfinance institutions may need to re-think about the contract design of their products by introducing longer loan maturity with convenient grace periods, without compromising the loan repayment disciplines, to better accommodate the need of the ultra poor.

## Session 26

### Complexity, External Validity, and Policy Influence

*Chair: Penny Hawkins, DFID, UK*

#### Reading between the Lines: Context, Social Capital and Complexity in Impact Evaluation

*Heewoong Kim, UNDP Egypt*

The rise of rigorous impact evaluation (RIE) has significantly changed the debates surrounding development cooperation. Rather than focusing on certain issue areas, impact evaluation shifted the orientation to effectiveness and results. It led to programs designed based on a theory of change and emphasized M&E and documentation of lessons learned. Furthermore, piloted with a handful of interventions, RIE provided a common framework to review, analyse and debate the development programs.

However, despite such rise of RIE, there remain some critical gaps in RIE, especially in terms of representation of context, social capital and complexity in RIE, to *make impact evaluation matter*. While successful RIEs have been conducted for simplified interventions, the program itself has multiple variables: e.g., a simple conditional cash transfer program can vary based on its targeting, amount and process of transfer, type of conditions, source of monitoring, etc. There are new innovations in RIE to capture the value of social capital, voice and values, however, these experiments are still in a preliminary stage. Current RIE seem to focus more on easily quantifiable variables such as asset transfers, days of training. Also, there is lack of consideration on the social, political and economic context that enables certain interventions to succeed over others. This paper looks into these gaps of context, social capital and complexity for RIE by comparing some RIE with other project evaluation reports. It also attempts to make suggestions by emphasizing context-specific small sample experiments (field labs), mix methodologies and other tools to make IE matter.

#### External Validity Clues: On What to Look for and How

*Vegard Iversen, University of Manchester and Heather Lanthorn, 3ie*

Impact evaluations may fixate on internal validity of identification, at the expense of external validity (e.g. Pritchett and Sandefur, 2013). In a timely and thoughtful response to this fixation, Woolcock (2013) presents a new framework built around ‘key facts’ to assess claims of external validity of research

designs and findings, allowing for careful consideration of when and where the programmatic and policy lessons learnt may apply. The framework parallels similar attempts by Cartwright (2011, 2013).

We take a different approach to considering external validity, taking the Theory of Change (ToC) as our starting point. In doing so, we highlight several gaps in the prior efforts. First, the need to account for a broader set of program design and supply-side considerations. Second, the need to account for the demand-side of programs and interventions, including individual and community demand and possible barriers related to local collective action capacity.

Our ToC approach engages social science theory and evidence to identify and help sharpen the focus on critical assumptions and critical variation across settings—that is, to assess whether Program Z, initially implemented successfully in Setting X, might be ‘replicated,’ in new Setting Y.

We are concerned, moreover, that Woolcock’s decision to anchor his framework in discussions of causal density and complexity may blur and obscure where daylight is required. We address this concern in relation to considering how to make the design and results of impact evaluation useful for programmatic, policy, and investment decision-making across settings.

### **Addressing Attribution of Cause and Effect in Small Impact Evaluations: Towards an Integrated Framework**

*Daniel Phillips, 3ie*

There exists considerable consensus among impact evaluators conducting large n impact evaluations involving tests of statistical difference in outcomes between the treatment group and a properly constructed comparison group. However no such consensus exists when it comes to assessing attribution in small n cases—that is when there are insufficient units of assignment to conduct statistical analysis of what difference the intervention has made. This presentation examines various evaluation approaches that could potentially be suitable for small n analysis and find that a number of them share a methodological core which could provide a basis for consensus.

## Session 27

### **The Role of Think Tanks in Impact Evaluation**

IDRC panel. Abstracts not available at time of printing.

## Session 28

### **Approaches to infrastructure impact evaluations**

*Chair: Michael Peter Barrow, OPSD Deputy Director General, ADB*

### **Micro-Impacts of Infrastructure in Facilitating Social Capital: A Case Study of an Irrigation Project in Sri Lanka**

*Yasuyuki Sawada, University of Tokyo*

While randomized controlled trials (RCT) became one of the main tools to evaluate the effectiveness of development programs and policies aimed at reducing poverty, it is difficult by nature to design

randomized infrastructure projects to identify the causal impacts. In this presentation, we will discuss how natural experiments and other evaluation methods can be utilized to evaluate infrastructure programs rigorously using a unique panel data set from an irrigation project in Sri Lanka under a natural experimental situation where a significant portion of irrigated land was allocated through a lottery mechanism. Also, we capture a set of new poverty outcomes such as social capital and trust using artefactual field experiments. By combining these pieces of data, we quantify causal micro-impacts of infrastructure on social capital.

### **Impact of Private Sector Participation on Access and Quality of Services: Systematic Review of Evidence from the Electricity, Telecommunications, and Water Supply Sectors**

*A. Thillairajan, Indian Institute of Technology Madras, India*

Individual studies on the impact of such Private Sector Participation (PSP) have yielded mixed results. This study synthesizes evidence on the impact of PSP on access and quality outcomes. Studies to be included for the review were selected using well defined inclusion, exclusion, and quality appraisal criteria. Meta regression analysis of 90 observations from 17 econometric studies was used to synthesize the evidences. Evidence on access was stronger as compared to that of quality. This indicated that PSP as a tool for improving quality of services has not been effective. Among the three sectors studied, telecom sector showed the strongest evidence of a true empirical effect between access and PSP. This indicated that PSP has been an effective strategy to improve telecom access. In the case of electricity and water supply, the evidence indicated a positive and negative impact on access respectively, though the effect was not significant. Overall, the results suggested that the effectiveness of PSP in improving access and quality in infrastructure services delivery has been limited, and underlines the need for creating appropriate monitoring and incentive structures to make PSP interventions more effective in improving outcomes.

## **Session 29**

### **Evaluating Energy and Mining Investments**

*Chair: Adora Navarro, PIDS, Philippines*

#### **Pico-Solar Lights: Exposure, Uptake, and Impact**

*Elly White, SolarAid*

SunnyMoney, UK charity SolarAid's social enterprise works to build a sustainable market for clean, safe, affordable energy through its innovative School Campaign model. With support from the Ministry of Education, field teams in rural Africa work alongside schools to promote and sell solar lights. Headteachers, as trusted members of the community, share the message of this new technology with students, parents and communities to encourage the use of the pico-solar lights for child study after dark in areas with no electricity access. The solar lights are then sold through the headteachers to members interested. The Light Library is a new model designed and piloted by SunnyMoney, to selected schools in rural Senegal. The objective was to increase access to, awareness of, and use of solar lights. The model was followed by the School Campaign in the regions selected for the Light Library implementation and in a comparison region to test the effect on demand.

This presentation explores the poverty effects of replacing kerosene lamps, or alternative traditional lighting methods, with pico-solar lights and the opportunities they can present for educational



improvement and environmental impact. It also reviews the impact of the Light Library model on uptake of solar lights in rural African communities. The research so far has led to data from over 12,500 interviews, surveys and observations collected using structured interviews in 36 regions of the five countries SunnyMoney is working in.

Results of the School Campaign show that building a market is possible and low-income families are reached; the lights saved considerable household income, children were doing more hours of study after dark, and improved health was reported. While only offering a basic level of modern energy access, pico-solar lights can have significant impact on household income levels in rural Africa, and other development-related activities.

The Light Library model increased child study hours, improved community interaction with the school and increased enrollment rates. It enabled families to save money from avoiding costs of lighting and have an opportunity to test out this new technology in a risk-free environment. The follow up School Campaign resulted in over double uptake in the Light Library schools with 35% of the school population purchasing a solar light for home use, compared to 15% in the comparison region.

Results showed that building a market is possible and low-income families are reached but that providing access and exposure, as well as raising awareness and trust in the new technology was key to uptake and adoption. Research to inform program design and delivery is key to effectiveness as well as communication with policy-makers, governments, donors and practitioners. A joint paper has been written with the World Bank to share this learning externally and allow other African governments or NGOs to implement or adapt the Light Library model. In order to provide more evidence and learning SolarAid is currently undertaking the first large-scale randomized control trial into the impact of pico-solar lights on poverty alleviation.

### **Impact Evaluation of the School Electrification and Education (SEEd) Project of the Alliance for Mindanao Off-Grid Renewable Energy Program (AMORE): A Utilization of Propensity Score Matching (PSM)**

*Nelson C. Faustino and Michael Arie P. Medina, College of Forestry and Environmental Science  
Central Mindanao University, Philippines*

The School Electrification and Education (SEEd) Project of the Alliance for Mindanao Off-grid Renewable Energy Program (AMORE) is a USAID funded project which aims to help students in off-grid areas in Mindanao improve their learning experiences through the provision of educational materials in multimedia format. These provided both teachers and students with improved teaching and learning experiences amid conditions characterized by poverty and conflict. Part of the package also includes a solar energy system for the rural school beneficiaries in order to facilitate the use of the above learning materials.

This presentation will provide a detail on how Propensity Score Matching (PSM) was employed in order to obtain a matched sample of the intervention group (student beneficiaries) and the control group (non-beneficiaries). The matched sample represents a subset of the respondents from both groups with more or less similar socioeconomic background thereby providing an unbiased estimation of the intervention effects. The t-test was then employed to compare the matched samples from the intervention and control groups in terms of their performance in the National Achievement Test (NAT).

The t-test revealed a significantly higher NAT mean percentage score (MPS) among the intervention group compared to the control group. These results showed that the AMORE-SEEd project has impacted the student beneficiaries in terms of their improvement in the NAT.



## **Welfare and Environmental Impacts Biogas Technology Adoption: Lessons from Tigray Regional State of Ethiopia**

*Nigussie Abadi, Department of Natural Resource Economics and Management, Mekelle University, Mekelle, Ethiopia*

Many Ethiopians face quality of life and livelihood challenges associated with the sub-optimal sanitation, dependence on traditional fuel, and decreasing agricultural productivity. To mitigate this livelihood challenges the government of Ethiopia has recognized the need for a national policy framework that encourages the uptake of biogas technology. However, despite the expectations of improved livelihood and environmental outcomes from the biogas technology, there was no rigorous impact evaluation of such technology. To this end this study tries to bridge the gap by investigating the impact of biogas technology adoption on welfare of the households and on the environment using 200 sample households in rural Tigray regional state of Ethiopia. We used a propensity score matching to analyze the impact of biogas technology on food security outcomes (per capita expenditure per adult equivalent), health outcomes (medical expenditure, number of households affected by indoor smoke relate illness and number of days absent from productive work due to indoor related illness) and environmental outcomes (fuelwood consumption, time spent for fuelwood collection and use of inorganic fertilizers). The Average Treatment Effect (ATT) results show from nearest neighbor (NN) and Kernel matching show that biogas adopters had higher per capita expenditure per adult equivalent. In terms of its impact on health, our results from both matching show that biogas adopters spent less medical cost, less affected by indoor related diseases and less number of days absent from work. We also found that biogas adopters consume less fuelwood, travel less to collect fuelwood and use less in organic fertilizer than their counterparts. Our finding has an important implication to provide win –win for poverty reduction and protection of the environment through the promotion of renewable energy sources such as biogas technology.

## **Assessing Mining Impacts on Households with Propensity Score Matching (PSM) for Responsible Mining Policy Implications: The Case of Surigao Provinces in Caraga Region, Philippines**

*Raquel M. Balanay, Caraga State University, Philippines and Jose M. Yorobe, Professor, Department of Agricultural Economics at University of the Philippines Los Baños*

Because of selection bias and endogeneity, PSM is used in the assessment of mining impacts on households for responsible mining policy implications in the Surigao Provinces of Caraga Region, Philippines. Nearest neighbor, kernel and radius matching procedures have been utilized to check for consistency of findings in the analysis of outcomes from 699 households that were selected through multi-stage random sampling from mining and non-mining barangays of Surigao Provinces.

The final probit model for propensity score estimation has implied that mining is identified with lower assets, educational attainment and years in the community, but higher food expenditure and number of employed household members, which are affirmed in the findings of average treatment effects on the treated (ATT) under radius matching. Mining is implied to cause higher health costs under nearest neighbor and higher food expenditure under kernel with epanovich bandwidth of 0.06. The two more restrictive matching procedures than radius also indicate that mining can induce higher educational attainment.

Health expenditure, being positively signed under nearest neighbor, may reflect the health hazards of mining's externalities. However, household income remains insignificant throughout the analysis, which indicates a shortfall in catalyzing the income effects of mining. The study recommends that responsible mining policies should ensure considerable investments for human capital development particularly in education, entrepreneurship and health to catalyze the income effects of mining under a healthy and safe environment in the Surigao Provinces of Caraga Region, Philippines.

## Session 30

### Assessing Health Impacts

Chair: *Litul Baruah, KPMG, India*

#### Impact Evaluation of ICT Application for Strengthening Health Service Delivery in Rural PNG

*Sakiko Tanaka, Senior Economist, ERD, ADB and Kitur Urarang, National Department of Health*

Papua New Guinea has widespread poverty and weak health indicators, particularly for maternal and child health. The infant mortality rates were 57 per 1,000 live births and the maternal mortality rate were 733 per 100,000 live births in 2006. Many children also suffer from diseases that are preventable through early immunization. The ADB Rural Primary Health Services Delivery Project (RPHSDP) for Papua New Guinea was approved by ADB on 30 September 2011 to help the Government of Papua New Guinea strengthen the rural health system in selected areas of the country.

An impact evaluation is planned to measure the impact of information communication technology (ICT) (i.e. sending SMS message to pregnant women/new mothers) on improving attendance of maternal and child primary health care (number of visits of antenatal care, attended delivery, well-baby clinic visit within one year, and child immunization) as a component of the RPHSDP.

