Evaluation of Market Transformation Programs: An Introduction

Ken Tiedemann, BC Hydro

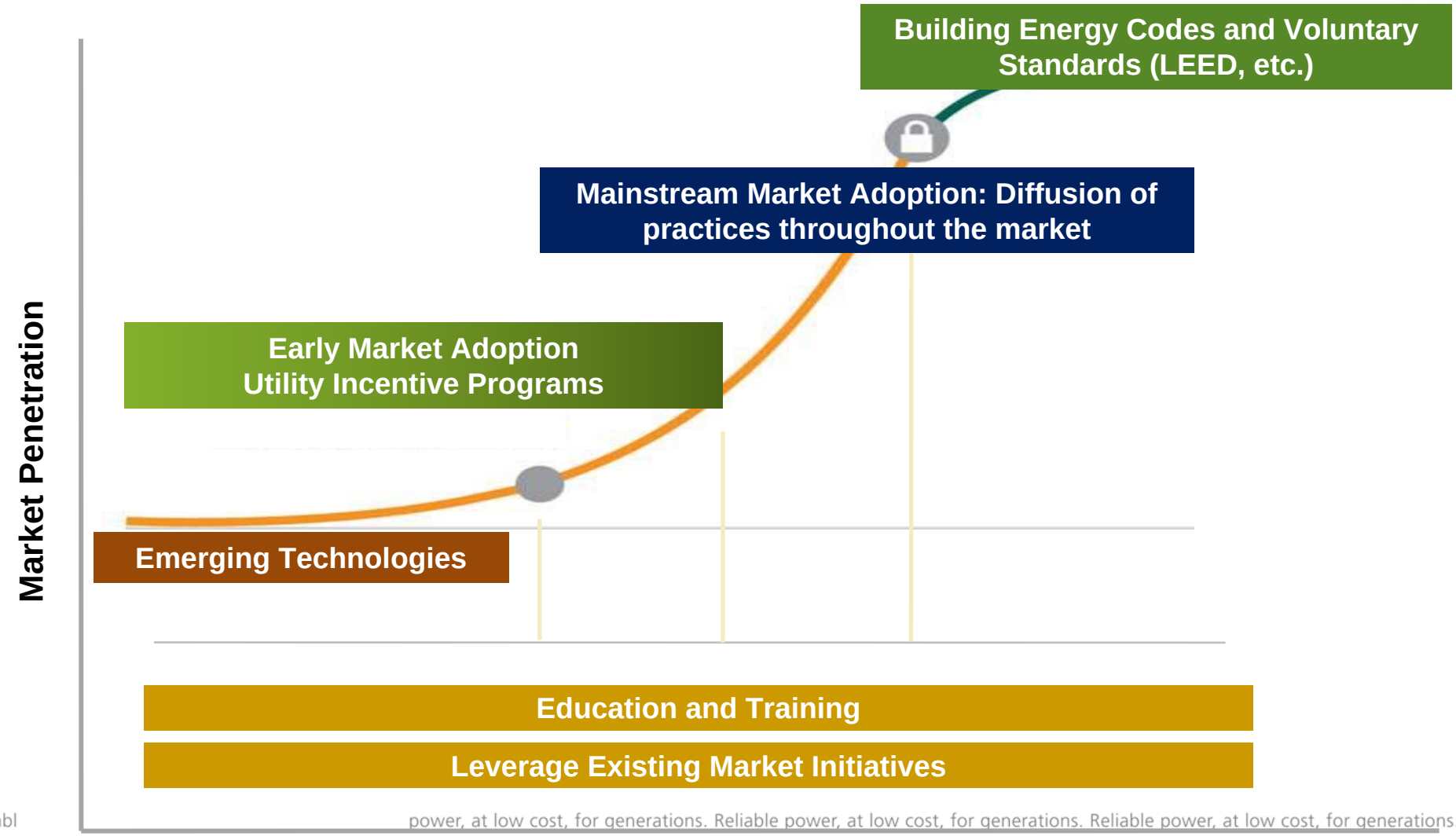
ACEEE Market Transformation Conference, April 2011, Washington, DC

BC hydro

Overview

- Module 1. Evaluation principles
- Module 2. Types of program evaluations
- Module 3. Program theory and program logic

Market Transformation: Adoption Curve



Evaluation of Market Transformation Programs

Evaluation Principles

BC hydro

Some Definitions of Evaluation

- "… the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming," Michael Quinn Patton, Utilization-Focused Evaluation, 3rd ed.
- "… the systematic assessment of the operation and/or the outcomes of a program or policy, compared to a set of explicit or implicit standards…," Carol H. Weiss, Evaluation: Methods for Studying Programs and Policies, 2nd ed.
- "… the use of social research procedures to systematically investigate the effectiveness of social intervention programs…," P. H. Rossi, H. E. Freeman and M. W. Lapse, Evaluation: A Systematic Approach, 6th ed.

