Evaluating Public R&D Investments

Presented by: Rosalie Ruegg, TIA Consulting, Inc.

ruegg@ec.rr.com

National Science Foundation (NSF)
Industry/University Cooperative Research Consortium (IUCRC)
Annual Meeting of Evaluators

Key Bridge Marriott

Arlington, VA

June 9, 2011

Outline

1. **R&D Evaluation Overview**

   - Why evaluate

   - Types of evaluation, evaluation questions , & study designs

   - A variety of evaluation methods and techniques

2. **Focus on Benefit-Cost Analysis**

   - Extended to portfolio analysis

   - Extended to multiple types of benefits

   - Issues

3. **Interactive Exercise: Evaluation Framework**

   - Mapping evaluation questions to a logic model

   - Identifying suitable methods to address the various evaluation questions

1. R&D Evaluation Overview

- Why evaluate
- Types of evaluation, evaluation questions, & study designs
- Evaluation methods and techniques

Why Evaluate?

- **To provide feedback for program management**

  - Is the program having the desired/intended effects?
  - Are changes needed?

- **To meet requirements for accountability**

  - Agency requirements
  - Congressional: GPRA
  - Executive (OMB): PART
  - Other?

- **To develop policy insights**

  - How does our program work?
  - Do we understand what will work better?
  - Can our evaluation inform future investment decisions ("Science of Science Policy")

Types of Evaluation & Evaluation Questions

Types of Evaluation:

- Process Evaluation -- performed on program structure and operations

- Impact Evaluation -- performed on program results

Types of Evaluation Questions:

- Descriptive -- who, what, where, how many, how much

- Normative -- how do actuals compare against program standards or goals

- Impact (cause-effect) – what outcomes/impacts has the program caused (more difficult to answer)

Evaluation Study Designs

Study design depends on type of question asked, degree of rigor required, resources available, and other issues of feasibility. Alternatives for RD&T study design include:

(For more info, see Web Center for Social Research Methods, www.socialresearchmethods.net)

- **Experimental design –** randomized trials; most rigorous, but not often used for RD&T studies due to feasibility issues.

- **Quasi-experimental design** -- elements of experimental design, but no randomized trials.

- **Counterfactual analysis** -- strengthened by documentation of plausible cause & effect relationships, logical time order, rigorous comparisons, and treatment of rival explanations.

- **Observational/descriptive/non-experimental design** – least rigorous, but adequate for answering descriptive questions.

A Variety of Evaluation Methods

Over time evaluators have taken different approaches to evaluation studies, and some of these have become recognized as "methods of evaluation." Why a variety?

- Different methods better answer different questions
- To strengthen confidence by providing multiple lines of evidence
- To reach different audiences
- For flexibility given different funding and timing constraints and preferences

