

Data Collection in M&E

Aniceto Orbeta, Jr

Philippine Institute for Development Studies

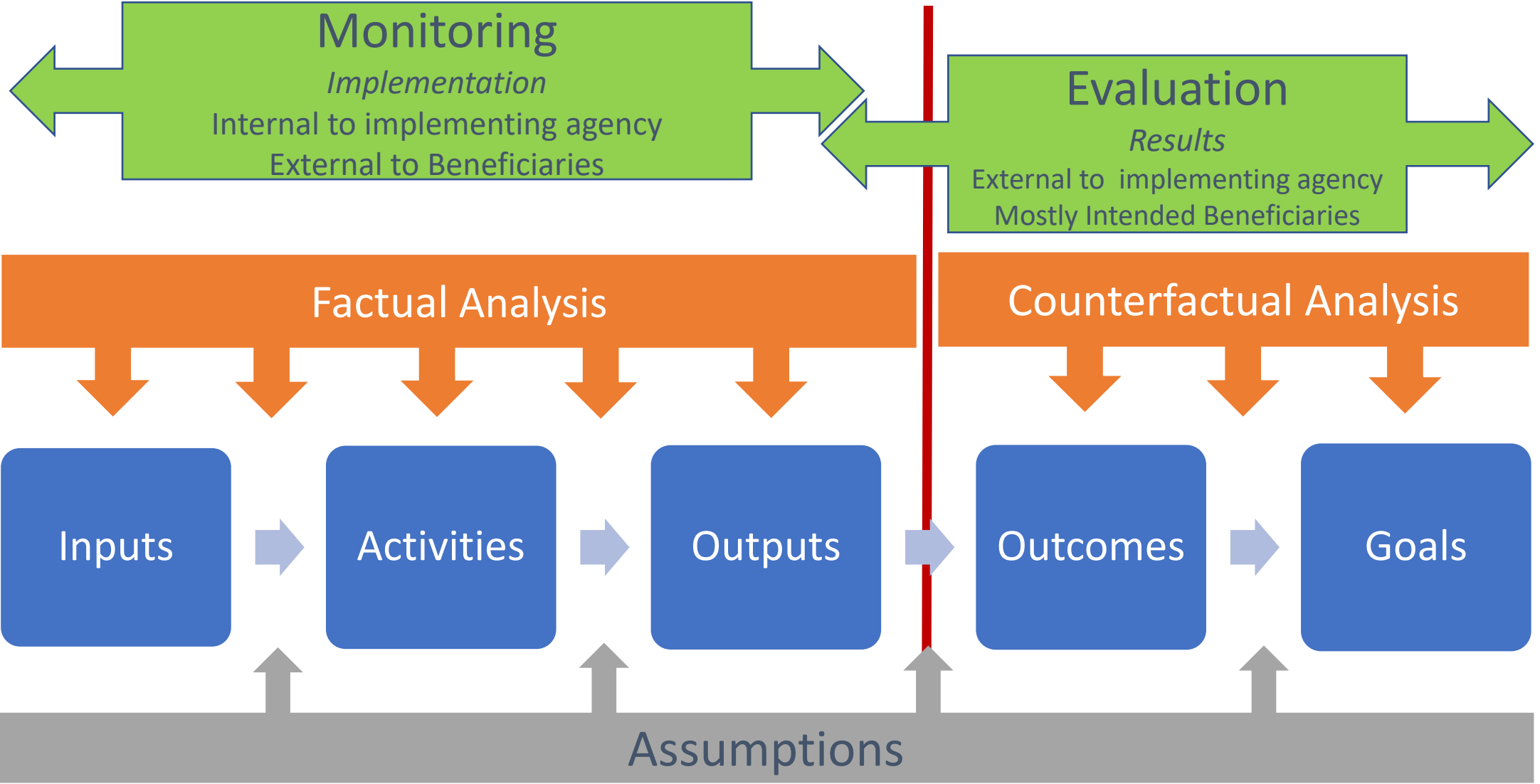
M&E Network Webinar Series

August 25, 2021

Outline

- TOC and Analysis Types in M&E
- Causality and the Evaluation Problem
- Data Collection Issues
 - Existing (secondary) data
 - New (primary) data
 - New data sources
 - Examples from impact evaluations