It is envisaged that the IE will follow a cluster randomized design. Baseline data will be collected from the 120 health centers on a rolling basis. As part of the program, health workers in these centers will be provided with mobile phones. In the first 3 months of the project, patient data, on pregnant women, will be collected by 24 health workers in 24 health centers and sent to a national database (GIS). This data will be aggregated at a health center level, analyzed and separated into four similar groups of 6 health centers. The first group receives Intervention A, the second Intervention B, the third receives Intervention C and the fourth represents the Control. 'Intervention A' involves sending SMS messages to pregnant women at certain intervals before and after the child's birth date, encouraging them to undertake an attended birth and provide immunizations to their babies after birth. 'Intervention B' involves sending the SMS messages with the promise of small phone credit if they attend as requested. 'Intervention C' involves sending SMS messaging and a health professional calling a patient just before the birth to encourage them to attend. The fourth represents the control which receives no intervention. This process will be repeated until data is collected from all HWs in all 120 HCs and groups are assigned.

Approximately one month after the due date (which is now available on the GIS), respondents who have mobile phone numbers in all groups will be contacted on their mobile phones to ask them if they had an attended birth or a home birth. Approximately 1 year after the birth of the child, respondents will be asked if they had provided immunization to their children and baby well clinics. In addition, qualitative survey of semi-structured interviews with new mothers, key informants at all levels (village head/health volunteers, health facility, District Health Office, Provincial Health Office and National Department of Health) will be undertaken to investigate the ICT process (if applicable) and impediments to visiting HCs.

The aim of the presentation is to discuss the purpose, methodology, limitations and implementation challenges of the IE for health sector in PNG.

## Performance-Based Financing in the Context of Selective Free Health-Care: An Impact Evaluation Using Routine Data in Burundi

*Jean-Benoît Falisse, DPhil candidate, Oxford Department of International Development; Juvenal Ndayishimiye, MD, MPH, Cordaid Burundi; René Queffelec, Cordaid Burundi*

Performance-based Financing (PBF) approaches are increasingly popular but still few impact evaluations are available. PBF started in 2006 in Burundi, at the time when user fees were removed for pregnant woman and children under 5, and was gradually extended to the rest of the country. This roll-out is exploited to assess the impact of PBF on the utilization of health-care services.

PBF is associated with substantial increases in anti-tetanus vaccinations of mothers (+23.9 percentage points in coverage target), deliveries at the health facility (+18.4 pp), and post-natal consultations (+19.4 pp). The number of inpatient visits, antenatal visits, and vaccinations of children is not correlated with PBF. Health-facility level figures from PBF-supported provinces show that preventative care indicators are not growing. It is also found that more qualified nurses headed to PBF-supported provinces.

The dataset is limited, does not include the quality of care, and is too poor to assess whether the impact of PBF is resource-driven or due to the PBF incentive mechanism. Since the impact of PBF is mostly visible on free services, the interaction effect between PBF and the selective removal of user fees is probably crucial. A possible explanation is that the removal of user fees increases accessibility to health care and acts on the demand side while PBF gives medical staffs incentives for improving the provision of services.

The study needs to be complemented with research that addresses the interaction between PBF and other health policies and the sustainability of the PBF incentive mechanism.

## Does Coupling of PhilHealth Indigent/Sponsored Program and Conditional Cash Transfer Induce Those with TB Symptoms to Consult/Seek for Treatment?

*Mel Lorenzo M. Accad, University of the Philippines School of Economics, Philippines*

### I. Introduction

Using the Household Members Survey Tuberculosis component of Family Health Survey 2011 in the Philippines, this study compares the Average Treatment Effects on the Treated of coupled/complimented Conditional Cash Transfer and PhilHealth Indigent/Sponsored Program vis-à-vis CCT alone or PhilHealth Indigent/Sponsored Program alone.

### II. Policy Background

Benefits design and selection process of PhilHealth Indigent/Sponsored Program

Benefits design, selection process and conditionality of Conditional Cash Transfer Program

### III. Data and Methodology:

Household Members Survey Tuberculosis component of Philippine Family Health Survey in 2011 using propensity matching score, average treatment effects on the treated

### IV. Results

Preliminary estimation results show that there's significant yet negative average treatment effects from having PhilHealth Indigent/Sponsored Program alone. There's positive but not significant average treatment effects from having CCT alone. Lastly, there's marginally significant and negative average treatment effects from having both PhilHealth Indigent/Sponsored Program and CCT.

### V. Policy Recommendations

### **The Achilles Heel of TB Prevention after a Century of International Medical Programs**

*Ann V. Millard, School of Public Health-McAllen Campus, Texas A&M University Health Science Center and Mary Ann J. Ladia, Institute of Clinical Epidemiology, University of the Philippines Manila-National Institutes of Health*

Although medical advances against tuberculosis (TB) have been significant, the only hope to prevent and eradicate it depends not on drug treatment alone but also a major component of human behavioral regimens extending over a period of six months for each patient, with close monitoring of a treatment partner. The infectious disease was nearly wiped out in the United States until a combination of complacency and cuts in government funding allowed it to re-emerge in forms that became increasingly difficult to diagnose and treat. In some parts of the United States, resurgence of the disease occurs periodically, and similar patterns have developed in the Philippines. An ethnographic study in the Philippines by one author (MAJL) identified reasons for the persistence of the disease in Baguio, a city that has a long history of TB treatment programs and sanatoria. Similarly to the United States, physicians in Baguio City may not always follow uniform methods of TB diagnosis and therefore may not always be able to advise patients and their families similarly about crucial behavioral precautions. Persons affected by TB have varied understandings of the disease, treatment regimens, and behavioral protocols with the result that they sometimes inadvertently infect those they care most about at home and households where they work. The huge international investments in research, prevention, and treatment of TB cannot be effective as long as the human element is insufficiently addressed. Has the fascination with high-tech research, supported by tremendous sums of money, undercut medical education and diffusion of behavioral protocols to the extent that TB has spread instead of being contained and prevented? It is ironic that the increasingly accurate scientific knowledge of TB has been accompanied by a loss of the attention to human behavior and cultural differences that is crucial to containing and preventing TB epidemics in the future.

## **Session 31**

### **Impact Evaluations In and On Conflict: Methodological Innovations and Implications from the Field**

*Chair: Annette Brown, 3ie*

#### **Evaluating Impacts on Youth Propensity for Violence: Challenges and Opportunities**

*Jon Kurtz, Mercy Corps, Director of Research and Learning (or depending on travel availability, Margaret Lada, Mercy Corps Afghanistan Research Manager)*

What interventions work to reduce the risk of youth of engaging in political violence? Mercy Corps has undertaken multiple impact evaluations to examine this policy-relevant topic. We've learned that answering this question poses unique demands and challenges to impact evaluation—both methodological and practical. Working in conflict-affected areas constrains the design options for impact evaluations. Insecurity limits the mobility of data collectors, and the ability to utilize third party surveyors. And gathering accurate data on controversial issues like attitudes and behaviors towards violence poses major difficulties.

During this session, Mercy Corps will share innovative and promising techniques we and our research partners have applied to address these challenges, including: i) overcoming social desirability bias through the use of list experiments to gauge support for extremist groups, and objective measures such

as Cortisol tests for assessing stress and aggression; ii) sampling and interviewing techniques to limit security risks to the implementing agency, enumerators and respondents; and iii) impact evaluation designs to overcome beneficiary selection processes that are politically sensitive and not amenable to randomization.

The presentation will draw on impact evaluations of Mercy Corps' youth-focused programs in Afghanistan and Liberia. The Afghanistan evaluation used a quasi-experimental pipeline design; the Liberia study—still underway—is a randomized impact evaluation with multiple treatment arms. The presenter will conclude by linking Mercy Corps' experiences to recommendations from the growing body of work on how to conduct impact evaluations in and on conflict, and with populations at risk of violence.

### **Capturing Change during Changing Times: Conducting Impact Evaluations with Highly Mobile, Conflict-Affected Youth**

*Juliette Seban, International Rescue Committee, research and evaluation advisor for economic interventions*

How can we best capture program impact during a time of change for the study population, while working within rapidly changing environments? Working with conflict-affected and highly mobile adolescent and youth populations, IRC has conducted four impact evaluations on youth and adolescent skills development and livelihood programs in Burundi, Liberia and Kenya. In Burundi, the program brought together returning refugee youth with those who remained during the conflict. In Kenya, the program worked with refugees from Somalia and vulnerable Kenyans while they navigated life in an urban slum. The impact evaluations each sought to better understand the extent to which market-driven economic programming is able to improve the income and well-being of young people, investigating the variables that drive success. All four impact evaluations were randomized controlled trials with a qualitative component for two of them. Qualitative components were designed to take the studies beyond the question of “does it work” to understand how and why, particularly around sensitive topics.

This session will reflect on challenges experienced across these evaluations such as selection of appropriate outcomes and measurements for adolescent and youth populations; and managing take-up and the potential for high attrition rates. Solutions explored to address these challenges will be presented including attendance tracking and management through biometrics; partnership with non-traditional actors such as businesses in the evaluations; and coordination between and capacity-building of local program and research teams. The presentation will also highlight methodological and ethical issues around measurement of non-cognitive and violence outcomes.

### **Evaluating the Effects of a Sport-for-Development Job Training Program on At-Risk Youth in Urban Central America**

*Mike Duthie, Social Impact, Senior Director of Impact Evaluation*

Since October 2011, Social Impact has been implementing randomized control trial (RCT) impact evaluations of sport-for-development programs in Honduras and Guatemala. The focal point of these studies is the USAID-funded A Ganar program, which utilizes soccer and other team sports to help at-risk youth find jobs, learn entrepreneurial skills, and/or re-enter the formal education system. A Ganar focuses on building ‘employability’ and life skills and reducing risk behaviors through lessons developed through sports. The research designs utilize mixed-methods RCTs (one with multiple treatment arms) in tracking three thousand youth over a 5-year study period. Through the use of field-tested measurement tools and randomized response games, the team is testing the efficacy of various approaches to quantify difficult-to-measure constructs (e.g. self-esteem, gender norms, risk behavior, etc.). Bolstering the survey research is a multipronged qualitative component consisting of: (1) multi-perspective, longitudinal case studies comparing in-depth interviews with participants and a sample



of their social network (parent, employer, trainer, peer, etc.), and (2) exit interviews conducted with participant youth at the conclusion of the program. Given the crime-afflicted environment (5 youth out of a recent cohort of 675 were killed over the 10 months between baseline and initial follow-up), the evaluation has developed various strategies to ensure the safety of respondents and field staff. The presentation will include an overview of evaluation methodology, focusing on measurement, strategies for limiting attrition, and tips for conducting research in high-threat environments.

### **Collecting the Best Data in Challenging Settings**

*Nassreena Sampaco-Baddir, Innovations for Poverty Action Philippines, Country Director*

Through our Post-Conflict Recovery Research Initiative, IPA is working with local governments, leading academics and NGOs on the ground to conduct some of the first large-scale rigorous randomized evaluations of projects focused on peace-building and the unique set of challenges facing post-conflict and fragile states as well as countries with sub-national conflict. With a broad range of contexts in need of evidence, IPA aims to identify programs that work both in response to uniquely challenging post-conflict contexts as well as programs that address issues more widespread across developing economies. IPA has conducted evaluations in Liberia, Sierra Leone, northern Uganda and Cote d'Ivoire which address a range of issues including youth employment, intimate partner violence, ex-combatant reintegration, community driven development and female entrepreneurship. We are currently focusing on peacebuilding, state building and reintegration work in the Philippines (funded by ADB) and Colombia as well as partnering on the Evidence for Peace Initiative with 3ie and the World Bank.

In order to fill evidence gaps on what works in conflict-affected states, it's necessary to ensure that data collected is accurate and that the data collection process does not put respondents or surveyors at risk. In this panel, we will touch on innovative methods for data collection from studies in conflict-affected areas, including recent IPA work in Mindanao. Mobile technologies, list randomization, and complementary, systematic qualitative work in addition to careful survey design inform our work in conflict-affected and fragile settings.

## **Session 32**

### **Intended and Unintended Consequences of Social Protection**

*Chair: Sri Wening Handayani, Principal Social Development Specialist, RSDD, ADB*

#### **Intra-Household Resource Control, Subjective Expectations and Human Capital Investment. Evidence from a Randomized Policy Experiment**

*Alex Armand, University College London and Institute for Fiscal Studies*

This paper studies how the interaction between intra-household allocation of resources and parental beliefs about the returns to education influences human capital investment among poor households. For this purpose, I study a conditional cash transfer program in the Republic of Macedonia, aiming at improving secondary school enrollment among children in poor households. For identification I exploit the random allocation of payments either to mothers or household heads, together with a unique information on parental subjective expectations of returns to schooling. I show that targeting mothers leads to an increase in secondary school enrollment only for children whose parental returns are sufficiently high at the beginning of the program. This effect is associated with an increase in individual

expenditure shares on education for this group. I find no differential impact for other inputs, such as monitoring of school attendance and time use. Overall, I show that the effect of channeling resources to mothers is strictly related to heterogeneity in parental perceived returns to schooling.