Examples of Methods of R&D Evaluation

Expert review
Survey
Indictor metrics
Interim Performance Rating Schemes
Case Study
Benchmarking
Social Network Analysis
System Dynamics Modeling
Bibliometrics
Historical tracing
Econometrics and Statistical Analysis (used variously)
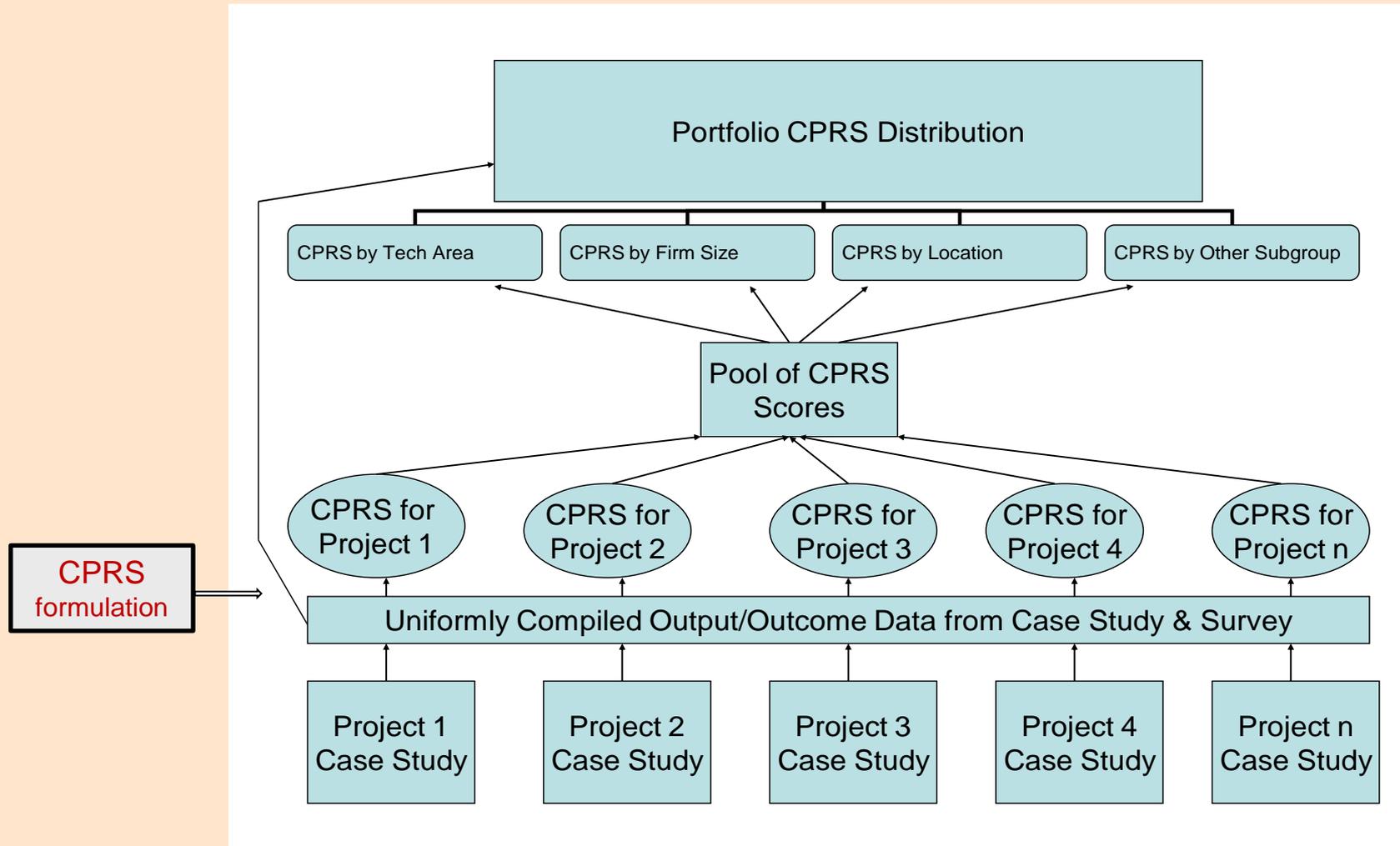Benefit-Cost Analysis
Others—it is an emerging field

Supporting techniques: logic models, sensitivity analysis, probability analysis, simulations, visualization tools, data collection tools, etc.

# Combinations of Methods to Increase Evaluation Effectiveness, e.g.,

- ## Interim performance rating systems

  - ATP's Composite Performance Rating System (CPRS) combined indicator metrics from survey, information from case study, and use of statistical weighting techniques to bridge from project case study to portfolio analysis of all completed projects.

  - Dept of Agriculture's Portfolio Review Expert Panel (PREP) Process combined the use of numerical ratings by Expert Reviewers, informed by evidentiary materials and with ratings based on OMB R&D criteria (relevance, quality, and performance) to score portfolio progress & provide recommendations for the Agency.

- ## Historical tracing

  - DOE/EERE's recent Historical Tracing studies combine bibliometrics -- patent & publication citation analysis; interview; review of documents, databases, and licensing; and statistical techniques, to comprehensively depict linkage of EERE's R&D outputs forward to all areas of impact, and to noteworthy commercial innovations in the targeted area.

8

ATP's CPRS: Multi-Tier Approach to Bridge from Project Case Study to Portfolio Analysis