Evaluation Is …

- … a systematic assessment of a program
- … an examination of program processes and program outcomes
- … an examination of a program's achievements of goals and objectives
- … an assessment of other impacts and effects which might have been unexpected
- … and, hopefully, a contribution to the improvement of programs

Evaluation Purposes

Program evaluations are often done for the following reasons:

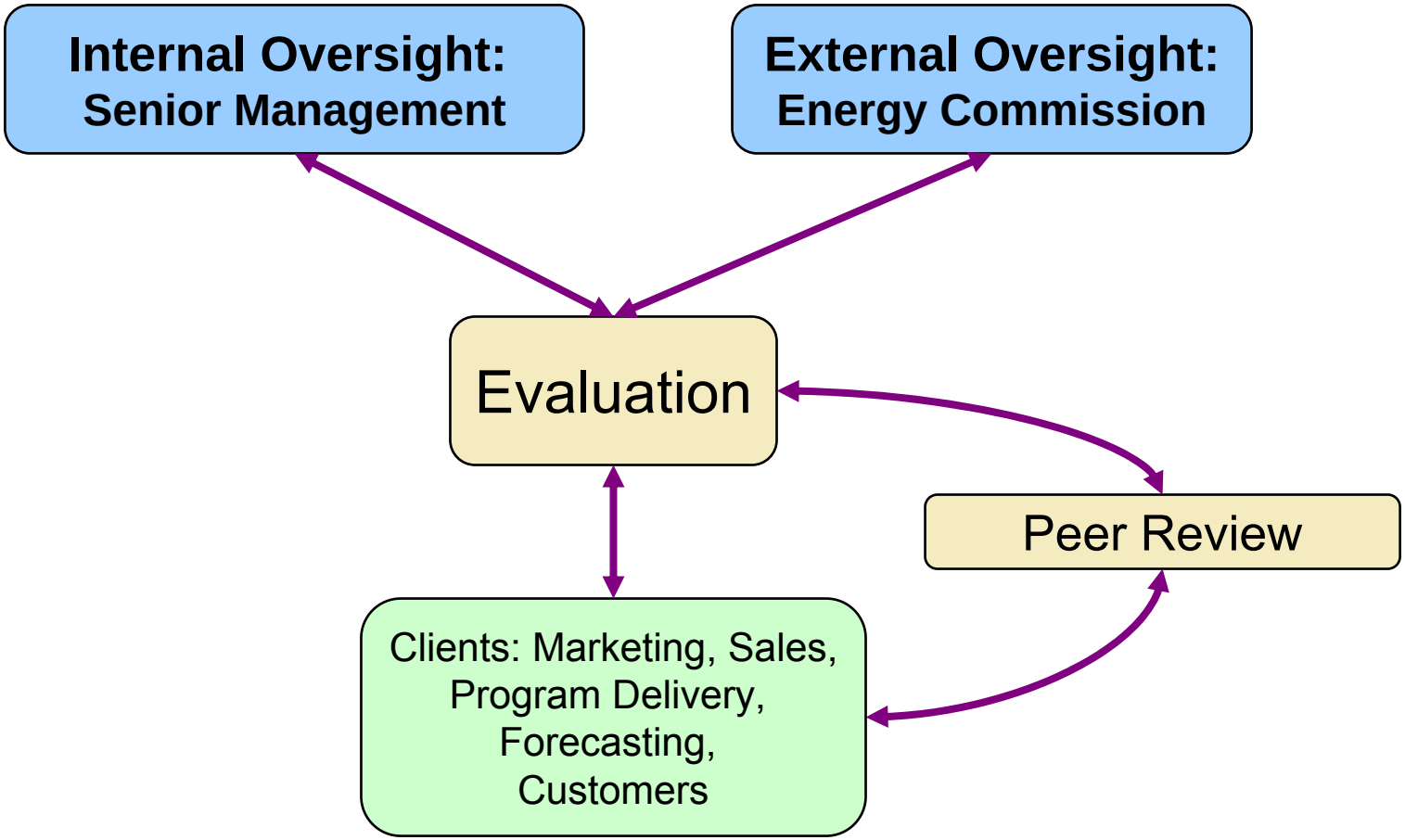
1. Document stated program activities, outputs, impacts and effects
2. Provide feedback to program management
3. Determine if program is being implemented as planned
4. Assess actual program impacts and effects
5. Identify opportunities for program improvement
6. Meet a statutory or regulatory requirement

Evaluation Principles

Energy program evaluation sometimes emphasizes four main principles:

1. Undertaking baseline studies and periodic data collection to understand the nature and size of the pre-program market and changes in the market over time
2. Leveraging existing program, market and customer data to minimize program evaluation costs
3. Using multiple lines of evidence to increase the credibility, validity and reliability of evaluation findings
4. Review and approval of completed evaluation studies by stakeholders

Evaluation, Accountability & Reporting



Quantitative Evaluation

- Quantitative evaluation and reporting refers to data and analysis that is quantitative in nature
- There is some type of metric that is related to variables of interest (number of products purchased, number of program participants, energy savings, demand savings), and data is collected through surveys, on-site measurement, metering
- Advantages of properly planned and executed quantitative evaluation and reporting include:
  - Permit generalizations to be made about large populations on the basis of representative samples
  - Under suitable identifying conditions they allow for causal inferences on impact of variables on outcomes
  - Allow other researchers to confirm or disconfirm research findings

Qualitative Evaluation

- Qualitative evaluation and reporting refers to data and analysis that is qualitative in nature
- Emphasis is on understanding the perspectives of target participants through direct interaction or direct observation, and data is collected through observations, in-depth interviews and focus groups
- Advantages of properly planned and executed qualitative evaluation and reporting:
  - Permits deeper understanding of customer perceptions, needs, values and decision processes
  - Allows for interaction between hypothesis formation and hypothesis testing
  - Generates insights that can be subsequently examined quantitatively