### Does Social Protection Promote Diversification as Autonomous Climate Change Adaptation Strategy? Evidence from Rural Ethiopia

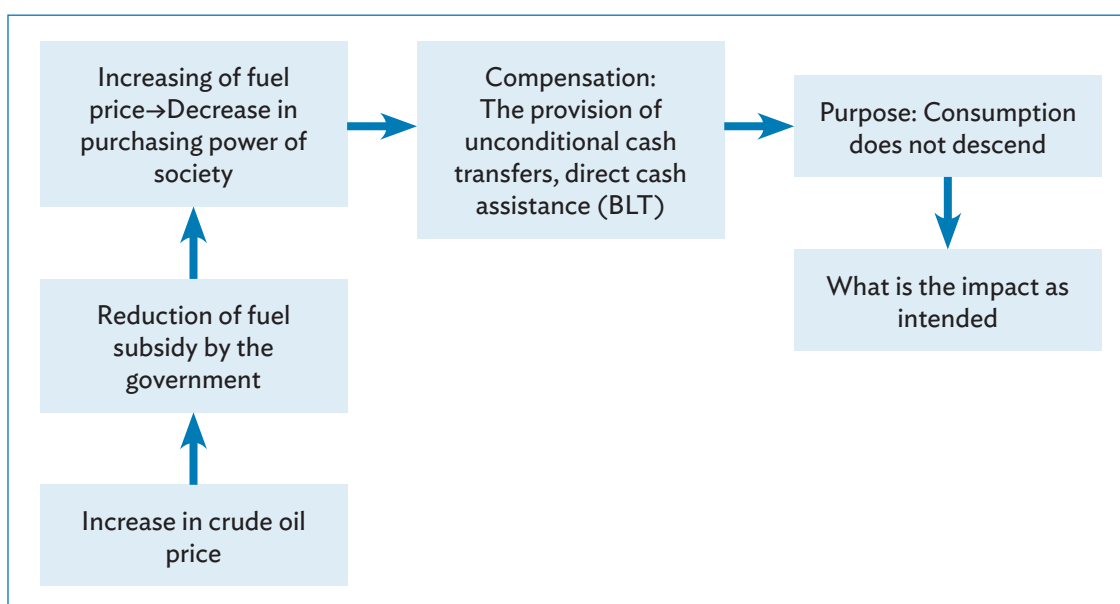
Zerihun Berhane Weldegebriel, *International Doctoral School in Local Development and Global Dynamics, University of Trento, Italy*

It is widely predicted that climate change will have an adverse impact on Ethiopian agriculture and exacerbate the inextricable problem of food insecurity. Social protection schemes such as the Productive Safety Net Program (PSNP) have been designed to address food insecurity and promote livelihoods. This scheme can potentially contribute to households' autonomous adaptation by reducing vulnerability to climatic shocks.

This paper examines the impact of the social protection schemes on autonomous adaptation strategies by taking the case of household income diversification. It uses non-experimental approaches namely difference-in-differences combined with Propensity Score Matching for a panel of 1306 rural households from the two recent rounds of the Ethiopian Rural Household surveys (ERHS) for the period 2004 and 2009. Taking advantage of the extensive data available on a range of activities and incomes, the paper makes a conceptual distinction between nonfarm and off-farm income, and uses the Adaptive Social Protection framework to examine the impact of the Productive Social Protection Program (PSNP). The results indicate that receiving transfers from the PSNP, on average increases income from nonfarm activities while significantly decreasing income from off-farm activities. These results confirm the hypothesis that social protection can promote positive adaptation strategies and serve as effective way of reducing the vulnerability of smallholders to climate change induced shocks.

### The Impact of Transfer Program (BLT) to Cigarette Consumption in Indonesia

Inayati Nuraini Dwiputri, *Faculty of Economics and Business-Universitas Gadjah Mada*





- I. Introduction
- II. Methodology
  - A. Data
    - IFLS in year 2000 (IFLS3)
    - IFLS in year 2007 (IFLS4)
    - Unit of analysis: Household
  - B. Potential of Bias:
    - Non-random program selection bias, there are criteria of poor households by Central Bureau of Statistics (BPS)
    - Unobserved factor and time invariant lead to endogeneity problem
    - So, RCT (Randomized Control Trial) could not be applied
  - C. Estimation Techniques, to make correction with endogeneity:
    - 1) Fixed Effect in the household level
    - 2) Instrument Variable (IV)
    - 3) DID (Difference in Difference)
  - D. Model:
    - 1) DID:
      - $Y_{ijt} = \alpha_0 + \beta_1 BLT_{ijt} + \lambda T_{ij} + \delta (BLT_{ijt} * Y_{ij}) + \beta_2 X_{ijt} b + v_{ijt}$  (1)
      - $Y_{ijt}$ : outcome cigarette consumption of household -i, village- j, year-t
      - BLT: dummy variable, 1 = household -i, village -j, year-t who got BLT; 0: households do not
      - $T_{ij}$ : dummy, 1= year 2007; 0= year 2000.
      - $X_{ijt}$ : control variable
      - $V_{ijt}$ : error term
    - 2) Fixed Effect in household level:
      - $Y_{ijt} = \alpha_0 + \beta_1 BLT_{ijt} + \lambda T_{ij} + \delta (BLT_{ijt} * T_{ij}) + \beta_2 X_{ijt} + a_i + v_{ijt}$  (2)
      - $a_i$ : household fixed effect
    - 3) IV for BLT program: education level of household head (it's appropriate with criteria of recipient of BLT)
      - We compare all the result and make policy recommendation

## Session 33

### Generating More and Better Employment

*Chair: Bart Edes, Poverty Reduction, Social Development and Governance Division Director, RSDD, ADB*

#### What Works in Youth Employment? A Systematic Review of the Evidence on Interventions to Improve Labor Market Outcomes of Youth

*Jonathan Stöterau, RWI, Germany*

Young men and women have been particularly affected by the economic crisis. In 2008–2009, the global youth unemployment rate saw its largest annual increase and it is projected to increase even

further (ILO, 2013). Policymakers therefore seek new ideas and evidence on what works best to improve the labor market conditions of young people.

This presentation will show results from a 3ie-supported Systematic Review of the available evidence on the effectiveness of youth employment programs. Our study is the first to document recent impact evaluations across a broad spectrum of youth interventions in a systematic, rigorous manner.

Through an extensive search of the available literature 32,117 potentially relevant primary studies were identified. Our resulting dataset comprises more than 1900 effect sizes from 53 studies that were considered of adequate content and methodological rigor to inform the meta-analysis. Additionally, for each of these studies, we collect detailed information about the intervention design and the research methods employed for its evaluation.

These data allow us to analyze effects of youth interventions in different contexts and across different intervention types, from skills training to public works programs. Further, the presentation will review the current state of methods to evaluate labor market effects of such interventions and highlight gaps in the available evidence. Finally, we discuss how to deal with the challenges that arise in synthesizing the evidence across a heterogeneous range of interventions.

### **The Impact of the People’s Republic of China’s New Labor Contract Law on Socioeconomic Outcomes for Migrant and Urban Workers**

*Zhiming Cheng, University of Wollongong, Australia and Wertheim Fellow, Harvard Law School, Harvard University*

This paper examines the effect of having a labor contract on a range of employee outcomes (hours worked, social insurance coverage, subjective well-being and wages) for a sample of urban and migrant workers in the People’s Republic of China. A methodological contribution is that we use propensity score matching, which allows us to draw causal inferences about the relationship between having a labor contract and each of these variables. We find that the effect of having a labor contract on employee outcomes is generally large and larger than the findings in other studies suggest. As such, our results suggest that the People’s Republic of China’s Labor Contract Law, which has made signing labor contracts mandatory, has been effective in improving the outcomes for Chinese workers

### **Incomplete Information and Migration Decisions: Do Job Fairs and Information Work?**

*Emily A. Beam, National University of Singapore, Singapore*

I implement a field experiment in the rural Philippines to measure the impact of information and job fairs on overseas job search. Information about average overseas wages increases short-run migration demand and expected overseas wages, indicating that individuals underestimate the returns to migration, but

it does not increase overseas job search. Similarly, I find that job fairs are ineffective in linking attendees with overseas jobs, at least for the marginal attendee, but this additional labor-market exposure has unexpected but persistent effects domestically, increasing the likelihood of search in Manila and formal sector employment in the ten months after the fair.

## Can Arts-Based Interventions Enhance Labor Market Outcomes among Youth? Evidence from a Randomized Trial in Rio de Janeiro

*Jochen Kluge, Humboldt-Universität zu Berlin, RWI Essen and IZA Bonn*

This paper provides findings of a small-scale, innovative labor training program that uses expressive arts and theatre as a pedagogical tool. The corresponding life skills training component is combined with a technical component teaching vocational skills. To our knowledge, this is the first paper to rigorously evaluate the effectiveness of a training program constructed around expressive arts. Using a randomized assignment of favela youth into program and control groups, we look at the short-run treatment effects on a comprehensive set of outcomes including employment and earnings as well as measures of personality traits and risk behavior. We find positive short-run employment and earnings impacts five months after the program finalized; no impacts are found for shorter periods. These short-run impacts are economically very large, compared to those typically found in the literature: a 33.3 per cent increase in the probability of being employed, and a 23.6 per cent increase in earnings. We find no evidence of significant program impacts on other outcomes, including personality-related traits, providing evidence that these traits may not be malleable for young adults in the short-run. We argue that the estimated labor market impacts are due to a combination of both skills formation and signaling of higher quality workers to employers.

## Session 34

### Evaluating Governance Interventions

*Chair: Ramesh Subramaniam, Deputy Director General, SERD, ADB*

#### How Alternative Development Spills over into More Honest Communities: An Experimental Approach in Colombia

*Juanita Vasquez, PhD candidate, University of Gottingen*

First, the context of Colombian coca growing areas is presented. These are marginalized areas where illegality is the norm and big efforts have been put into place in order to eradicate coca. We focus on Colombia's biggest alternative development program and build on the impact evaluation carried out by Econometria, a consultancy firm, who allowed us to introduce and honest experiment within the impact survey in order to capture effects on honesty and legality. These topics are usually hard to capture in "standard" surveys, as people tend to have problems answering truthfully. We present a novel way to enhance impact evaluations in a simple and straightforward way. We will close with our findings that alternative development programs have an impact in reducing dishonesty. We go further into explaining the impact on dishonesty and find that living in areas with more coca increases dishonesty as well as in areas where there are more human rights violations.

#### On Strengthening Management Capacity Female Local Government Members: Learning from a Randomized Experiment in Bangladesh

*Atonu Rabbani, University of Dhaka*

Community based organizations can play important roles in economic development especially when traditional market and/or state based mechanisms are weak. In this paper, we will take advantage of a randomized experiment to shed light on empowerment of female local political leaders through an awareness building campaign and formation of women—centric community development groups. Civic engagements and knowledge building are positively associated with social development (as we show in

this paper). Such factors can make a political leader more accountable and ensure delivery of socially desirable goods and services. We show that this allows the female local government members (FLGMs) to interact with community women groups (CWGs) to prioritize sanitation and hygiene related issues compared to the FLGMs in the control areas. FLGMs receiving treatment are also more likely to have “plans” to address sanitation and hygiene issues and are more likely to work with CWGs to make such plans. We also investigated other actors (namely, volunteer groups and general public) and interaction between those actors and the FLGMs. We found a much weaker causal links in these cases. Hence, we conclude that our results are not spurious and causal interpretations between the interventions and management outcomes are valid. To the best of our knowledge this is probably the first effort to use such framework to understand organizational and management capacities of local political leaders.

### **Incentivizing Development: A Field Experiment with Third Tier Community Organizations in Pakistan**

*Asha Gul and Hamna Ahmed, Research Fellows, Center for Research in Economics and Business (CREB), Lahore School of Economics, Pakistan*

We will present the experimental design for a forthcoming randomized controlled trial (RCT) with NGOs (also called Third Tier Organizations—TTOs) in Pakistan. These are mature community institutions at the Union Council level, formed as apex organizations of many local first and second-tier organizations. Presently, there are almost 800 TTOs (all will be part of the RCT), covering about 1.3 million rural households across Pakistan. These institutions are set-up and supported by the funder. The funder must find effective ways to engage with these institutions to ensure (i) sustainability i.e. effectiveness in serving the community without funder’s support, (ii) inclusion of underrepresented groups and (iii) flexibility to respond to local needs.

The aim of this RCT is to test the impact of two sustainable measures, which the funder can take to engage with these institutions to facilitate their transition into independent organizations capable of sustaining collective action in their communities on their own. These measures are (i) self-reporting and (ii) non-financial incentives. We propose to test the impact of these measures on TTOs’ governance and service delivery separately through four treatment arms: reporting on governance in T1 (service delivery in T2), incentives to best performing TTOs reporting on governance in T3 (service delivery in T4).

The RCT would help understand whether these measures improve performance of TTOs, which measure is more effective in boosting TTOs’ performance, is it better to incentivize on governance or service delivery? The Baseline data collection will be undertaken from August to September 2014.

## Session 35

### **Community-Driven Development in Asia**

*Chair: Nazmul Chaudhury, World Bank*

#### **Community-Driven Development Projects: An Overview of How the Millennium Challenge Corporation is Broadening the Evidence Base in CDD**

*Michelle Inkley, Millennium Challenge Corporation*

The Millennium Challenge Corporation (MCC) was established in 2004 as an independent US foreign aid agency that works in partnership with competitively selected countries that demonstrate a commitment to good governance, economic freedom, and investment in their citizens. Currently, MCC

has CDD investments in two countries. This presentation will focus on the experimental evaluations for the Philippines and Indonesia CDD projects.

For the Philippines CDD project, KALAHIS CIDSS (KC), the cost-benefit analysis focused on expected benefits from the infrastructure subprojects (faster travel times from improved roads, better health outcomes from clinics, etc.), while highlighting the absence of monetizable data on the social benefits of KC. In addition to measuring changes in income, the MCC-funded evaluation of the KC project includes two methods designed to capture social benefits, including community participation and decision-making: (i) a controlled simulation of community decision-making, and (ii) observations of village assemblies.

The Indonesia Community-Based Health and Nutrition to Reduce Stunting Project (Nutrition Project) seeks to reduce and prevent low birth weight, childhood stunting, and malnourishment of children. The Nutrition Project builds on the successful and rigorously-evaluated *Program Nasional Pemberdayaan Masyarakat—Generasi Sehat dan Cerdas* (PNPM-Generasi) by enhancing this demand-side, community-driven development project. The enhancement includes building the capacity of maternal and child health care workers around feeding and growth monitoring, the distribution of micronutrients to pregnant women and children, training service providers on how to promote proper sanitation and hygiene behavior, and national and local awareness campaigns to induce behavior change around breastfeeding and complementary feeding.