CPRS = Composite Performance Rating System
Source: Ruegg (2006) NIST GCR 06-891

DOE/EERE's Historical Tracing Forward and Backward

Forward tracing from R&D to downstream outcomes

R&D Program

Innovation 1

Innovation 2

Dead-end

Innovation 3

Innovation 4
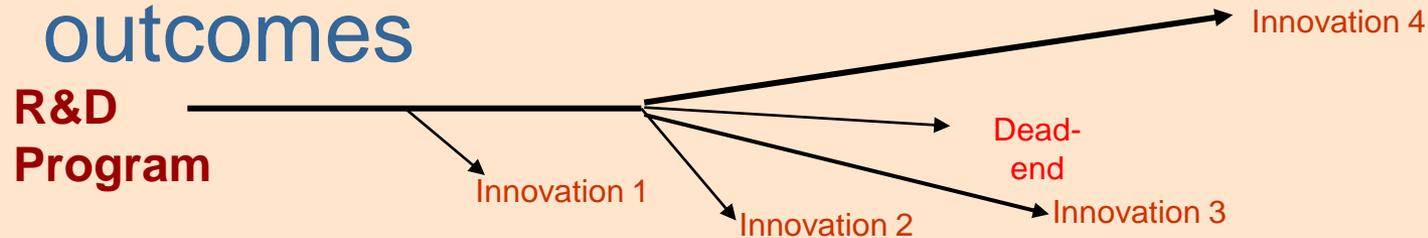
Backward tracing from a selected outcome to upstream R&D
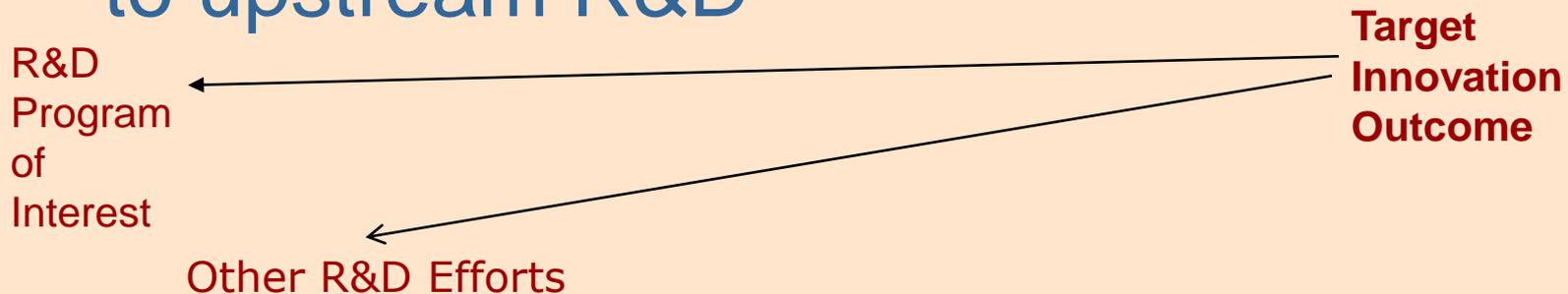
R&D Program of Interest

Target Innovation Outcome

Other R&D Efforts

2. Focus on Benefit-Cost Analysis

- Extended to portfolio analysis

- Extended to multiple types of benefits

- Issues

Benefit-Cost Analysis Extended

B-C method was traditionally applied at the project level, but has been extended by ATP and DOE:

➢ ATP – extended B-C analysis to evaluate portfolios of technology clusters.

➢ DOE- extended B-C analysis to address multiple categories of benefits for portfolios consisting of entire technology programs and subprograms.
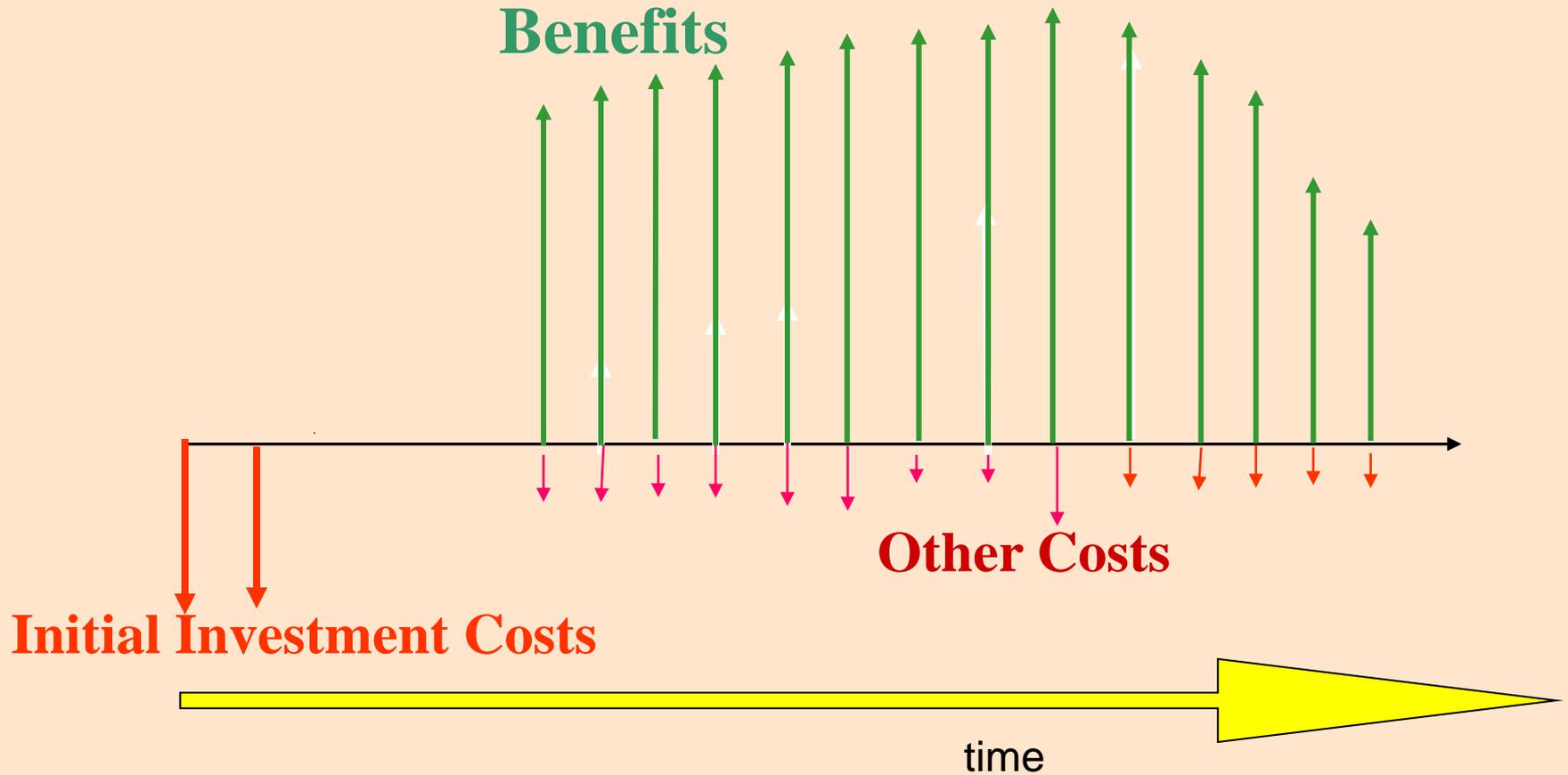
Benefit-Cost Analysis Description