Internal Validity

- According to Trochim, “Internal Validity is the approximate truth about inferences regarding cause-effect or causal relationships. Thus, internal validity is only relevant in studies that try to establish a causal relationship …for studies that assess the effects of social programs or interventions, internal validity is perhaps the primary consideration.”
- In energy program evaluation, we would like to be able to conclude that the program or treatment made a difference – it reduced peak, reduced energy, or both
- Key question: are observed changes attributable to the DSM/MT program or intervention and not attributable to some other possible cause

External Validity

- Again re Trochim, “external validity is the degree to which the conclusions in your study would hold for other persons in other places and at other times.”
- There are two major approaches to external validity: the sampling model and the proximal similarity model
- Sampling Model: (1) identify the population you would like to generalize to; (2) draw a representative sample from that population and conduct research with the sample; (3) generalize results back to the population
- Proximal Similarity Model: (1) think about different generalizability contexts and develop theory about which contexts are more like our study and which are less so; (2) place different contexts in terms of their relative similarities; (3) generalize the results of our study to other persons, places or times that are more alike

Research Designs

- Research design refers to the structure of the research elements – we use a concise notation to refer to these research elements
- Observations or measurements are represented by an O to represent observations of different measures and/or at different times
- Treatments or programs are represented by X with multiple X's for multiple treatments
- Assignment to a group is represented by a letter at the beginning of each line with R for random assignment and N for nonequivalent groups
- Each treatment or comparison has its own line, and subscripts can be used to distinguish different groups

Some Common Research Designs

- **Experimental study**. Participants are randomly assigned to a treatment or a control group, with pre and post measurements

  R: O X O

  R: O O

- **Quasi-experimental study**. Participants choose a treatment or comparison group, with pre and post treatment measurements

  N: O X O

  N: O O

- **Observational study**. After the treatment, participants and non-participants are measured

  N: X O

  N: O

Evaluation of Market Transformation Programs

Types of Evaluations

BC hydro

Four Main Types of Evaluations

1) **Process -** the how and the why ➔ program improvement
2) **Market -** the what and the where ➔ program improvement
3) **Impact -** shorter term direct effects ➔ attribute to program
4) **Market Effects (Transformation) -** longer term effects ➔ attribute to program

Some Key Evaluation Questions

- Energy program evaluations include process, market, impact and market transformation evaluations
- For process evaluations, key issues include: (1) how satisfied are participants with the program, and (2) how effective and efficient are program activities?
- For market evaluations, key issues include: (3) how much of the market has been captured by the program, and (4) what is the remaining market potential?
- For impact evaluations, key issues include (5) what are gross impacts on energy and peak demand, and (6) what are net impacts on energy and peak demand?
- For market effects (transformation) evaluations, key issues include (7) what is the pre-program baseline, and (8) what are the market effects on energy and demand

Process Evaluation

- **Purpose/Objective**
  - Improve program design, implementation, operation, and delivery
- **Timing** - Usually during the program
- **Methods**
  - Examples - interviews, surveys, review tracking system files
- **Key Outcomes** -
  - Documentation of program operations/logic
  - Assessment of process/operation/design effectiveness against expected or planned performance/customer satisfaction
  - Recommendations/suggestions to improve program

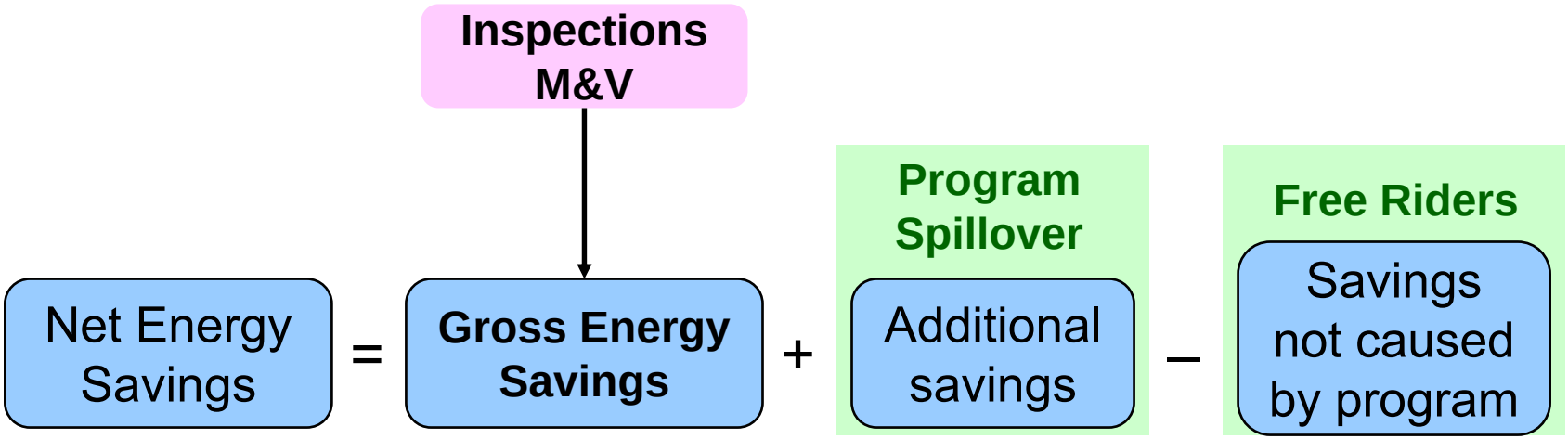
Market Evaluation

- **Purpose/Objective**
  - Assess share of the market captured and the remaining potential
- **Timing** – After at least one year of program activity
- **Methods**
  - Examples - interviews, surveys, shelf stock studies, demand and supply models
- **Key Outcomes -**
  - Documentation of stocking, sales and installation behavior
  - Assessment of market barriers
  - Recommendations/suggestions to improve program