### **KALAHIS CIDSS-NCDDP Municipal Talakayan: Learning Together towards Results**

*Dir. Benilda Redaja, National Project Manager, Department of Social Welfare and Development*

The “Municipal Talakayan (Dialogue)” of the KALAHIS CIDSS-NCDDP is an innovative M&E strategy which evaluates the overall local development through a process that is participatory, transformative and empowering. It brings together local stakeholders for a democratic dialogue which (1) provides snapshot of the conditions and level of development; (2) increases local capacities on the use of data for participatory planning, implementation and evaluation; and (3) enhances awareness and understanding of the stakeholders’ roles and responsibilities in local development and governance.

Municipal Talakayan has the following features: (1) relatively simple and less technical using available local and project data; (2) provides analysis and venue of presenting development status, gaps and directions at municipal level inducing local policy change; (3) captures community perspective; and (4) complements impact and process evaluations in determining policy reform at national and local levels. This type of monitoring and evaluation can be a useful, cost-effective complement to impact evaluations—the experience of how it is being used should be useful to others working on complex projects and evaluations where timely data is needed to understand implementation.

### **Promoting Technology, Greater Transparency and Accountability in the Operations of KALAHIS CIDSS-NCDDP with the Aid of Project Information Management System (PIMS)**

*Assistant Secretary Camilo G. Gudmalin, Deputy National Program Director for Operations, Department of Social Welfare and Development*

Monitoring and Evaluation for community-driven development (CDD) is more necessary and challenging because of its dispersed implementation and difficulty in managing the flow of data and information into an MIS system. At the same time, the objective of good governance and building social capital necessitates evaluating less easily measurable goals (e.g. transparency, accountability, and empowerment). Along these challenges, the KALAHIS CIDSS-NCDDP is developing a technology

called as the Project Information Management System (PIMS), a facility that will confidently satisfy the information requirements, feedback systems, and decision-making needs of the Program. With PIMS, the Program will obtain and manage huge volume of data coming from the community by utilizing the following major components:

- 1) SMS and android application for data capture and submission
- 2) Business Intelligence and GIS for data accessibility and visual analysis
- 3) Geo-tagging for sub-projects' progress monitoring
- 4) Interactive website and web forum for improved quality of exchange interactions among stakeholders

The session will showcase how a user-oriented information system such as PIMS supports an efficient and successful M&E system for CDD program. A system that will not only capacitate the LGUs and communities on simplified data collection and processing, enriched public access to information and use in promoting local development, but also one that enables KALAHY CIDSS-NCDDP beneficiaries to convey their demands for better outcomes.

### **Community-Driven Development Projects: Assessing Quality of Participation in Decision-Making**

*Jansen Mayor, Millennium Challenge Account-Philippines, and Edith Felix, Innovations for Poverty Action*

Kapit-Bisig Laban sa Kahirapan—Comprehensive Integrated Delivery of Social Service (KALAHY-CIDSS) is a community-driven development (CDD) program in the Philippines. As a CDD program, it espouses the principles of community empowerment and participatory governance in delivering projects to address poverty. By giving communities control of the development process, resources, and decision-making authority, the CDD approach enables communities to decide for themselves how best they can improve their conditions.

As part of the evaluation of KALAHY-CIDSS, we recently conducted a study to look at the quality of community participation in the decision-making process. Two methodologies were employed to do this. The first was direct observation of the conduct of Barangay (community) Assembly (BA). BAs are twice a year mandated meetings that each barangay should do to enable community members to join in the decision making process. The second methodology we used was a simulation of the decision making process through a structured community activity (SCA). In the SCA, the communities were given a small grant that they need to spend on some infrastructure. The condition of the grant is that the community officials had to call a community meeting to decide on where they are going to use the grant. Both methodologies used the same instruments to tally participation and record the meeting content. The result of the study offers interesting insights on the dynamics of decision-making in a community.

This presentation will present the results which we believe could help current and future implementation of a CDD type project.

### **To What Extent Can Community Empowerment Contribute to Health and Education Outcome? Findings and Lessons from an Impact Evaluation of PNPM Generasi in Indonesia and Next Steps**

*Yulia Herawati, Audrey Sacks, and Rob Wrobel, World Bank*

In this presentation, we will present the findings and lessons from a randomized control trial of a community empowerment program, PNPM Generasi, in Indonesia. The results suggest that the program contributed to several positive impacts including a reduction of malnutrition among children aged 0 to



3 and significant increases in the utilization of basic health services offered at community health clinic. The impact on longer term outcomes has been very limited. There is no evidence of impact on wasting, infant and neo-natal mortality, the incidence of diarrhea, and measures of learning outcome. These results are consistent with global evidence of Conditional Cash Transfer programs with an evaluation period of three to four years. In order to measure the potential impact of the program on these longer term outcomes, we plan to conduct another round of the RCT in 2015—eight years after the baseline was conducted—with the addition of a qualitative component. In addition, this presentation will raise questions about the effectiveness of performance-based incentives based on the experiences of PNPM Generas. We will conclude with a discussion of the lessons we learned about the limitations of quantitative RCTs and the need to carry out mixed methods evaluations.

## Session 36 ICT and Education

*Chair: Jouko Sarvi, Advisor, RSOD conc Practice Leader (Education), RSGS (Chair of Education CoP), ADB*

### **The Impact of Mobile Technologies on Education in Low- and Middle-Income Countries: A Mixed-Methods Systematic Review**

*Laurenz Langer, University of Johannesburg*

This presentation reports the findings of a mixed-methods systematic review of mobile technology's impact on education in low- and middle-income countries. It further presents an innovative approach to adapt systematic review in the context of international development, i.e. the concept of mixed-methods research synthesis.

A mixed-methods approach to systematic reviewing has the ability to answer not only traditional effectiveness questions of 'what works', but further provide insights in contexts and mechanisms to explain how and why an intervention might (or might not) work. Given the complexity of mobile learning programs in developing countries, the conducted systematic review subscribes to such a mixed-methods model of reviewing.

The review features two distinctive modules: an effectiveness module to aggregate the impact of mobile technologies on learning and teaching outcomes using statistical meta-analysis as a method of synthesis; and a configurative module to formulate a program theory of mobile learning for development. This configurative module uses thematic synthesis to investigate the contexts and mechanism that influence program effectiveness and consequently draws from qualitative research evidence.

The findings of both modules are combined to answer the question of mobile learning for development's effectiveness, and, in addition, to construct an evidence-informed program theory of how mobile technologies might most effectively be applied in resource-poor education systems. The presentation will show in detail how the mixed-methods approach to systematic reviewing was operationalized, i.e. how different inclusion criteria, quality assessment tools, and synthesis methods were developed for each review module. Lessons for policymakers and practitioners will similarly be presented.

### **The Impact of a One Laptop per Child Program on Learning. Evidence from Uruguay**

*Alina Machado, Instituto de Economía—Universidad de la Republica, uruguay*

We present evidence on the impact of one of the largest deployments of an OLPC program and the only one implemented at a national scale: Plan Ceibal in Uruguay on students' math and reading scores. Unlike previous work in the field, we have unique data that allow us to know the exact date of laptop delivery for every student in the sample. This gives us the ability to use a continuous treatment, where days of exposure are used as a treatment intensity measure. Given that there is some variation in the date of laptop delivery across individuals within the same school, we can identify the effect of the program net of potential heterogeneity in the rate schools gain improvements on student's achievement over time in the absence of the OLPC program across the country (i.e. we allow each school to follow a different learning growth curve over time due to unobservable time-varying heterogeneity). We also run an alternative specification where we allow for different learning growth curves over time between schools located in Montevideo and the rest of Uruguay. Our results suggest that the program had no effects on math and reading scores. The zero effect could be explained by the fact that laptops in class are mainly used to search for information on the internet.

### **Assessing the Efficacy of Computer Technology in Education: Evidence from a Pooled Study of Computer Assisted Learning Programs among Rural Students in the People's Republic of China**

*Weiming Huang (Stanford University)*

There is large heterogeneity among the studies that investigate whether computer technology improves education and what students benefit most from it. The overall goal of this study is to assess the effectiveness of computing technologies to positively raise educational performance and non-cognitive outcomes and identify what program component are most effective. To achieve this aim we pooled the data sets of five separate studies about computer technology programs that include a total of 16,342 students from 171 primary schools across three provinces in the People's Republic of China. We find that overall computing technologies have positive and significant impacts on student academic achievement in both math and in Chinese. The programs are found to be more effective if they are implemented out-of-school, avoiding what appear to be substitution effects when programs are run during school. These programs do have heterogeneous effects by gender. Specifically, boys gain more than girls in Chinese. We did not find heterogeneous effects by student initial achievement levels. We also found that the programs that help students learn math—but not Chinese—have positive impacts on student self-efficacy.

### **A Quasi-Experiment on the Effect of e-Learning on Educational Progress and Improvement in Mathematics in Sri Lanka**

*Bilesha Weeraratne, Institute of Policy Studies*

The predominant approach in teaching mathematics in Sri Lanka involves traditional methods. However, many developed countries have adopted e-learning to complement other methods of teaching mathematics with mixed results. In an Asian Development Bank funded project a quasi-experiment is carried out to test the effect of e-learning systems on educational progress and improvement in mathematics achievement and cognitive ability of eighth grade students in Sri Lanka. The experiment involves 800 eighth grade students in 20 schools for two academic terms. In each school 20 students form the treatment group, while another 20 students comprise the control group. As the intervention, the treatment group students are taught mathematics through an e-learning software in the school computer lab by the mathematics teacher during 2–3 days of the 5 day instruction for mathematics,

while the remaining days involve regular class room instruction by the same mathematics teacher. The control group is given regular class room instruction by their respective mathematics teacher during all 5 days of the week. The experiment involves a base line, end-of-term, and end-line tests administered to all 800 students and survey questionnaires administered to students, their teachers and school administrators. The tests encompass material taught in the second and third academic terms of the eighth grade mathematics curriculum. This paper intends performing an impact evaluation of the ongoing quasi-experiment using test scores and survey data in a difference-in-difference framework. The outcomes to be analyzed include students standardized test scores and item response theory adjusted test scores, and students self-reported attitude towards learning mathematics.

## Session 37

### Improving Newborns and Infant Health and Nutrition

*Chair: Somashekhar Nimbalkar, Head, Neonatal Intensive Care Unit, Shree Krishna Hospital and Pramukhswami Medical College, India*

#### India's ICDS Program and Child Under-Nutrition: Evaluation of Program Effectiveness and Impact

*Sakti P. Pal, President, Development Evaluation Society of India (DESI); Consultant, Independent Evaluation Office, Government of India; Former Adviser (Evaluation), Planning Commission, India*

India's Integrated Child Development Services (ICDS), a Flagship Program with universal coverage aims at reducing malnutrition, morbidity and mortality among women and children by providing a "package" of services through the network of *Anganwari* (service delivery centers) and other agencies. ICDS is expected to improve the nutritional status of children/women and contribute towards attainment of Millennium Development Goals (MDGs 1, 2, 4 & 5) relating to mortalities and morbidities caused by poverty, malnutrition and illiteracy. A comprehensive evaluation of ICDS was undertaken at the instance of Planning Commission, GOI under the direction and guidance of the author. This presentation is confined to assessment of impact of ICDS on *nutritional status of children* and *identification of factors influencing the outcome*.

The study design made use of the Program Logic Model and both secondary and primary data to test the validity of assumptions implicit in the *causal chain*. The findings tend to suggest that wherever program implementation had followed the Causal Chain impact of ICDS on nutritional outcome was found to be positive. Inter-State variation in impact could be explained by the nature and extent of deviations of the relevant inputs, activities and outputs from "norms" during implementation. Use of "propensity score matching" brought out positive impact, which was clouded because of confounding factors, implying validity of the Program Logic of ICDS.

ICDS has been reformed for better outcome. The important recommendations of this study considered in program restructuring related to: improvement of service conditions of delivery personnel, increased allocation of funds to equip *Anganwari* Centers for quality service delivery, scientific growth monitoring, indexing of cost norms to inflation, convergence of complementary services and targeting services to the poor and needy.

## **A Systematic Review of Evidence on the WHO Essential Newborn Care Training Program**

*Marian Theresia Valera, UPEcon Health Policy Development Program, Philippines*

Over the years, substantial progress has been made towards the achievement of Millennium Development Goal (MDG) 4, which calls for a two-thirds reduction in under-five mortality by year 2015. In response to this challenge, the World Health Organization introduced a low-cost intervention in the form of the Essential Newborn Care (ENC) training program.

Methods: (1) Inclusion Criteria: Controlled before-and-after (CBA) studies with an active baseline design were reviewed. Only studies where outcomes were measured and compared before the ENC program has been introduced and again after the training has been conducted and applied will be included. The primary outcome measured in the analysis is early (7-day) neonatal mortality rate. (2) Search Methods: Among the 363 studies whose abstracts were subjected to a rapid review, 41 were considered as potentially eligible studies based on the outcomes measured. Upon further review of the full texts, three controlled before-and-after studies that assessed the effectiveness of the ENC program were selected for the final review. (3) Data Analysis: The effect size on early neonatal mortality (summary relative risk (RR) with 95% confidence intervals, CI) was estimated by comparing early NMRs post- and pre-ENC program initiation.

Conclusion: The empirical evidence on the effectiveness of the WHO ENC training program in the developing country setting is associated with a slight reduction in early neonatal mortality in developing countries.