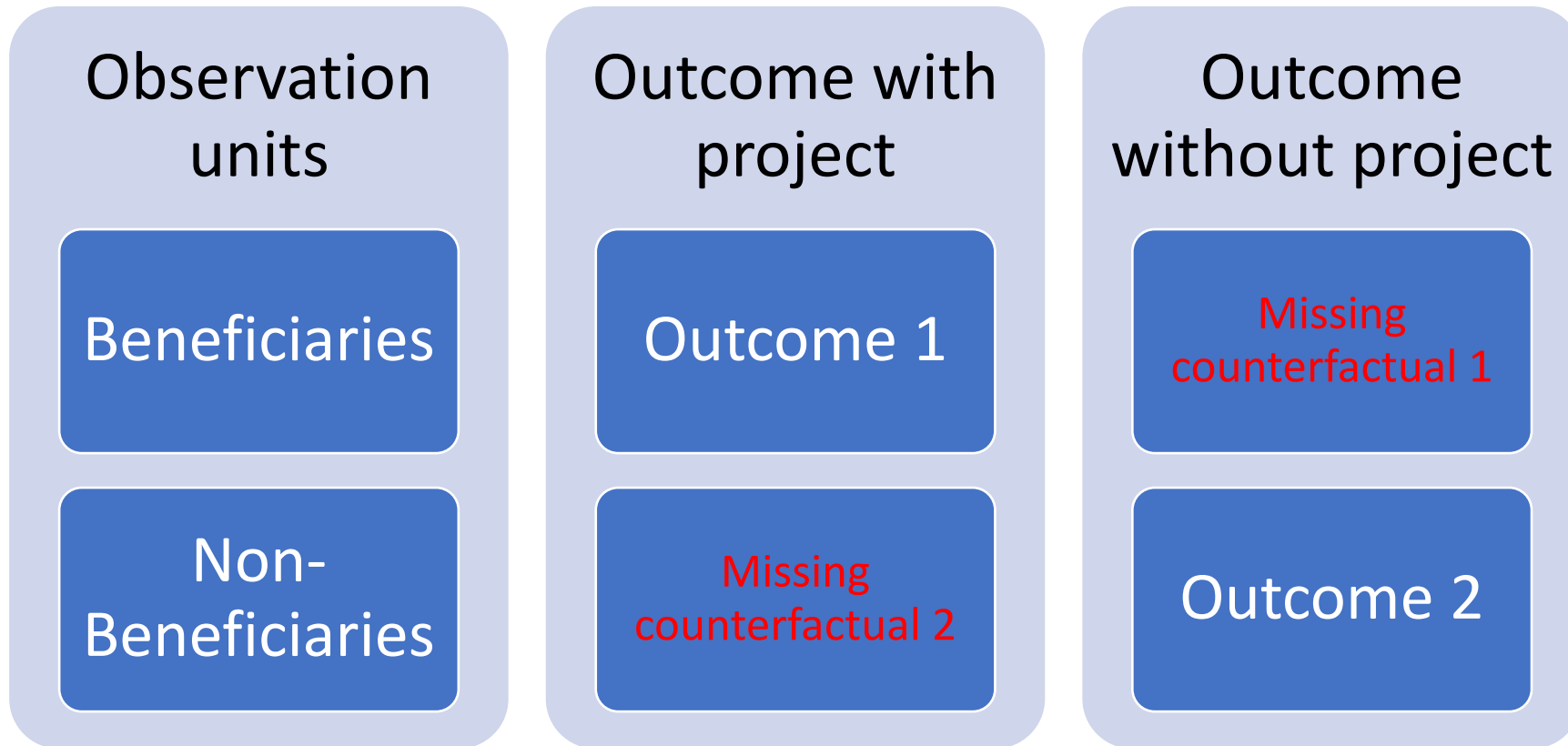
Theory of Change and Analysis Types in M&E