Impact Evaluation

- **Purpose/Objective**
  - To estimate the energy savings caused by a program net of savings that would have occurred in the absence of the program
- **Timing** - Usually at least one year after program launch
- **Methods**
  - Data collection - program files, surveys, billing information, M&V, deemed savings
  - Engineering algorithms, statistical/econometric analysis
- **Key Outcomes -**
  - Net energy and demand savings
  - Estimates of direct program spillover
  - Estimates of free riders
  - Estimates of non-energy impacts (increased productivity, reduced GHG emissions)

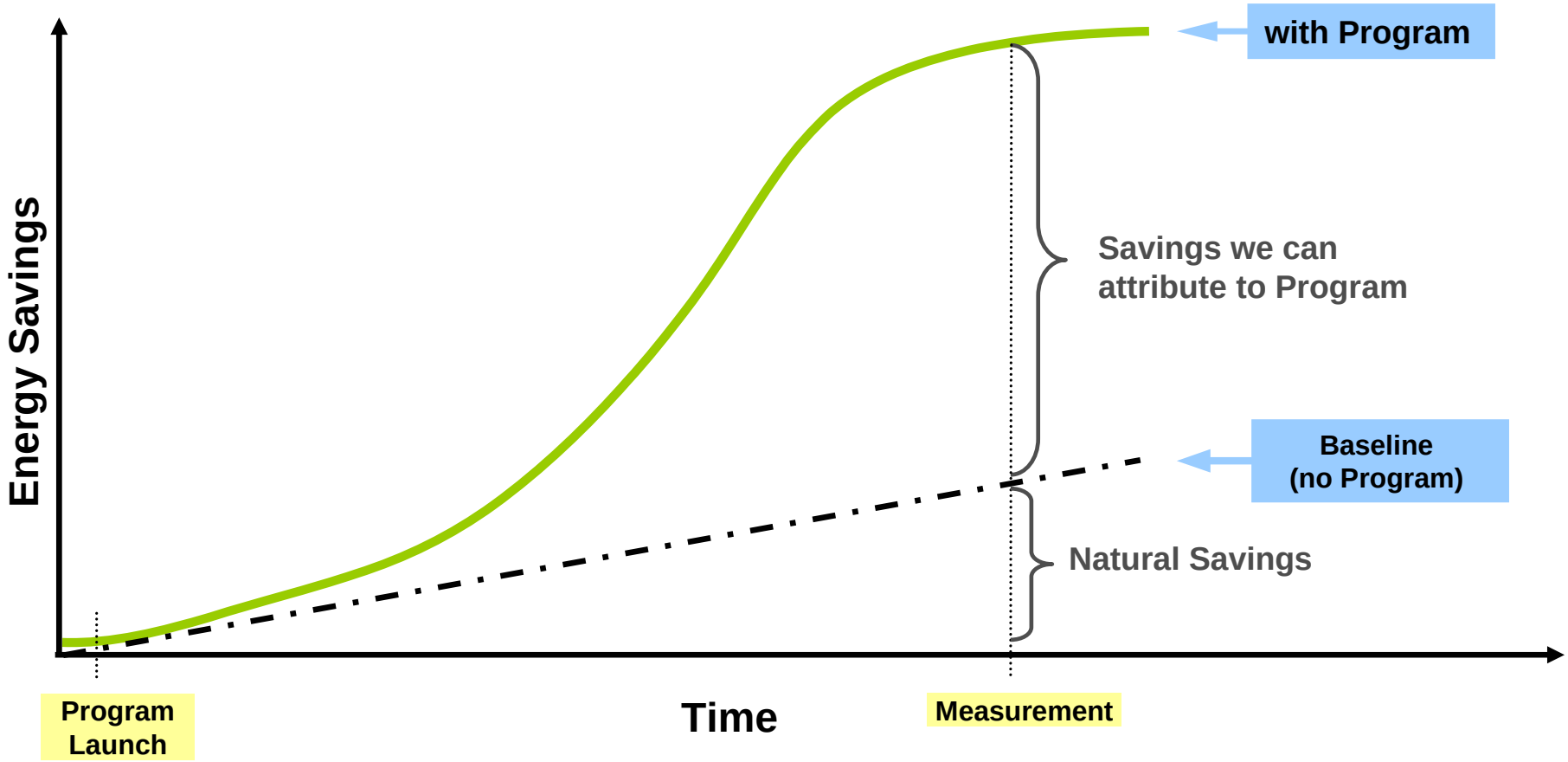
