SESSION 8

HOW BIG DATA CAN STRENGTHEN DEVELOPMENT EVALUATION: OPPORTUNITIES & CHALLENGES

Michael Bamberger
OUTLINE

1. The challenges facing development evaluation
2. Data analytics: powerful new tools for the analysis of evaluation data
3. How big data can strengthen evaluations
4. Caveat emptor – understand what you are buying
CHALLENGES FACING DEVELOPMENT EVALUATION
DESIGN CHALLENGES

Many theories of change do not address complexity and emergence

Identifying a credible counterfactual

Identifying unintended outcomes

Developing complexity-responsive evaluation designs
DATA COLLECTION CHALLENGES

a. Cost and time of data collection
b. Collecting data on difficult-to-reach groups
c. Monitoring implementation processes and behavioral change
d. Collecting qualitative data
e. Quality control of data collection and analysis
g. Collecting data on geographic and longitudinal impacts of projects
h. Collecting data on the longitudinal dimensions of projects
i. Sample design
DATA ANALYSIS CHALLENGES

a. Integrating multiple sources of data into a common platform
b. Reducing costs, time and difficulty of data analysis
c. Providing real-time feedback from the analysis
d. Analysis of large data sets
e. Analysis of data from complex programs
f. Analysis of qualitative data
g. Predictive analysis
h. Causal analysis
THE TOOLS OF BIG DATA SCIENCE
The Four Stages of the Data Analytics Cycle

A. DESCRIPTION AND EXPLORATION
- Documenting what is happening
- Creating integrated data platforms
- Identifying new patterns
- Data visualization

B. PREDICTION
- What is likely to happen?
- Which groups are likely to succeed and fail?
- Data visualization

C. DETECTION
- Identifying outliers and groups likely to fail
- Providing actionable information, often in real-time
- Data visualization

EXPERIMENTAL
and PREDICTIVE MODELING and MACHINE LEARNING

EVALUATION/PRESCRIPTION
- Explaining why things happen
- Recommending how to improve performance
- Data visualization
THE DATA ANALYTICS CYCLE

A. Descriptive and exploratory analysis
   – What is happening now?

B. Prediction:
   – What is likely to happen?

C. Detection
   – Who is likely to succeed and fail?

D. Evaluation/ prescription
   – Why things happen
   – Recommendations: how to improve
THE TOOLS OF BIG DATA ANALYTICS

1. Data exploration
2. Descriptive analytics
3. Prediction, predictive analytics and data mining
   - Regression
   - Bayesian analysis
   - Machine learning
   - Neural networks
   - Decision trees
4. **Machine learning** (Artificial intelligence)
   - Pattern recognition
   - Neural networks
   - Clustering

5. **Social media analysis**
   - Themes, trends and network searches
   - Trending topics
   - Behavior [e.g. registering to vote, vaccinating children,]
6. Social network analysis
   - Structure of personal, community, political and other networks
   - Analysis of power, influence and communication flows
   - Defining network structure and measuring change

7. Experimental and quasi-experimental designs:
   - On-line, digital experiments
   - Quasi-experimental designs in the field (propensity score matching)
HOW DATA SCIENCE CAN STRENGTHEN EVALUATION PRACTICE
Stages of the evaluation cycle

1. Project identification and appraisal
   - Creating integrated data platforms

2. Planning and designing the evaluation
   - Addressing complexity

3. Project implementation monitoring
   - Real-time feedback

4. Continual feedback from stakeholders and vulnerable groups

5. Project completion report
   - Monitoring and addressing emergence

6. Planning and evaluating project sustainability

THE PROJECT EVALUATION CYCLE AND POSSIBLE INTERVENTION POINT FOR BIG DATA
BIG DATA APPLICATIONS AT EACH STAGE OF THE EVALUATION CYCLE

1. Project identification and appraisal
   - Satellite images
   - Social media analysis
   - Sociometric analysis
   - Crowdsourcing
2. Planning and designing the evaluation
   - On-line theory of change
   - Using satellite images and remote sensors to strengthen sample design
     • Strengthening comparison group designs through propensity score matching
   - Using systems mapping to evaluate complex programs
3. Monitoring project implementation
   - Early warning systems
     • Social media – real-time feedback
     • Satellite tracking of population movements, drought etc
     • GPS mapping
   - Implementation process tracking
     • Real-time feedback
     • Video/audio feedback on meetings
     • Satellite tracking
   - Detecting priority intervention groups and areas
     • Predictive analytics and detection of problem cases
     • Social media
3. Monitoring (Continued)
   - Collecting contextual data
     - Satellite can track changes and movement over large areas
     - Crowdsourcing provides rapid feedback from large groups
   - Quality control of data collection
     - Quality control apps for GPS-enabled phones:
       - Tracking location of interviewers
       - Random audio control to listen-in to interviews
3. Monitoring (Continued)

- Sample selection for evaluation studies
  - Using satellite images to strength comparison groups
  - Using satellites to select low-income households based on roof materials

- Monitoring behavioral change
  - Social media analysis
  - Analysis of audio and video images
  - Remote sensors to track use of water, sanitation and other services.
4. Sustainability analysis

- Obtaining longitudinal data sets from satellite images, social media and integrated data platforms
- Periodic short phone surveys to update information on the status of project infrastructure, social organization, etc.
- Request staff and project participants to take photos of project infrastructure
- Create user groups on Facebook and other media to provide feedback
CAVEAT EMPTOR
UNDERSTAND WHAT YOU ARE BUYING

- Extractive versus inclusive uses of big data
- The fallacy of large numbers
- Correlation and causality
- Different perspectives on theory
- Quality control:
  - Bias
  - Exclusion
  - Construct validity
  - Working with proxy variables