## **Constrained Provider Choice: The Effect of a TBA Ban on Mortality**

*Edward Okeke, MD Ph.D., RAND*

Reducing the high rates of maternal and child mortality are two of the five Millennium Development Goals, yet, many countries continue to lag well behind 2015 MDG targets. This has led to a search for alternative strategies to reduce maternal and child mortality, particularly mortality associated with pregnancy and childbirth. A controversial strategy that is being considered in several countries is a ban on the use of traditional birth attendants (TBA). TBAs attend more than 50% of births in developing countries, with attendance being as high as 90% in some countries. Their use however remains the subject of fierce debate with many arguing that the poor quality of care provided by TBAs leads to preventable deaths. Despite the lack of empirical evidence, several countries have either implemented, or are considering implementing, bans on use of TBAs. In this paper, we evaluate the effects of a 2007 TBA ban implemented in Malawi on institutional deliveries and infant mortality. Using data from the 2010 Demographic and Health Surveys combined with a unique dataset on the location of all health facilities in Malawi, we estimate the effect of the ban using double-difference and triple-difference econometric models. We find that the ban increased the probability of an institutional delivery, and increased the probability that a skilled provider attended a birth. Our results also indicate a small reduction in infant mortality. To our knowledge, this is the first study to study the effect of a ban on TBAs on institutional deliveries and mortality.

## Session 38

### Promoting Adoption of New Crops and Techniques in Agriculture

*Chair: Bekele Shiferaw, PEP-Global office, Nairobi, Kenya*

#### Evaluating the Impact of Women's Training in Home Gardens in Bangladesh: Combining Quantitative and Qualitative Data

*Pepijn Schreinemachers and Marie Antoinette Patalagsa, AVRDC—The World Vegetable Center*

Home-based vegetable production has been recognized as a nutrition- and gender-sensitive intervention that has the potential to improve nutrition in developing countries, yet evidence to support this practice is lacking. Our study tested whether training women to grow improved home gardens contributes to increased production and consumption of vegetables and strengthens their position within the household. Data were collected in two southern districts of Bangladesh. Qualitative data to assess gender impacts came from focus group discussions and in-depth interviews with men and women. Quantitative data came from a questionnaire survey of 582 women. Selection bias was minimized by comparing women who had received training in 2012 with women who had been identified to receive training in 2013. Training increased the per capita production of vegetables (mostly leafy green vegetables) from 20 to 37 kg per year (+86%). The diversity of production and frequency of harvesting also increased. In terms of nutrient yields, training increased the supply of plant proteins by 171%, iron by 284%, vitamin A by 189% and vitamin C by 290%. There also was an increase in the diversity of vegetables consumed. Improved home gardens reduced women's dependence on their husbands for buying vegetables. Many men helped in the garden and appreciated the lower food expenses. As traditionally they do not work together, both men and women felt that working together in the garden improved their relationship. The study therefore provides evidence that home gardens have a positive effect on vegetable production and consumption and improve intra-household relationships.

#### Impact of Integrated Aquaculture-agriculture Technologies Value Chain Participation on Welfare of Marginalized Indigenous Households in Bangladesh: A Panel Data Analysis

*Abu Hayat Md. Saiful Islam, Bangladesh Agricultural University, Bangladesh*

Integrated aquaculture-agriculture (IAA) (e.g. Rice-fish system) is a potential sustainable intensification system to produce more food from less land and water. But for widespread diffusion and poverty reducing policy making, it is necessary to know the impacts in terms of household welfare. Thus in the light of ongoing debates about the sustainable development paradigm, this paper examines the link between participation in IAA technologies value chain and welfare effects on marginalized indigenous households using a large three-year household panel survey of IAA value chain participators (treated) and non-participator (control) households from four agro-ecological regions of Bangladesh. The distributional effect of IAA participation is also investigated. The study applied standard fixed effect, heckit panel model, control function approach, IV fixed effect and Hausman Taylor regression to control for endogeneity of IAA participation as well as to control for unobserved heterogeneity. The causal impact estimation reveals that IAA value chain participation has a significant positive impact on household income, expenditures and consumption frequency specially fish consumption in all those estimates. We also found evidence that IAA value chain participation has stronger impact on welfare of relatively better off households who participated in IAA production related value chain activities than landless extreme poor households who participated in IAA upstream and downstream value chain activities. Such findings illustrate the potential of IAA to contribute to extreme poverty reduction and

improvements in livelihoods and food and nutrition security in Bangladesh and possibly other countries, which have similar socioeconomic, agro-ecological, and institutional settings.

### **Impact of the Adoption of Improved Technology for Rice Parboiling on Nonfarm Income and Children Schooling in Benin**

*Lidia Dandedjrohoun, University of Abomey-Calavi, Benin*

Parboiling is a post-harvest operation of the paddy (unhusked) rice which enhances both the nutritional quality and the milling yield. The widespread use of traditional methods of rice parboiling in Benin leads to poor quality of local parboiled rice which is not competitive in the market. In order to address this problem the National Agricultural Research System of Benin in collaboration with AfricaRice Center (AfricaRice) have developed and introduced an improved rice parboiling equipment to end users. An educational video developed by AfricaRice was used for its diffusion. Several studies had shown technical performances of the improved equipment. However, the technical performances of a given technology introduced in rural area or not, are not sufficient to show its impact on adopters. Then, the key question is to know whether the adoption of the improved parboiling technology has really improved the livelihood of the adopters' households.

This paper aims to assess the impact of the adoption of the improved technology on the women parboiler households' nonfarm income and children schooling expenditure. To achieve this goal, the counterfactual outcomes framework is used on data collected from a sample of 200 women selected in Benin.

The primary results of analysis indicate that potential adopters' households have increased their per year nonfarm income by 67582 FCFA. As regards the impact on children schooling, adoption of improved parboiling equipment generates an increase of 4829 FCFA for the total children schooling expenses.

## Session 39

### **Information, Financial Literacy, and Financial Inclusion**

*Chair: Betty Wilkinson, Director, Central West Public Sector and Finance Management Division, ADB*

#### **Financial Life-Skills Education for Children and Youth: Systematic Review and Meta-Analysis**

*Daniel D Shephard, Aflatoun Child Social & Financial Education, The Netherlands*

The education sector has noted weaknesses in life-skills education and mismatches in the transition from education to civic and economic life. In addition, governments, the financial sector and civil society have highlighted the lack of financial capabilities among youth and adults. This lack negatively affects the individuals and families who are consequentially less capable of absorbing unexpected financial shocks and pursuing their goals.

This paper presents results from the first systematic review and meta-analysis of RCTs to investigate the state of the evidence regarding the effectiveness of financial life-skills education for children and youth. The review identified 21 studies and conducted meta-analyses on financial knowledge, attitudes, and behaviors. Subgroup analyses are presented by population age (children, youth), intervention type (in-class, online, mass-media), and intervention intensity.



The identification of 21 studies refutes earlier claims that there are very few experimental studies of financial education. Findings challenge the relative importance placed on knowledge, attitude, and behavior outcomes—with some programs having stronger effects on knowledge and others are having stronger effects on behavior or attitudes. The relative strength of implementation models is explored through subgroup analysis, highlighting good practices, trade-offs, and common pitfalls.

Reflections will be shared on broader implications for education programs that provide life-skills relevant to society and the labor market. The next steps for research and practice are also discussed, with a focus on how the current evidence should be used to scale and test evidence-based models.

### **When Can Financial Education Affect Savings Behavior? Evidence from a Randomized Experiment among Low Income Clients of Branchless Banking in India**

*Santadarshan Sadhu, IFMR Finance Foundation, India*

Despite many important innovations in the banking technology in the recent past, more than half of world's adult population still does not use formal financial services to save or borrow. Lack of knowledge about benefits of formal banking, lack of familiarity with banking products and the difficulty of travelling to bank branches for opening and accessing such accounts are often quoted as major barrier for bringing the low value depositors into the formal banking.

This paper presents a randomized field experiment with clients of a doorstep banking provider that offers access to banking in remote area in the state of Uttar Pradesh in India. We measure the impact of a classroom based financial education program, delivered in conjunction with a doorstep account, on savings behavior among low-income households. The methodological and conceptual contribution of the study is to explore the causal relationship between financial education and savings and between financial education and financial capabilities, in combination with branchless banking.

Our results show that the financial education intervention had a significant impact on savings. Individuals who received the financial education training saved 29% (\$27) more than the control group. These savings appear to come in part from a decrease in temptation goods by 32%. Moreover, we find improvements on attitudes related to financial planning, but we do not find impact on financial knowledge or time preferences. These findings suggest that financial education can improve savings outcomes, even if it does not affect overall financial literacy or deep preference parameters.

### **The Economics and Psychology of Long Term Savings and Pensions: A Randomized Experiment among Low-Income Entrepreneurs in Maharashtra, India**

*Shailendra Singh Bisht, IBS Hyderabad, India; Anup Kumar Roy, Institute for Financial Management & Research, India*

There has been growing interest in the potential of pure savings to help households achieve better financial outcomes. This has been partly motivated by research in behavioral economics which demonstrates that individuals might have self-control problems, or time-inconsistency.

Our research uses a two-part randomized trial to study the adoption of a long-term savings instrument. The target population consists of low-income households in Maharashtra with access to banking facilities. In the first part, we exposed a fraction of our target population to three framing treatments that delivered information in different ways. While the interventions revealed a clear interest in take-up, they also revealed large institutional hurdles for households. In the second part of the study, we methodically lift these hurdles by offering incentives and assistance to analyze their impacts on take-up.

The results from the first experiment reveal very small but significant effects of the marketing treatments. Further, the treatments that offered less flexibility indicated greater take-up implying that self-control problems might play a role in financial decisions. The second experiment revealed that immediately delivered incentives encourage adoption much more strongly than delayed incentives. These findings have policy relevance in a fiscal scenario where monetary incentives are being used to increase take-up of commitment savings devices.

## Session 40

### Donor Agency Experiences with Impact Evaluation

*Chair: Christopher Edmonds, Senior Economist, PARD, ADB*

#### Learning from Evaluation at MCC

*Jack Molyneaux, Millennium Challenge Corporation*

The Millennium Challenge Corporation (MCC) was established in 2004 as an independent US foreign aid agency that works in partnership with competitively selected countries that demonstrate a commitment to good governance, economic freedom, and investment in their citizens. MCC emphasizes evidence-based investment decisions and rigorous measurement of results. MCC's first batch of independent impact evaluations, released in 2012, assessed farmer training programs in Armenia, El Salvador, Ghana, Honduras, and Nicaragua. The results of these evaluations were mixed: some revealed important impacts on farm incomes, but found little corresponding impact on measures of household income. Much of the learning from these evaluations was related to how difficult it is to design and implement high quality projects and evaluations. We reflect on MCC's early evaluation policy and practice and the evidence from the first five evaluations that informed what was working and what wasn't working in MCC's evaluation practice. We then discuss how some of these reflections, as well as additional learning from MCC's growing evaluation portfolio, have resulted in concrete action by MCC to strengthen its practice around results measurement, particularly its use of impact evaluations.

#### JICA's Strategy for Impact Evaluation

*Satoshi Shigiya, Deputy Director, Evaluation Department, Japan International Cooperation Agency*

The management of JICA is currently promoting "Evidence based management" to improve the quality of services delivered to developing countries. JICA's Evaluation Department has a strategic role to promote this initiative and believes that impact evaluation is one of the critical and effective tools to promote evidence based management.

While JICA has been conducting impact evaluations since 2001, it still has a long way to go to fully incorporate the concept of evidence based management into its project management cycle. Major internal challenges include resource constraints—especially human resources in the head office as well as field level—and lack of awareness of the value of impact evaluation among the staff. It is also facing general challenges stemming largely from impact evaluation such as external validity and efficient and timely feedback of the result to the operation.

This presentation will provide an overview of JICA, a stocking of its impact evaluations to date and the strategy to promote impact evaluation.

## Session 41

### Evaluating *Pantawid Pamilyang Pilipino Program (4Ps)*

Chair: Julie Dumlao, MCA, Philippines

#### Does Pantawid Foster Dependence or Encourage Work? Evidence from a Randomized Experiment

Aniceto Orbeta, *Philippine Institute for Development Studies (PIDS)*

The *Pantawid Pamilyang Pilipino Program (4Ps or Pantawid)* is the social protection program of the Philippine government. Like conditional cash transfer programs in many other countries it provides cash grants to beneficiaries provided they comply with specified conditionalities such as keeping their children in school, getting health check-ups, and attending family development sessions. One of the main concerns about the program is the fear that its monetary assistance could foster welfare dependency. On this point, a commonly held belief is that income transfers from the Government tend to reduce work incentives and encourage dependency on public assistance.

This paper provides more detailed evidence on the impact of Pantawid on the work efforts of members of beneficiary households using experimental design data from the first wave of the Pantawid impact evaluation. Using a more disaggregated approach yields some intriguing findings than those in the aggregated analysis provided in Nazmul, Friedman and Onishi (2013). Consistent with their results and those of previous studies, we find not surprisingly that Pantawid has indeed no significant negative impact on work efforts. Surprisingly, however, we find that, contrary to the critics' expectations, the CCT program seems to increase the desire to work more on the part of household heads and his female spouse, all adult members 18 years and above and middle-aged workers 35–54 years old. Thus, the present study provides additional evidence to the growing literature showing that there is no evidence on work disincentives of CCT programs among adult members regardless of how one divides up the study population. In regard to child labor, the program significantly reduces the number of hours of work for pay of elementary school-aged children (6–11 years), but it did not significantly affect the incidence of child labor. The call, therefore, for de-funding

#### Estimating the Impact of 4Ps on School Participation of Filipino Children Using Propensity Score Matching (PSM)

Celia M. Reyes, *Philippine Institute for Development Studies*; and Christian D. Mina, *Philippine Institute for Development Studies, Philippines*

The presentation highlights the evaluation of the impact of the conditional cash transfer (CCT) program in the Philippines on school participation of children using propensity score matching (PSM). The presentation commences with a brief background outlining the trends in poverty and education indicators, the significant contribution of the 4Ps, and the rationale for conducting the impact evaluation (IE). Reasons for the application of PSM as the impact evaluation method are provided. The presentation takes off with the steps followed by the authors in coming up with the impact estimates as well as the data sets used. Different versions of the proxy means test (PMT) models and matching algorithms used in conducting the PSM, including the sensitivity analysis of the matching results, are presented. Impact estimates are then discussed and compared with estimates from earlier assessments. Descriptive and correlation analyses that show the comparison between the key characteristics of the 4Ps families and those of the non-4Ps families are also presented. Based on the findings, recommendations on how to make the 4Ps a more effective government intervention in improving the school participation of Filipino children and promoting inclusive growth are discussed. Policy influence concludes the presentation.