Impact Evaluation - Net Effects



Market Effects (Transformation) Evaluation

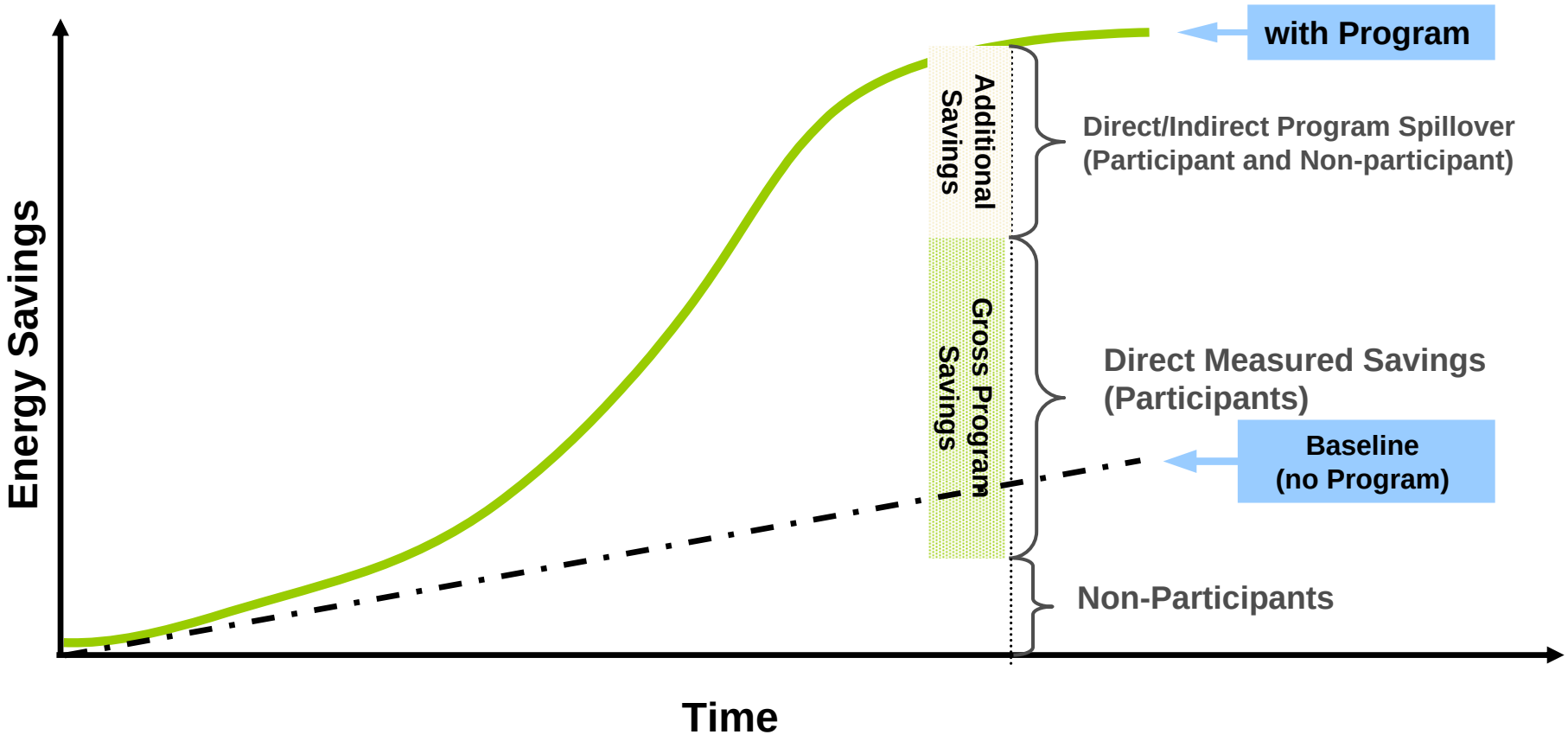
- **Purpose/Objective**
  - To estimate the magnitude of energy savings associated with a change in the structure or function of a market or behaviour of market participants that is caused by an EE program
- **Timing** - Usually at least one year after program launch
- **Methods**
  - Data collection - surveys, interviews, in-store research, baseline measurement
  - Engineering algorithms, statistical/econometric analysis
- **Key Outcomes -**
  - Net incremental energy and demand savings

Bringing it all together(1)