Causality and the Evaluation Problem

- Causality and attribution is central in impact evaluation
 - We want to say program P **causes** a change in outcome of interest Y
- Causal impact requires at least two outcomes (from two causes)
 - Outcome Y1 when received program P
 - Outcome Y0 with status quo (not receive program P) (called **counterfactuals**)
 - $\text{Impact} = Y1 - Y0$
 - This should happen at same instant, for the same subject
 - This results in a missing data problem
 - This is the evaluation problem

There are two missing counterfactuals



All impact evaluation methods offer solutions to this missing data problem

Data Collection Issues

- What data items are needed?
 - TOC the basis for defining indicators to be measured (theory-based evaluation); indicators across the causal (result) chain
 - Evaluation questions
 - Evaluation design and methods

Data Collection Issues

- Sources of data
 - Existing (secondary) data
 - Census
 - Nationally-representative surveys
 - Administrative data
 - New (primary) data
 - New data sources

New (Primary) Data Collection Issues

- Sampling and Sample Size (Power calculation)
- Data generation instrument (questionnaire)
- Fieldwork – conduct of interviews
- Quality control procedures
- Data capture and storing

New Data Sources

- Internet as data source
- Use of unstructured data – e.g., web scrapping, text-as-data
- Device generated data – e.g., google and apple mobility data; financial transactions; messages sent
- Online surveys

Pandemic and Data Collection

- Impact on representativeness
- Impact on how to ask questions
- Impact on field operations

Data Collection Issues: Examples from IEs

Microfinance

Kondo, Orbeta, Dingcong, Infantado; ADB, 2008

- **IE Design:** Difference-in-Difference (DID)
- **Target Population:** Microfinance borrowers
- **Sample:** 2,276 households; 116 barangays; 31 munis; 26 prov
- **Survey Period:** October 2006-Jan 2007
- **Outcomes of interest:** Household welfare (income, consumption, savings, assets)
- **Information gathered**
 - Household questionnaire
 - Barangay profile questionnaire
 - MFI Profile questionnaire

Pantawid Pamilya: Wave 1

WB, DSWD, 2012

- **IE Design:** Randomized Control Trial (RCT)
- **Target population:** Poor Households 2.5 years in program (treatment: received benefits in Jun 2008 – April 2009; control: receive benefits after 18 months)
- **Sample:** total of 3,742 households, 1,418 used in analysis rest are for spillover analysis; 4 provinces, 8 municipalities; 130 barangays; 65 treatment, 65 control
- **Survey period:** Oct 2011 - Feb 2012
- **Outcomes of interest:** education, health, beneficiary behavior, coverage of social protection programs
- **Information gathered**
 - 4 Household modules (Household info, Mother, school-age children, 0-5 years old)
 - 1 School module
 - 1 Health facility module
 - 1 Barangay module
 - 1 LGU module

Pantawid Pamilya: Wave 2

Orbeta, Abdon, del Mundo, Tutor, Valera, Yarcia; DSWD, 2014

- **IE Design:** Regression Discontinuity Design (RDD) (using PMT score as sorting variable)
- **Target population:** Households around the poverty threshold; Treatment households 2-4 years in program
- **Sample:** 5,041 households; 26 provinces, 30 municipalities (10 for each of Luzon, Visayas, and Mindanao); 160 barangays; 3,108 households (62%) within PhP6,250 bandwidth, 1,933 (38%) outside
- **Survey period:** October-December 2013
- **Outcomes of interest:** education, health, child labor, employment of adults, household spending, outlook for the future
- **Information gathered**
 - 5 Household modules (Household info, Mother-reproductive, School-age Children, 0-5 year old children; Cognitive assessment for school-age and women)
 - 1 Barangay Module