## The Impact of the Philippines' Conditional Cash Transfer Program on Consumption and Women's Labor Force Participation

*Melba V. Tutor, Department of Social Welfare and Development, Philippines*

The Pantawid Pamilyang Pilipino Program provides cash grants to poor households qualified on predetermined investments in human capital. This study analyses the program's impact on consumption using the 2011 Annual Poverty Indicators Survey. Average treatment effect on the treated (att) is estimated through propensity score matching methodology. Heterogeneous impacts are examined among the bottom 20% of income distribution.

The study finds that among the total sample, per capita total expenditures is not affected by the program. In terms of monthly per capita, only carbohydrates and clothing significantly increased. As expenditure shares, education and clothing registered significant positive impact. No impact is observed on health spending, both in per capita terms and as a share of expenditure. The impact of Pantawid Pamilya on consumption is more pronounced among the poorest-fifth of households.

Results show that households have responded to program conditionalities but there is very little room to improve consumption of other basic needs. The recent program modification of increasing education grants to older children and extending support up to secondary school completion will help households sustain induced behavioral changes over time. The stronger impact on the poorest-fifth of households underscores the need to improve the targeting mechanism to address leakage issues.

One of the cornerstone strategies of conditional cash transfer (CCT) programs to address intergenerational poverty rests on the primary role of women in managing household affairs. CCTs target mothers as direct program beneficiaries and make them drivers of increased demand for essential services like health and education. The PSM methodology is used to examine the effect on women's employment. Several variations of matching techniques are implemented to check robustness of the results.

## Session 42

### Infrastructure Matchmaking Clinic

*Chair: TBC*

A presentation of 3–4 selected ADB infrastructure interventions, with a moderated open discussion of how to conduct impact evaluations of these projects. Academics with an interest in being involved in these studies are encouraged to attend to meet directly with the ADB project team leaders.

## Session 43

### Policies for Better Maternal and Reproductive Health

*Chair: Maya Vijayaraghavan, Senior Evaluation Specialist, IED, ABD*

#### Do Skilled Birth Attendants Improve Birth Outcomes? Evidence from Nigeria

*Edward Okeke, MD Ph.D.*

We present preliminary results from an ongoing 3ie-funded evaluation of a skilled birth attendant program in Nigeria. This unique program, known as the Midwives Service Scheme (MSS), deployed

more than 2,000 midwives to 652 rural primary health care facilities in early 2010. To evaluate the effects of this program on utilization of maternal, neonatal and child health (MNCH) services and on health outcomes we are surveying more than 8,000 women and 400 primary health care facilities across twelve states in Nigeria. We are also carrying out in-depth qualitative interviews in three states. In this presentation we will discuss some of the challenges in evaluating a large government program that was not designed with an evaluation in mind; we will also present some preliminary results from our analysis of the data. At the moment we have data from five states and expect to have data from at least five more states by September. The results of this innovative evaluation of a large government-led supply-side program will inform policy making around efforts to reduce maternal and infant mortality in low- and middle-income countries.

### **Measuring What Works: Using Difference-in-Difference Techniques to Measure Impact of Women’s Groups on Maternal Health Uptake in Rural Nepal**

*Ram Chandra Silwa, Green Tara Nepal a Nepalese NGO, Nepal; Sheetal Sharma, PhD candidate, Bournemouth University; Edwin van Teijlingen, Bournemouth University; and Padam Simkhada, The University of Sheffield*

There is a paucity of high quality evaluations in maternal health in low-income countries. This paper reports a community-based health-promotion intervention in Nepal targeting rural women, in groups and on a one-to-one basis. Measuring improvement in such quasi-experimental designs is challenging due to confounding factors, including socioeconomic status and geography.

Three surveys were conducted at baseline (2008), midline (2010) and final (2012). In total 1,236 women of childbearing age, with their last child ( $\leq$ two years), were interviewed in this controlled before-and-after, cross-sectional study. Difference-in-Difference estimation was used to assess the effects of the intervention on the outcome variables while controlling for a constructed wealth index and personal characteristics.

The proportion of women who sought antenatal care at least once during their last pregnancy increased from 88.7% to 98.0% (OR=4.08, [95% CI 1.14–14.62]  $p < 0.05$ ). In addition, a greater proportion of women reported taking iron/folic acid during pregnancy 79.3% to 96.0% (OR=2.76 95% CI 1.01–7.56],  $p < 0.05$ ) as well as seeking postnatal care 73.8% to 85.9% (OR=2.01 [95% CI 0.98–4.10],  $p < 0.05$ ).

Clinically relevant (but not statistically significant) increases were seen in the proportion of women with a prenatal visit in the first trimester (54.2% to 60.5%), who attended more than four visits (72.4% to 86.0%). Institutional delivery (71.4% to 76.0%); and had a skilled attendant at birth (75.4% to 82.0%).

This evaluation showed that this community-based intervention has a greater impact on the uptake of community-based antenatal care and less so on facility-based care. Confounders (low educational level, low household income, and multiparity) are risk factors for low use of intrapartum services. Hence additional measures are needed to improve the use birth facilities and skilled birth attendant.

### **Impact Evaluation of Women and Girls’ Empowerment Projects in Some Egyptian Villages**

*Salma Galal, Egypt Research and Evaluation Network, Egypt*

Multiple organizations are working in a comprehensive approach on women and girls’ empowerment in diverse sites in Egypt. An international NGO implemented projects in 10 villages. The activities comprised improving education in schools, providing saving groups and small loans, issuing identity cards for women, raising awareness of the community against early marriage, female genital mutilation (FGM) of girls and enhancing their education targeting 10,000 persons.

Method: Impact evaluation was done two years after ending the projects using a combination of quantitative, qualitative methods and the most significant changes stories. Since no baseline surveys were done in the communities at the beginning of the projects, the impact evaluation relayed on the comparison between a sample of intervention and control villages. A sample of 206 women and 198 girls were randomly chosen from five intervention villages and 91 women and 100 girls from two control villages.

Results: More girls continued their secondary education and fewer girls were beaten in intervention villages. Knowledge increased about early marriage and FGM and practices decreased in intervention and control villages. FGM as a taboo issue was discussed in public and between parents. More religious leader adopted the new ideas.

Conclusion: Sensitive issues like women empowerment, battering, FGM and gender activities are best addressed through a holistic approach that takes into consideration the basic needs of the community and by using local resources.

Impact of impact evaluation: New projects implemented baseline surveys to improve the comparison, monitoring, outcome and impact evaluation.

### **Saving Lives of Mothers and Newborns through Capacity Development and Partnership: Experiences from the Safe Motherhood Promotion Project in Bangladesh**

*Masamitsu Kurata, Japan International Cooperation Agency*

This evaluation reports the findings from a quasi-experimental impact evaluation of the Safe Motherhood Promotion Project (SMPP) conducted in the Narsingdi district of Bangladesh. SMPP was undertaken with technical and financial support by the Japan International Cooperation Agency (JICA) to improve the availability and utilization of quality health services for women during pregnancy and childbirth and neonates by targeting both health facilities and communities. Combining survey data originally collected in project area and individual birth records obtained from Bangladesh Demographic and Health Survey as comparison, we calculate difference-in-differences (DID) estimators. The results show the proportion of women who attended at least one ANC by formal provider increased by 23.0% (from 48.9% to 71.9%) in the project area, whilst the national average change was 5.9% (from 50.0% to 55.9%), and thereby they DID impact for this case is estimated as 17.1% (=23.0%–5.9%) at less than 1% significance level. Similarly, we confirm that the DID impact using regional average as comparison is 11.4% at less than 10% significance level. The DID impact for women’s visits to at least four ANC and their use of SBA are also confirmed to be positive, though they are not statistically significant. The paper also highlights that the methodology employed in this evaluation can be widely applied for donor funded projects which usually collect data only in the project area due to several real world constraints. This design allows aid practitioners to conduct inexpensive but rigorous impact evaluation.

## **Session 44**

### **Keeping Kids in School: Evidence from Field Experiments in the People’s Republic of China**

*Chair: Steve Bell, Abt Associates, USA*

Like other middle-income nations, the People’s Republic of China is now entering a critical period of economic transformation. Increasingly, traditional low-wage industry in the People’s Republic of China is being replaced by higher wage and higher valued industry. As the economy shifts and as wages rise,



individuals will need more schooling to hold the necessary skills and knowledge that these kinds of jobs require. Unfortunately, evidence shows that rural students in the People’s Republic of China today—the major part of the People’s Republic of China’s future labor force—are receiving dramatically less schooling than they will need to take on these higher wage jobs. While nearly 90% of urban students enroll in high school, less than 50% of rural students do. Even more concerning, dropout from rural junior high schools is also high, ranging from 20 to 30% in poor rural areas. If this inequality and under-education is allowed to persist, rural students will struggle to find and hold onto good jobs, income inequality will likely remain or increase, and economic growth in the People’s Republic of China is likely to stagnate.

All four papers included in this session present evidence on high rates of rural secondary school dropout as well as evaluations of interventions designed to reduce dropout from the secondary education system. The proposed session includes four studies that use field experiments to analyze innovative approaches to improve student educational attainment in a developing country. The first paper analyzes a randomized controlled trial conducted in two rural Chinese provinces evaluating the short-run effectiveness of Early Commitment of Financial Aid in increasing high school matriculation among poor 9th grade students and increasing intentions to attend high school among poor 7th grade students. The second paper examines the long-run effectiveness of Early Commitment of Financial Aid in reducing junior high dropout and increasing high school matriculation among poor rural students three years after the initial ECFA contract. The third paper presents findings on another issue that is likely a contributing factor to high rates of junior high school dropout in rural People’s Republic of China: pervasive mental health issues that are found to be highly correlated with school dropout. The fourth and final paper summarizes findings from eight large-scale survey studies conducted among 24,931 rural students examining the prevalence and correlates of rural secondary school dropout and conducts mixed-methods analysis based on 52 in depth qualitative interviews with rural students to examine the deeper reasons that may be promoting dropout in these communities.

**Giving Kids a Head Start: The Impact and Mechanisms of Early Commitment of Financial Aid on Poor Students in Rural People’s Republic of China**

*Hongmei Yi, Chinese Academy of Sciences, People’s Republic of China*

Due to financial constraints, poor junior high students may drop out of junior high or fail to matriculate into high school. Early commitment for financial aid (ECFA) may reduce dropout and increase high school matriculation among poor junior high students. In this study, we estimate the impact of two ECFA programs: one at the start and one near the end of junior high school (seventh and ninth grade, respectively). We find that the ninth grade program increases high school matriculation by 16%. We also find that the seventh-grade program increases students’ plans to attend high school by 15% but fails to decrease dropout rates or improve math achievement. In examining why ECFA motivates behavioral change for ninth graders but not seventh graders, we show how seventh graders engage in wishful thinking (they appear to change plans without reference to whether their plans are realistic). In contrast, because better students have self-selected themselves to stay in school through ninth grade, there is a significant and material effect of the intervention on ninth grade students.

**The Long-Term Impact of Early Commitment of Financial Aid (ECFA) on Junior High School Students’ Matriculation into High School**

*Fan Li, KU Leuven, Belgium*

In many developing countries, students from poor, rural areas are less likely to enroll in high school and more likely to drop out of junior high school than their non-poor, urban peers. Policymakers in some countries have sought to encourage poor, rural students to attend high school by providing scholarships,

tuition waivers, and other forms of financial aid. Although studies show that Early Commitment of Financial Aid (ECFA) is one such intervention with positive impacts on increasing matriculation in the short run, little is known about the long-term impact of ECFA. This study makes use of a randomized controlled trial to examine whether providing ECFA (conditional on matriculating into high school) at the start of junior high can reduce dropout from junior high and increase the rate of disadvantaged students' matriculation into high school three years later. We find that for the average student, providing ECFA three years before graduation has very limited effect on increasing matriculation into high school or reducing junior high dropout. However, we also find that providing ECFA does measurably increase the rate of matriculation into academic high school for students who have a higher than average expected cost of going to academic high school.

### **Mental Health and Dropout Behavior: A Cross-Sectional Study of Junior High Students in Northwest Rural People's Republic of China**

*Qinghe Qu, Chinese Academy of Sciences, People's Republic of China*

Junior high dropout rates are up to 25% in poor, rural areas of the People's Republic of China. Although existing studies have examined how factors such as high tuition and opportunity costs contribute to dropout, fewer studies have explored the relationship between dropout rates and mental health in rural People's Republic of China. The overall goal of this study is to examine the relationship between dropout and mental health problems in rural Chinese junior high schools. Correlational analysis was conducted among 4,840 students across 38 junior high schools in rural People's Republic of China. Ordinary least squares (OLS) regressions were used to determine the types of students most at risk for mental health problems (measured with the Children's Manifest Anxiety Scale) and whether mental health problems are correlated with dropout behavior. We find that mental health problems are widespread in the sample of rural children, with 74% of students at risk. The student and family characteristics that correlate with dropout (poverty and low achievement) also correlate with mental health problems. More importantly, even after controlling for these background characteristics, mental health problems remain correlated with dropout rates. This finding suggests that interventions focusing on mental health in rural areas may also help reduce dropout.