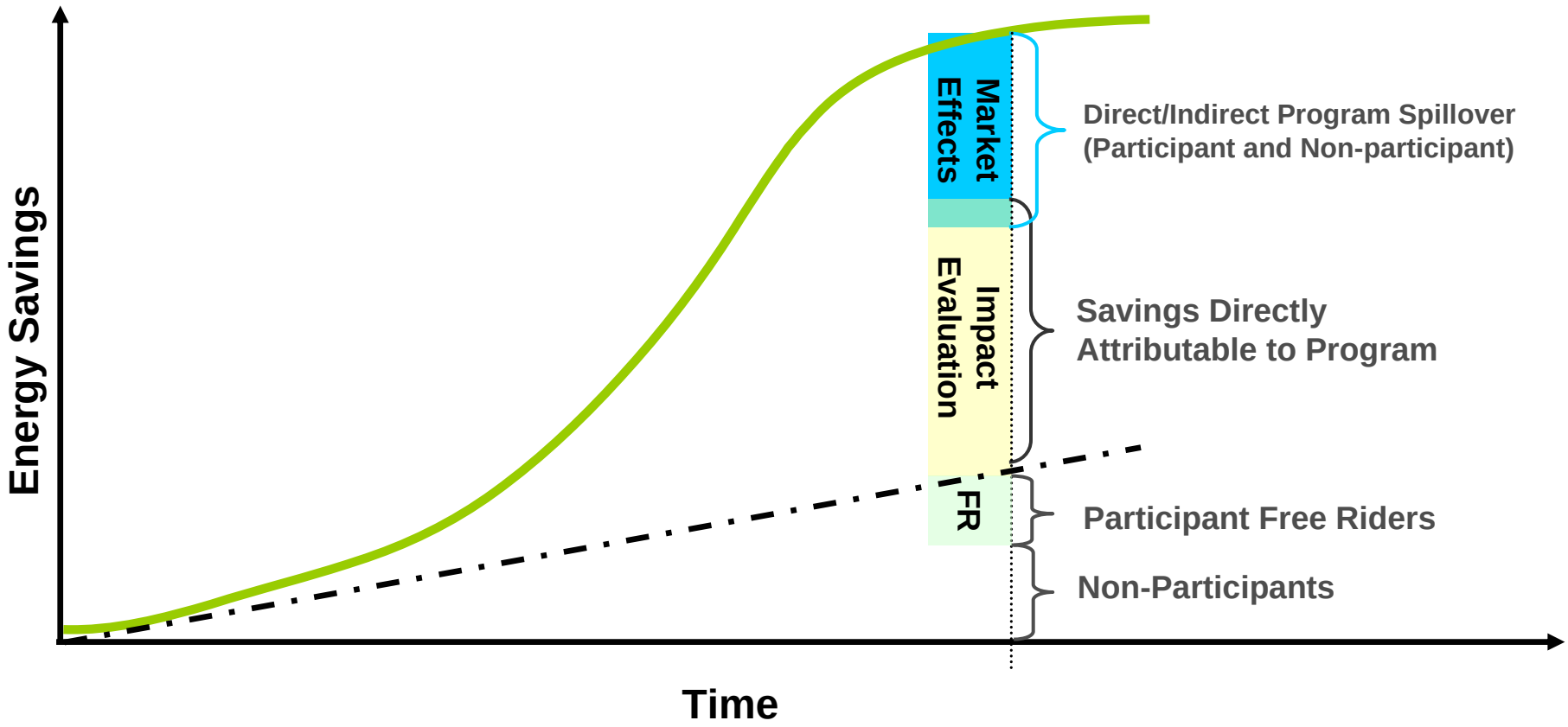


Reliable power, at low cost, for generations. Reliable power, at low cost, for generations. Reliable power, at low cost, for generations. Reliable power, at low cost, for generations.

Bringing it all together (2)



Bringing it all together (3)



Case Study: Residential CFLs in British Columbia

British Columbia has had a variety of CFL programs, including give-away, coupons, manufacturer buy downs, advertising, in-store promotions, advertising

Evaluation Objectives

- characterize the market for CFLs in BC Hydro's service area, from both supply-side and demand-side perspectives
- assess the impact of BC Hydro's residential CFL initiatives on energy savings (direct and incremental market effects)
- assess the impact of BC Hydro's residential CFL initiatives on peak demand (direct and incremental market effects)

Case Study: Residential CFLs in British Columbia

Data Collection -

- Supply side – baseline and annual shelf space surveys
- Demand side - baseline and annual consumer surveys, participant surveys, and comparison group survey in North and South Dakota

Data Analysis -

- Estimation of net direct (coupon + spillover – free rider) and incremental CFL purchases in BC Hydro service area
- Energy and Demand savings calculated using standard engineering equations

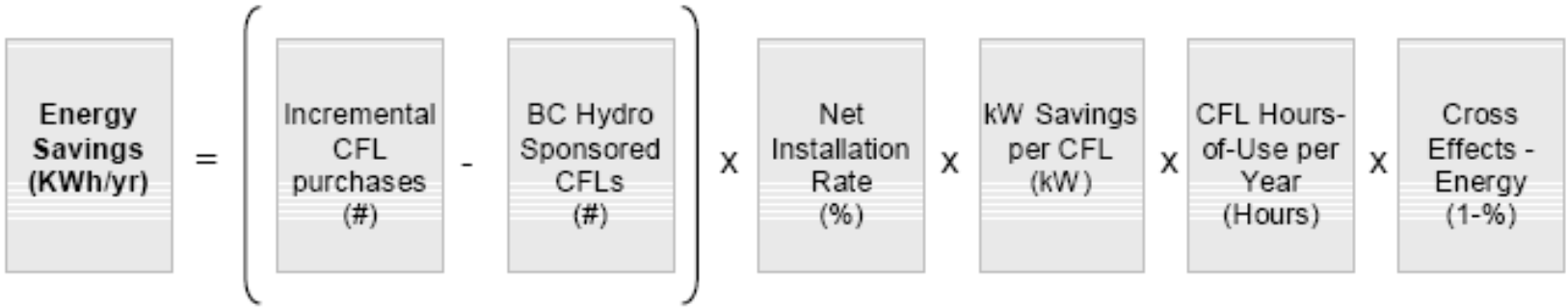
Case Study: Residential CFLs in British Columbia

Direct Effects Equation - Energy Savings:

Energy Savings (KWh/yr)	=	Net CFL Purchases (#)	X	Net Installation Rate (%)	X	Demand Savings per CFL (kW)	X	Annual Hours- of-Use (Hours)	X	Cross Effects - Energy (1-%)

Case Study: Residential CFLs in British Columbia

Market Effects - Energy Savings:



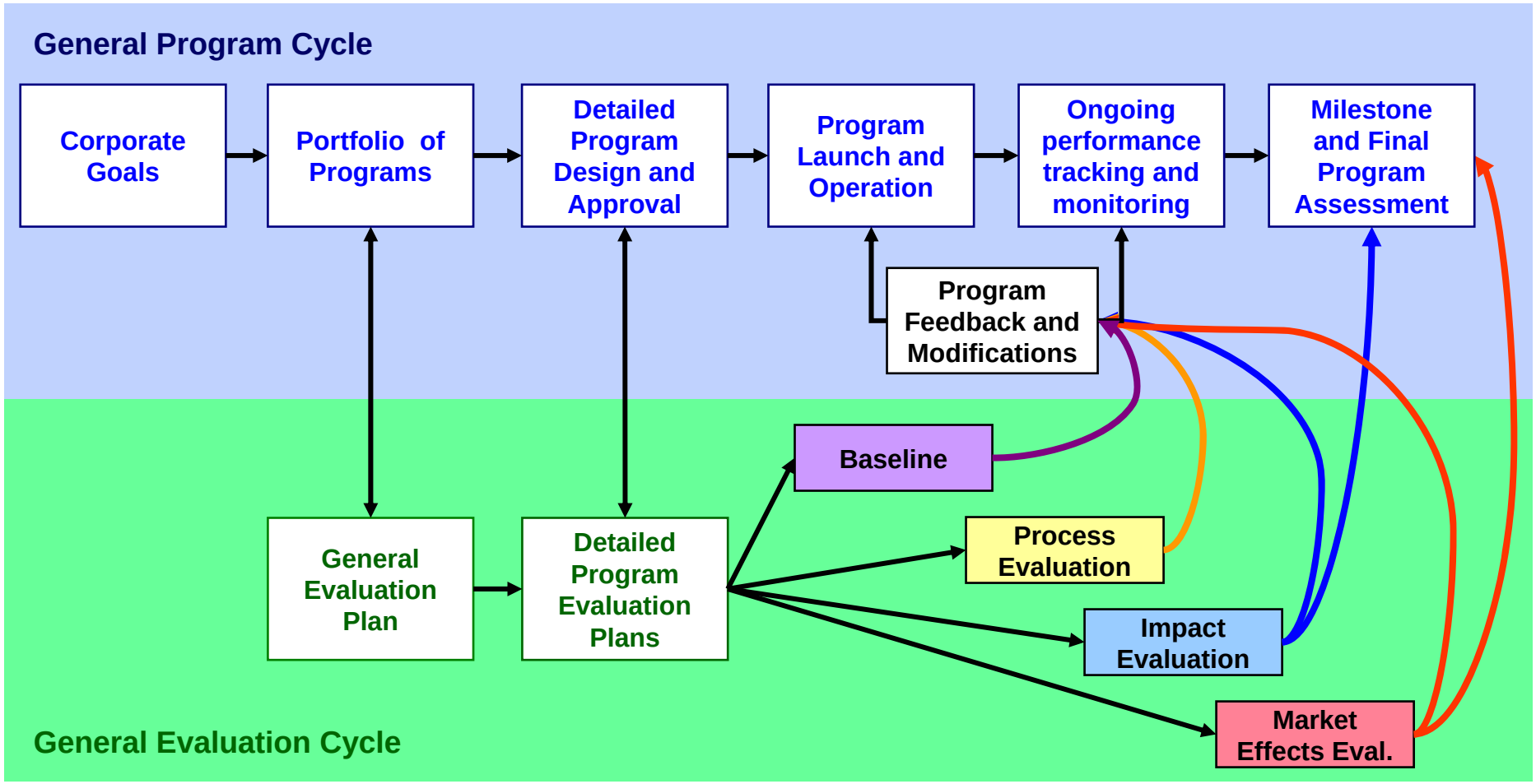
Case Study: Residential CFLs in British Columbia

Summary of Program Impact – Direct and Market Effects
CFL Installations, Energy Savings, and Peak Demand Savings
Fiscal Year 2005-06

	Installed CFLs (000)	Energy Savings Run Rate GWh/Yr [1]	Coincident Peak Demand Savings (MW) [1]
Direct Effects	93.0	4.2	1.9
Market Effects	628.3	19.2	14.8
Total Evaluated Savings	**721.3**	**23.4**	**16.7**

[1] Represents annual savings for CFLs installed during the fiscal year in question, as of fiscal year end.

Reporting: Program and Evaluation Cycles



Evaluation of Market Transformation Programs

Program Theory and Program Logic

BC hydro

Program Theory and Program Logic

- Program theory and program logic are the two key concepts in describing a program
- Although some experts use the terms as synonyms, it is useful to formally distinguish between program theory and program logic
- Program theory tells the program story by indicating by saying what a program will do, how it will do it, and what the results of the program will be
- Program logic is a graphical representation of the rationale for the program story which makes explicit the relationships among activities and goals

Description of Program Theory

- Program theory tells the reader the expected relationship between program activities and program goals in a way which will allow the reader to understand the essential nature of the program and why the program will work
- Development starts with a description of program inputs (resources), activities, outputs, short-term (outcomes) and long-term impacts (outcomes)
- Barriers to program goals and a description of how the program will overcome these barriers are often part of the program theory
- Program theory can also define program metrics, describe how these metrics will be tracked, and state what progress is expected over time

Components of Program Theory and Definitions

- **Market actors** are the economic agents who are targeted by a program of play a role in its delivery – examples include trade allies and customers
- **Inputs** (resources) are the elements required to initiate and sustain program activities – examples include person-years, dollars, skills
- **Activities** are the elements which convert inputs into outputs - examples include purchasing energy efficient appliances or installing additional insulation in homes
- **Outputs** are the immediate deliverables of the program – examples include marketing materials developed, brochures printed, applications processed, incentives paid and inspections undertaken