Pantawid Pamilya: Wave 3

Orbeta, Melad, Araos; PIDS, DSWD, 2021

- **IE Design:** RDD
- **Target population:** Households around the poverty threshold; treatment households in the program from 2-9 years
- **Sample:** 6,775 HH (3450 treatment, 3325 comparison), covered 30 city/municipalities nationwide; 6 village per municipality; 25 provinces
- **Survey Period:** November 2017 to February 2018
- **Outcomes of interest:** education, health, child labor, employment of adults, household spending, outlook for the future
- **Information gathered**
 - 4 Household Modules (Household Info, Mother -reproductive, school-aged children, 0-5 years-old children)
 - 1 health facilities module
 - 1 barangay module

Kalahi-CIDSS (MCC implementation)

Beatty, BenYishay, King, Orbeta, Pradhan; IPA, MCC, 2018

- **IE Design:** Paired Clustered RCT; paired, used muni characteristics (poverty incidence, population size, number of barangays, land size) for pairing
- **Target Population:** Municipalities with poverty incidence of 33-69 percent
- **Sample:** 198 munis (99 pairs of treatment control munis)
- **Survey Period:**
 - Baseline: April-June 2012
 - Interim: Feb-June, 2014
 - Third-round: July-October 2015
- **Outcomes of interest:** Social Capital; household welfare from subprojects
- **Information gathered**
 - **Baseline:** Household, barangay, KII with officials, FGD with barangay residents
 - **Interim:** Household, barangay, structured community activity, barangay assembly observations
 - **Third-round:** Household, barangay, SCA, Coordinating Team, Municipal, FGD with residents

Sustainable Livelihood Program

Orbeta, Ballesteros, Reyes, Paqueo, Corpuz; DSWD, PIDS, 3ie, 2020

- **IE Design:** Matched Single Difference (match using sex, age, education of household head and spouse, household size, per capita household income)
- **Target Population:** Pantawid Pamilya Beneficiaries
- **Sample:** 2,592 households; 39 municipalities/cities; treatment: Pantawid Beneficiaries availed of SLP assistance in January-June 2018; comparison: matched SLP-eligible Pantawid beneficiaries who are not SLP beneficiaries; using pre-intervention characteristics from Listahanan 2
- **Survey Period:** Feb 2020 to July 2020
- **Outcomes of interest:** hours of work, household per capital income, household savings, household expenditure
- **Information gathered**
 - Household module
 - SLP participant questionnaire
 - Business profile questionnaire
 - Municipal profile questionnaire

Judicial Reform: Small Claims Procedure

Orbeta, Paqueo, Sidiqui; IPA, 3ie, 2021

- **IE Design:** RDD (using amount of claim as sorting variable)
- **Target Population:** Small claim cases
- **Sample:** 15,844 small claims cases; 82 courts
- **Admin data:** eCourt records augmented by court information on claim amounts not available in eCourt records
- **Outcomes of interest:** Case duration, case disposal rate
- **Information gathered**
 - Timing of case activity, claim amounts

Judicial Reform: eCourts

Orbeta, Paqueo, Sidiqui; IPA, 3ie, 2021

- **IE Design:** matched DID
- **Target Population:** All cases
- **Admin data:**
 - Court -level analysis: 305 eCourts, 259 non-eCourts
 - Case-level analysis: 116 eCourts and 87 matched non-eCourts with digitized court dockets
- **Outcomes of interest:**
 - **Court-level analysis:** Clearance rate, Disposition rate
 - **Case-level analysis:** Case duration
- **Information gathered**
 - Timing of case-level activity
 - Court level case flow

Judicial Reform: Continuous Trial Procedure

Orbeta, Paqueo, Sidiqui; IPA, 3ie, 2021

- **IE Design:** Event study (introduction of CTP as events)
- **Target Population:** Criminal cases
- **Admin data:**
 - **Case level analysis**
 - eCourts Data: 544,226 cases; 336 courts
 - Continuous trial monitoring system (CTMS) data: 890,609 cases; 1635 courts
 - **Court-level analysis**
 - Monthly case flow from 2013-2019
- **Outcomes of interest:**
 - Court-level analysis: Clearance rate, Disposition rate
 - Case-level analysis: Case duration
- **Information gathered**
 - Timing of criminal case-level activity
 - Court-level criminal case flow

Thank You