### **Dropout in Rural People's Republic of China's Secondary Schools: A Mixed-Methods Analysis**

*Natalie Johnson, Rural Education Action Program, Stanford University, USA*

Students in rural People's Republic of China are dropping out of secondary school at troubling rates. While there is considerable quantitative research on this issue, no systematic effort has been made to assess the deeper reasons behind student decision-making through a mixed-methods approach. This article seeks to explore the prevalence, correlates and potential reasons for rural dropout throughout the secondary education process. It brings together results from eight large-scale survey studies including 24,931 rural secondary students across four provinces as well as analysis of extensive interviews with 52 students from these same study sites. The results show that cumulative dropout across all windows of secondary education may be as high as 63%. Dropout is significantly correlated with low academic performance, high opportunity cost, low socioeconomic status and poor mental health. A model is developed to suggest that rural dropout is primarily driven by two mechanisms: rational cost-benefit analysis or impulsive, stress-induced decision-making.

## Session 45

### Does Community Driven Development Work?

*Chair: Taejong Kim, KDI School of Public Policy and Management*

#### Does Community Driven Development Work? Findings from an Evidence Synthesis Report

*Radhika Menon, 3ie*

Community driven development programs typically involve providing block grants and technical assistance to communities for building local public goods. The active involvement of community members in selecting, designing and implementing projects is the defining feature of these programs.

Donors such as the World Bank have made sizeable investments in community driven development and reconstruction programs. The World Bank alone supports approximately 400 community driven development projects in 94 countries valued at almost \$30 billion. But has community driven development worked in achieving its objectives?

This presentation is based on an evidence synthesis report which draws on impact evaluations of around 20 community driven development programs in low- and middle-income countries.

The presentation will elaborate on the theory of change of community driven development programs and analyze the evidence on impact at every step of the causal chain. It will address several key questions: Have these programs used effective targeting mechanisms? How have they worked in inducing participation? How do factors like program design, length of project, quality of implementation, nature and intensity of training contribute to the impact of the program? How effective have these programs been in reducing poverty and improving welfare? How have they contributed to improving transparency, accountability and building social cohesion? What are the implications for research, policy and practice?

The presentation will use examples and case-studies to illustrate the findings of the report.

#### The Anatomy of Failure: An Ethnography of a Randomized Trial to Deepen Democracy in Rural India

*Kripa Ananthpur, Madras Institute of Development Studies. Chennai, India*

Programs that induce citizen participation to improve the quality of government at the local level are the subjects of large amounts of funding and intense debate. This paper combines a randomized control trial of a citizenship training and facilitation program in rural India, with an in-depth, four-year, ethnography of the intervention to understand the underlying mechanisms of change. The quantitative data show no impact from the intervention. Household and village survey data from 100 treatment and 100 control villages show considerable improvement across a wide variety of governance and participation indicators over time, but the differences in the changes between treatment and control villages are not statistically significant. The detailed qualitative data from a ten per cent subsample allow us to unpack the reasons why the intervention “failed” highlighting the role of variations in the quality of facilitation, lack of top-down support, and difficulties with confronting the stubborn challenge of persistent inequality. However, the qualitative investigation also uncovered subtle treatment effects that are difficult to observe in structured surveys. The paper thus demonstrates that a concerted effort to use “thick description” to uncover process of change using careful and detailed qualitative work can add value to standard impact evaluations.

## Impact Evaluation of a Local Empowerment Project: A Case Study of the Community and Social Development Project in North Central Nigeria

*Hassan Ishaq Ibrahim, Federal University Dutsin-Ma Katsina State Nigeria, Nigeria*

The Federal Government of Nigeria sought and obtained financial assistance from the International Development Association (IDA) of the World Bank Group to support the implementation of Community and Social Development Project (CSDP). The Project has been formulated and implemented from 2009 to 2013 in 26 States and the Federal Capital Territory (FCT) in Nigeria, using the Community Driven Development (CDD) approach. The Project Development Objective (PDO) is to support empowerment of communities for sustainable and increased access of poor people to improved social and natural resource infrastructure services. This paper seeks to; (i) examine the operations and processes that make up the CSDP e.g. targeting mechanism, governance & empowerment, capacity building and social capital infrastructure, environmental sustainability, gender and vulnerability (ii) determine CSDP impacts on participating communities (iii) examine cost effectiveness of CSDP interventions and (iv) assess changes in the well-being of individuals, households or communities that can be attributed to the CSDP. A combination of quantitative and qualitative methods of data collection was adopted in the data collection. A total of 44 communities and 660 households were selected for both benefitting and non-benefitting communities respectively. The Difference-in-Differences (DD) method was used for the impact evaluation. Findings suggest that CSDP implementation using the CDD approach was cost effective, environmentally friendly, and socially inclusive. Above all, the project impacted positively on all benefitting communities.

## Session 46

### Experiences in Producing, Managing, and Using Impact Evaluations

*Chair: K.E. Seetharam, Principal Knowledge Sharing and Services Specialist, RSDD, ADB*

#### Education of Potential Users of Impact Evaluation Results Matters!

*Marija Nashokovska, Foundation Open Society Macedonia, Macedonia*

Introducing the impact evaluation as one of the types of evaluation to an implementer of a development project, especially in situation when the evaluation is commissioned by the donor can be stressful. Generally, implementers are very concerned regarding the outcomes of the evaluation much before the start of the process. They are even more concerned, when they discover that conducting the impact evaluation means repetition of the process twice—once at the beginning of the project and second time at the end of the project. At least, that is my experience of participating in impact evaluation from both sides—as implementer and as impact evaluation specialist.

In order to ensure smooth process of conducting the impact evaluation, education of people working on the implementation of a project that is subject to impact evaluation is a must. Explanation of the steps undertaken during the impact evaluation and their importance for the outcomes of the evaluation process is essential. This is in particular important when it comes to evaluating dynamic social phenomena such as citizen participation in local communities to solve problems for instance.

This paper explores the necessity and structure of educational activities that should be undertaken in advance of conducting impact evaluation in particular for development projects. Tips from practical experience will be included aiming to support the understanding about what information and knowledge needs to be shared with the implementers, so they can be supportive and cooperative in the process.

## Impact Evaluation of Institutional Funding for International Development NGOs: The Case of the UK Civil Society Challenge Fund

*Anne-Marie O’Riordan, Triple Line and the University of Bath*

One aspiration for institutional funding of NGOs and other civil society organizations in international development is to promote experimentation, learning and innovation. A need for good quality feedback on outcomes and impact follows. But generating it in practice is not easy, particularly when grants are relatively small and for disparate activities being implemented across varying timeframes. This paper draws on the case of the UK government Civil Society Challenge Fund (CSCF) to explore the methodological and institutional difficulties that have to be overcome in evaluating impact and results. Section 1 sets the context by describing the origin and design of the CSCF, and review the monitoring and evaluation system it adopted. Section 2, provides a qualitative study of the end-of-project independent evaluation reviews of a sample of ten projects funded under CSCF between 2008 and 2013. Evaluation activity focused primarily on processes and outcomes, which were generally positive. In contrast, evidence produced on end-of-project impact tended to be limited in quantity and weak in quality. Section 4 discusses explanations for this, and considers scope for strengthening evaluation of such programs. More specifically, it (a) reviews the scope for meta-evaluation across institutional funding portfolios using qualitative data analysis software and (b) experimental work into partial blinding of qualitative evaluation to address potential confirmation bias in project level impact evaluation.

## Impact and Beyond

*Geert Geut, Deputy Director IOB, The Netherlands*

Over the last couple of years The Policy and Operations Evaluation Department (IOB) at the Dutch Ministry of Foreign Affairs has accumulated substantial experience with robust impact assessments, meant as building blocks for higher level evaluations of specific policy areas, such as education, SRHR and renewable energy, trying to close the “micro-macro gap”.

The result has been a growing insight in aid effectiveness, at the same time acquiring experience and knowledge about the applicability of concepts such as Theory of Change and Impact Pathways. The robustness of these forms of research has contributed to the formulation of evidence-based policy-making, although there are still major steps to be taken. At the same time, we have experienced constraints when it comes to the application of mixed methods and in terms of efficiency analysis.

Difficult areas and challenges remain: for instance, the direction of causality is not always clear cut, the measurement of complementarities of joint interventions is complex and sometimes arbitrary; assessing the effects of (the absence of) policy coherence requires additional tools (policy simulation), while additionally (in PPPs and private sector programs) is difficult to establish.

The main challenges lie in using heterogeneity for impact analysis, establishing optimum sample size (especially in the case of small n) and the sequential use of analytical methods.

We see the way forward in more work to be done in choosing the appropriate methods for key questions, in addressing specific questions from different stakeholders and choosing the optimum mix and sequence of methods and approaches. Finally, we might have to be more selective in the use of evaluations and look outside the direct realm of development aid.

## Special Sessions

### Impact Evaluations in the Philippines

Short presentations of Impact Evaluations conducted in the Philippines. Media, national researchers, and policy makers are encouraged to attend.

#### Establishing the Philippine Impact Evaluators Network for Rural and Agricultural Development

*Chair: Dario A. Cidro*

The creation and establishment of Philippine Impact Evaluators Network for Rural and Agricultural Development (PIENRAD) is timely and relevant as pressing challenges in agricultural and rural development need to be addressed. The impacts of climate change on agriculture, the globalization of agricultural production and marketing, the debate on organic versus conventional production, the shrinking agricultural extension initiatives, and the impacts of good governance on rural and agricultural development, are just some of the many issues besetting the agricultural and rural sector needing attention. The scholarship and practice therefore of impact evaluation will enhance the adoption of impact evaluation tools and strategies as key to meaningful evaluation. Currently, there are only two evaluation associations in the Philippines but both organizations does not cater specifically to rural and agricultural development. The Philippine Educational Measurement and Evaluation Association (PEMEA) focuses on evaluation in academic setting while the Philippine Development Evaluation Association (PHILDEV) focuses in advocating transparency and integrity in public sector governance. It is expected that PIENRAD will complement to the contribution of these two organizations—PEMEA and PHILDEV- to the scholarship and practice of impact evaluation particularly in the rural and agricultural development sector. Moreover, with the presence of International Initiative for Impact Evaluation (3ie) in the Philippines, the PIENRAD could be an active partner and conduit to 3ie's initiatives in the Philippines. Finally, we hope to institutionalize PIENRAD and become a leading organization in promoting the scholarship and practice of impact evaluation for rural and agricultural development.

##### **Key guide questions for the workshop:**

- 1) How do we increase the Network members?
- 2) Who are the potential partners of PIENRAD within and outside the country?
- 3) How do we secure information on the current needs and challenges of impact evaluators in the country particularly to those working in the rural and agri sectors?
- 4) Who the key scholars, practitioners, and institutions working on impact evaluation in the Philippines? In Southeast Asia?
- 5) How can ADB, 3ie and PIDS help the PÎENRAD?
- 6) What are the academic institutions, government agencies and NGOs working on impact evaluation?
- 7) What is the current status of the scholarship and practice of impact evaluation in the rural and agricultural development sector?



## Poster Session

### **Design Considerations for a 10-Year Alumni Tracking Study: Methods for Measuring the Impact of Social Change**

*Mirka Tvaruzkova, Senior Evaluation Officer, Institute of International Education (IIE)*

This poster discusses the methodological considerations for an innovative ten-year tracking study of a fellowship program funded by the Ford Foundation. The International Fellowships Program (IFP) provided higher education opportunities to individuals from marginalized social groups from 2002–2012. Approximately 4,370 fellows from 22 countries participated in the program. The Institute of International Education (IIE) is implementing a tracking study to analyze the pathways of these alumni over time, capturing their behavioral changes and professional accomplishments, as well as the potential impacts of their fellowship on their communities. In our benchmarking research we found that while most evaluations happen as a program is in process or as it is ending, less research has been done on program impact of international fellowships over time. Yet this is precisely the impact practitioners would like to see, the evidence that their program has promoted long-term change in the lives of their fellows and their communities. Program implementers also share the desire to document impacts that go beyond the individual experience, impacts at the organizational, community or even national and international levels. Our hypothesis is that the International Fellowships Program provided alumni with opportunities that have caused social impacts beyond the individual experience. We examine the social justice impacts of the fellowship and the extent to which IFP alumni have been able to address social injustices not only as individuals, but also as community leaders. Finally, we discuss the role of IIE in contributing to research related to higher education and social justice.

### **Ex-Post Regulatory Impact Assessment in Emerging Economies**

*Jyoti Katiyar, Research Associate, CUTS International*

Compliance with costly, antiquated and duplicative directives of regulatory and government agencies is a burden on regulated entities, especially in emerging economies, impeding overall growth and development. A framework that facilitates systematic and periodic review of regulatory tools and institutions to ensure that they meet intended objectives efficiently and effectively, and that their objectives remain relevant in dynamic landscape, is thus needed.

Ex-Post Regulatory Impact Assessment (RIA) is an evidence based framework supported by empirical findings to assess the quality of regulatory tools and institutions. It is informed by data and findings generated through wide public consultations from relevant stakeholders, thus minimizing biases and highlighting the true costs- direct and unintended- and benefits to all sections of the society. The findings are used to develop better policy alternatives, costs of which are outweighed by their benefits. It thus allows for continuous refinement and improvement in the policy design, tools, and implementation process.

Ex-post RIA of regulatory tools and institutions has benefitted various economies in the past. For instance, in UK, it has led to reconfiguration of sector regulations in the financial and communication sectors. In Australia, it has significantly contributed to the increase in GDP. As a result of intensive stakeholder involvement, it helps in increasing awareness and interest in regulatory management and reform, across the board.

The presentation will discuss the concept and process of ex-post RIA, its experience and benefits, and will call for its adoption by emerging economies in order to modernise the existing regulatory architecture.

## Capacity Building Program (Water and Sanitation) Intervention in Rural Development of Pakistan

*Ume Laila, KDI School of Public Policy & Management*

To examine the impact on incidence of diarrhea as the health outcome, amongst children living in households with access to water and sanitation facilities. The research hypothesis is that safe drinking water and improved sanitation can reduce the incidence of diarrhea among children under five.