Components of Program Theory and Definitions

- **Outcomes** are the immediate or the once removed impacts of the program – examples include changes in energy- using product awareness/knowledge/purchase
- **Outcomes** are not within the direct span of control of the program, because they require some additional action by a market actor – for example, a program can distribute CFLs, but the program cannot guarantee that they will be use or that they will be replaced after they burn out
- **Outcomes** can be intended or unintended, with intended outcomes sometimes called impacts and unintended impacts sometimes called effects
- There can be several layers of outcomes – short-term (sometimes called program purpose) and long-term (sometimes called program goal)

Description of Program Logic

- **Program logic** show the reader explicitly the flow between inputs or activities, outputs, and their subsequent, short-term impacts (purpose), and long-term impacts (goal)
- Sometimes the program logic is displayed with key elements in boxes and the flows between elements represented by arrows, and these logic models can often become quite complex

$$\text{Input}_1 \rightarrow \text{Output}_1 \rightarrow \text{Purpose} \rightarrow \text{Goal}$$
$$\text{Input}_2 \rightarrow \text{Output}_2 \rightarrow \text{Purpose} \rightarrow \text{Goal}$$

- Alternatively the program elements can be placed in a table with the linear relationships represented by rows in the table and different activities represented by columns

Simple Program Logic Model

	Activity 1	Activity 2	Activity 3	Assumption
Inputs/ resources				
Outputs				
Purpose				
Goal				

Reliable power, at low cost, for generations. Reliable power, at low cost, for generations. Reliable power, at low cost, for generations. Reliable power, at low cost, for generations.

Use of Program Theory/Program Logic

- Clarify and document program assumptions and goals and develop a shared understanding of the program among stakeholders
- Identify missing activities (which should be added) or extraneous activities (which should be eliminated)
- Provide a framework for program monitoring, tracking and control (program management)
- Provide a framework for examination of program activities and recommendations for change (formative evaluation)
- Provide a framework for identification of key hypotheses and data needed to test these hypotheses (summative evaluation)

Program Success, Theory Failure, Implementation Failure

- **Program Success**:

  Program → set in motion causal process → which led to desired effect

- **Theory Failure**:

  Program → set in motion causal process → which did not lead to desired effect

- **Implementation Failure**:

  Program → did not set in motion causal process → which would have led to desired effect

Multifamily HVAC Replacement: Description of Market Opportunity (modified from TecMarket Works 2004)

- A utility wants to reduce HVAC energy consumption in multifamily dwellings
- Target is owners of multifamily dwellings who are making HVAC equipment replacement decisions
- Basic idea is to get owners to install energy efficient heating, cooling and ventilation equipment to capture short-term energy and demand savings
- Program is operated through equipment supply and installation contractors who in turn work with their customers who are or could be made interested in installing energy efficient equipment at the time of equipment replacement

Multifamily HVAC Replacement: Pre-program Research Findings

- **Facility owners want**:
  - Reduce operating costs of their facilities
  - Reduce first costs for new equipment
  - Reduce hassle (transactions costs) involved in equipment replacement
  - Replace HVAC equipment before failure
- **Equipment contractors want**:
  - Additional business opportunities in the competitive replacement market
  - Be seen as innovative and professional suppliers of advice and equipment

Multifamily HVAC Replacement: Program Theory

- **Program Rationale**: Program can increase installation of energy efficient and cost-effective HVAC equipment by affecting the demand side and supply side of the HVAC replacement market, and replacement of equipment will lead to energy and peak savings
- **Market Actors**: Key market actors are owners of multifamily residential buildings and HVAC equipment suppliers
- **Demand Side Activities**: increase multifamily residential building owner knowledge and interest; reduce cost of audits, reduce first costs of equipment
- **Supply Side Activities**: increase installer interest and knowledge; increase audit capacity; increase installation capacity

Multi-family HVAC Replacement: Logic Model

Activities	Design marketing materials	Recruit and train contractors	Recruit and audit customers	Equipment purchased and installed
Inputs	Design and develop materials	Training program	Promotion and audits signed-up	Incentives provided
Outputs	Appropriate materials distributed	Trained contractors in place	Audits completed	Equipment installed
Purpose	Trade allies interested in participating	Trade allies market program	Customer agrees to installation	Equipment operating properly
Goal				Energy and peak saving

Reliable power, at low cost, for generations. Reliable power, at low cost, for generations. Reliable power, at low cost, for generations. Reliable power, at low cost, for generations.

Selected References

- Campbell, D. T., & Stanley, J. C. ,1966 and many updates, Experimental and Quasi-experimental Designs for Research, Rand McNally, Chicago, Ill.
- Hall, N., et al, 2004, The California Evaluation Framework, TecMarket Works.
- McLaughlin, John A. and Gretchen Jordan, 2004, Logic Models: A Tool for Describing Program Theory and Performance, in Handbook of Practical Program Evaluation, Jossey-Base, San Francisco, CA
- Rogers, Patricia, Timothy A. Hacsi, Anthony Petrosina and Tracey A. Huebner, 2000, “Program theory in evaluation: causal models in program theory evaluation”, New Directions for Evaluation, 87:47-55
- Trochim, William H., 2001, The Research Methods Data Base, 2nd ed., Atomic Dog Publishing, Cincinnati, OH
- Weiss, Carol. H, 1998, Evaluation Methods for Studying Programs and Policies, Prentice-Hall, Upper Saddle River, NY