Research variables: Diarrhea, Households with access to: a) Flush toilets, b) Ventilated pit latrines, c) Compositing toilets, and Socioeconomic and demographic variables

Data source: DHS survey 2007(95,441) households, and DHS survey 2013 (14,000) households. The DHS includes topics related to fertility levels and determinants, family planning, fertility preferences, infant, child, adult and maternal mortality, maternal and child health, nutrition, knowledge of HIV/AIDS and women's empowerment.

Research technique: First I will compare the prevalence of diarrhea among children less than five years from DHS 2006 and 2013 then I will investigate the causes for the decline, or same situation in the incidence of diarrhea by looking at the changes in the water and sanitation access. I will use the propensity score matching and divide the study in treatment or intervention group with access to improved sanitation, and without access to improved sanitation as control group. This will give me an idea about the mean improvement or decline between 2006 and 2013. I will calculate the average treatment effect. The aim of the method of matching is to construct a comparison group that resembles as much as possible the participant group with respect to observed characteristics that affect both the participation decision and outcome.

## The Critical Role of Monitoring and Evaluations Systems in Impact Evaluations: Lessons from a Case Study

*MacCarthy Honu-Siabi and Desiree Manicom, University of KwaZulu-Natal*

Monitoring and evaluation systems are crucial in providing data from which the impact of policies and programs can be assessed. The importance of monitoring and evaluation (M&E) during and after program implementation has driven many governments, organizations and donors to invest substantial resources in developing comprehensive monitoring and evaluation systems and tools (Cloete, 2009, p. 297). How these systems are designed, to a great extent, influence the types and qualities of data that will be provided. The design of monitoring and evaluation systems can help tailor evaluations to specific needs and uses. Some writers have explained that, even though many resources are invested into the development of M&E systems, not all of them actually get implemented or even if they do, they are only partially implemented (Groene & Branda, 2006). Mahmood et al (2011) pointed to the complexity of M&E systems as one of the difficulties experienced in its implementation. The NGO sector is one of the implementers of social programs and many use M&E systems to assess their program impacts as well as for accountability to funders.

The paper present the empirical evidence through the findings of a master's research project which critically analyze the M&E system of an NGO called SaveAct and its role in impact evaluations. The study uses a case study approach to analyze the M&E systems of an NGO in Pietermaritzburg called SaveAct. In-depth qualitative interviews are being used with the designer, implementers of the system and project beneficiaries to gather information about their experiences.

The study concluded that M&E systems and frameworks should be designed to be simple and clear. They should contain fewer tools and should be easily understood by the people who do the actual data collection. The tools with fewer questions were easier to implement than those with many questions.

The design of M&E tools and frameworks need to bear in mind which questions are to be addressed during impact evaluation so that such data is collected to improve the usefulness of the results.

### **An Introduction to Evidence Gap Maps**

*Birte Snilstveit, 3ie*

A range of organizations are engaged in the production of evidence on the effectiveness of social and economic development interventions, and over 2000 impact evaluations and 300 systematic reviews are now either completed or on-going. This growth in the production of evidence also presents challenges. How can decision makers best access the existing evidence in a specific field when it is scattered around different databases, websites and the grey literature? How can we present research in a format which is useful and accessible for a non-technical audience? And how can we ensure limited resources are spent efficiently and important evidence gaps are prioritized?

Evidence gap maps aim to respond to these challenges and present a new addition to the tools available to support evidence informed policy making. They draw on a range of methods for evidence mapping and synthesis, and present a visual overview of systematic reviews and/or impact evaluations in a sector or sub-sector. Evidence Gap Maps schematically represent the types of interventions evaluated and outcomes reported. They enable policy-makers and practitioners to explore the findings and quality of the existing evidence and therefore facilitate evidence informed decision-making. They also identify key “gaps” where little or no evidence from impact evaluations and systematic reviews is available, and can inform a strategic approach to building the evidence base in a particular sector. This presentation provides an introduction to evidence gap maps and methodology. It will be followed by separate presentations of examples of evidence gap maps produced to date.

### **Agricultural Interventions in Low- and Middle-Income Countries: An Evidence Gap Map**

*Jennifer Stevenson, 3ie*

Agriculture has wide-ranging global impacts which extend to economic growth, poverty reduction, food security, livelihoods, rural development and the environment. The sector provides a key source of income for an estimated 2.5 billion people in low- and middle-income countries. In addition, around 70% of the global extreme poor live in rural areas in low- and middle-income countries, relying directly or indirectly on agriculture for their livelihoods. Thus, interventions in the agricultural sector have the potential to improve the lives of a large number of people.

The evidence base on the effects of agricultural interventions is increasing, but there is a need to make this evidence easily available to decision makers, and to identify important gaps in the evidence which can inform future research efforts. This presentation will outline the findings of an evidence gap map of agriculture interventions in low- and middle-income countries. The evidence gap map presents an overview of what we know and don't know about effective agriculture interventions. It maps impact evaluations and systematic reviews evidence according to 11 key agricultural intervention areas, including access to finance and interventions to promote resilience and adaptation to climate change, and, and important agricultural, income and well-being related outcomes. The presentation will first highlight policy-relevant findings based on the included systematic reviews, before outlining areas with potential for future evidence synthesis, as well as ‘absolute gaps’ where future research could be focused.

## Primary and Secondary Education Interventions in Low- and Middle-Income Countries—An Evidence Gap Map of Systematic Reviews

*Daniel Phillips, 3ie*

Education is not only a direct determinant of an individual's well-being and life-opportunities, but is also fundamentally important in its potential to bring about social change and economic development. Though progress has been made towards increasing the number of children receiving an education in Lower and Middle-Income Countries (LMICs), in many instances increased enrollment has not translated into improved learning outcomes. As a result, in recent years there has been a shift in focus from access to education to learning for all. According to the recent *Education for All Global Monitoring* report (UNESCO, 2013), around 250 million children in LMICs still cannot read, write, or solve basic mathematics problems.

While there is an increasing body of research that examines educational interventions in LMICs, studies typically focus on a limited number of interventions or outcomes. The International Initiative for Impact Evaluation (3ie) has produced an 'evidence and gap map' designed to comprehensively capture systematic review evidence related to educational interventions designed to improve education access and learning outcomes for primary and secondary school children in LMICs. The team carried out a systematic search of published and grey literature which, after screening for inclusion, identified around twenty reviews for the evidence and gap map. This presentation will focus on highlighting key 'gaps', where little or no evidence from systematic reviews is available, and where future research should be focused. It will also outline the policy-relevant findings based on the included reviews.

## What We Know and Don't Know: Evidence Gap Map of Productive Safety Net Programs on Extreme Poverty

*Martina Vojtkova, 3ie*

Despite recent successes in poverty reduction, about 1.2 billion people around the world still live on less than 1.25 dollars a day. An even greater number—1.57 billion people suffer severe deprivations in health, education and standards of living. A lot more needs to be done to go the last mile and fully eradicate extreme poverty, a goal much talked about in the post-2015 discussions. To achieve this, a more evidence-based approach to poverty reduction programming is needed. To identify the best way forward, we first need to find out what we know, and don't know, about what works to get people out of poverty.

This presentation will introduce the results of an Evidence Gap Map of productive safety net programs on extreme poverty and related outcomes. Productive social safety nets are programs that include *livelihood or income generating components* to expand market opportunities and help stabilize consumption, increase and diversify incomes, build and protect assets, and improve food security of extreme poor individuals, households, and communities. The Evidence Gap Map presents a visual overview of existing and ongoing systematic reviews and impact evaluations on productive safety nets, organized along 6 broad intervention categories such as financial services, land reform and microenterprise support services, and important poverty alleviation outcomes along the causal chain. The presentation will focus on policy-relevant findings based on the included systematic reviews, and highlight key 'gaps' in both systematic review and impact evaluation evidence that can inform future research commissioning.

## **Creating a Public Impact Evaluation Repository—Lessons Learned and a Review of Evidence Collected**

*Drew Cameron, 3ie*

In May of 2014 3ie launched its comprehensive Impact Evaluation Repository. This repository is the first attempt to systematically identify and index all published impact evaluation studies and make the index publicly available to policy makers and development researchers. To build this repository, we used various search methods, including snowball sampling, crowdsourcing, and an extensive systematic search and screening of more than 45 online academic databases, organization websites, search engines, and research libraries. We screened more than 60,000 studies against a pre-specified set of inclusion criteria, and indexed the relevant meta-data. The Impact Evaluation Repository now includes nearly 2,400 records of evaluations in low- and middle- income countries that can provide essential evidence on what works in international development.

This paper begins by reviewing historical experiences with the development of research repositories, particularly for program evaluation and effectiveness studies. We then explain the different mechanisms through which repositories can benefit evidence-based policy making, and draw lessons for systematic search and screening processes based on our experience developing this repository. The majority of the paper provides an analysis of the state of impact evaluation evidence for international development and trends over time. We also present specific analysis of the state of evidence for the sectors of infrastructure, climate change, environment, natural resources, social development, rural development, financial inclusion, and institutional and policy reform. We present analysis of this evidence base specific to East Asia and the Pacific regions (including Southeast Asia). We conclude with recommendations for future investments in impact evaluation.

## **Conducting Impact Evaluation of Integrated Programming in Internal Displaced Persons (IDP) Settings**

*Sharon Low, MPH; Jean-Francois TRani, PhD; Qihui Chen, PhD*

**Background:** Somalia has been in a state of conflict since the 1990s. The population has been subjected to forced displacement and armed conflict, with children being most at risk of violence.

**Study design:** The study will quantify health outcomes and perception of the well-being of the women and children, with focus on education and aspirations in determining protection and reduced risks in being abuse, e.g. early marriages, forced labor as militia. This is a pilot study with baseline in 2014 and end-line in 2015.

**Sampling and data collection:** Sample size was based on a balance of logistical constraints unique to IDP populations and desired precision for outcomes. In this case, cluster is synonymous to camp and was selected randomly. Households were selected within each cluster using randomized household walk. Number of households per cluster at point of data collection was recorded to adjust for study power. Focus group discussions and in-depth interviews to understand mechanisms of protection will be explored during end-line, subject to security situation.

**Data analysis:** Intervention and control groups will be compared using propensity score matching (PSM), and combined with Difference in Differences (DD). PSM is regression-based and estimates “program effect” by eliminating observable confounders on which comparisons are made. If DD-PSM is available, then DD will take care of fixed, unobservable confounders.

**Conclusion:** Debates around effectiveness of humanitarian action are ongoing with little evidence on impact of integrated programming. The role of this study is to call attention to IDP and the feasibility of adapting sampling and data collection without foregoing the rigor needed.

## Health Policy Changes and their Impact on Equity in Health Financing in India

*Swadhin Mondal, Institute for Studies in Industrial Development (ISID), New Delhi, India*

Catastrophic healthcare payments, horizontal inequity, and progressivity of out-of-pocket spending (OOPS) are three crucial issues in health financing. In many countries, these issues have been addressed through health sector reforms. This study examines the impact of health policy changes on equity of financing among households by using four successive rounds of national sample survey (NSS) data on consumer expenditure in India. The horizontal and vertical equity of healthcare payment was measured by Wagstaff and AJL (Aronson, Johnson, and Lambert) decomposition method of redistributive effect (RE). This study shows that the vertical effect (V) of OOPS on income redistribution increased by 15% between 1994 and 2004, and then decreased by 80% in 2012. Horizontal inequity (H) and re-ranking have fallen over the period. Also, RE decreased by 63 per cent between 2004 and 2012. The value of V is lowest in 2012. This implies that government-funded healthcare services have a higher impact on low income group and produce higher equity in OOPS. This study concludes that the health policy changes made by the Government of India, especially 2005 onwards, have had a positive impact on equity in out-of-pocket payment and income inequality.

## Monitoring and Evaluation of Business Development Support to Small and Medium Enterprises: Design and Implementation of PRIME (Pioneering Real-Time Impact Monitoring and Evaluation)

*Fédes van Rijn, LEI-Wageningen University and Job Harms, ESE Erasmus University*

Private sector development is seen as a key contribution to economic growth and poverty reduction in low- and middle-income countries. For the Government of the Netherlands, who invests annually invests €420 million, this is a key pillar in investments in development aid. Two major programs are the Centre for the Promotion of Imports from Developing Countries (CBI) and PUM—The Netherlands Senior Experts.

Therefore, the reasons for evaluating this program are obvious. However, impact evaluation of these programs faces many trade-offs such as questions of attribution versus contribution, internal versus external validity, accountability versus learning, and impact evaluation as an internal or external process. This presentation will focus on finding a balance between these trade-offs using the PRIME program as an example, an innovative five-year impact evaluation partnership between CBI, PUM, LEI-Wageningen University and the Erasmus School of Economics.

The PRIME program is designed to develop an approach to data collection to monitor key indicators in the supported SMEs to assess the impact of PUM and CBI on development, and to give them more insight into the effectiveness of their support and the opportunity to learn about how they can manage for greater impact. Part of the presentation will focus on the design of the program as such; the set of key indicators that will be monitored and the peer-reviewed process of selecting these based on ongoing initiatives such IRIS and an extensive literature review. Another part of the presentation focuses on lessons learned so-far on methodology design and research implementation.

One of the main activities so far focused on the development of the indicators that will be used to monitor and evaluate CBI and PUM support. These indicators were developed based on:

- an extensive literature review: In the course of 2013 and 2014, the team reviewed a rich body of empirical studies and obtained valuable insights in the diverse ways in which prior economic studies have attempted to measure the impacts of advisory services on SMEs.
- Intensive discussion on the TOC of both programs
- Evaluating all M&E data currently collected and identifying the gaps.



## MAKING IMPACT EVALUATION MATTER

---

Two other interesting activities concern the kickoff of two case studies in the tourism sector in Bolivia and the aquaculture sector in Indonesia.

In the first quarter of 2014 the case study methodology was further elaborated.

Both missions are scheduled for the second quarter of 2014. Preparation included some very interesting interviews with staff, program managers, participants and experts from PUM and CBI