Features of Traditional B-C Analysis:

- Identify benefits and costs of a project (or program or portfolio)

- Express effects in money terms where possible

- Note timing of cash flows & apply appropriate discount rate

- Computation of economic performance measures, e.g.,

  - Net Present Value Benefits (NPV), - Benefit-to-Cost Ratio (B/C), & -Internal Rate of Return on Investment (IRR)

- Qualitative treatment of other effects or measurement in other units

Uses:
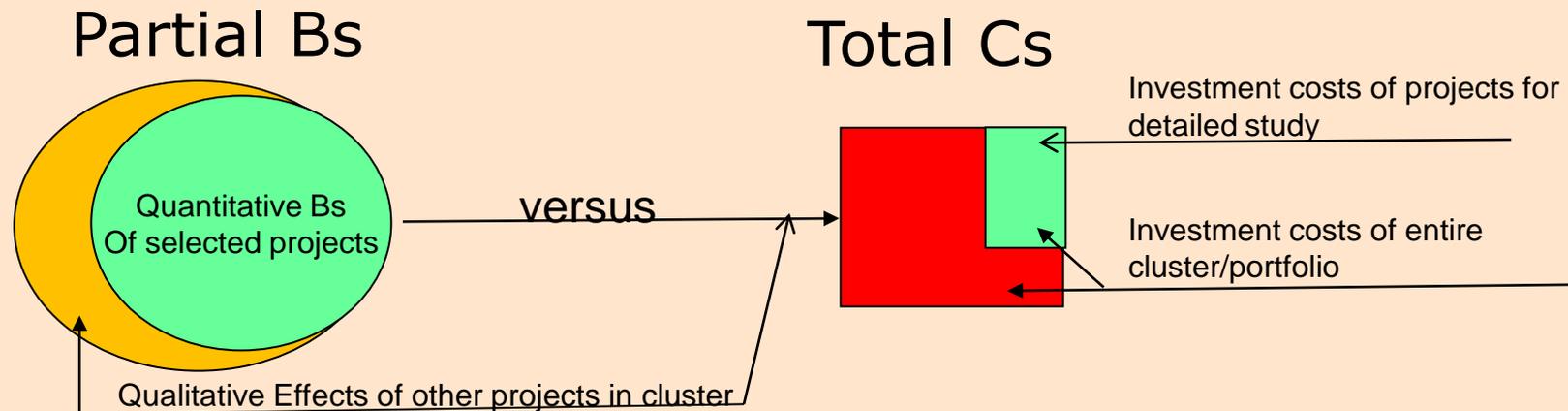
  - To demonstrate that a project (or program) was or was not economically worthwhile

  - To inform R&D investment decisions

Benefit-Cost Analysis: Working with Cash Flows

Benefits

Other Costs

Initial Investment Costs

time

Benefit-Cost Analysis Extended from a Single Project to a Portfolio

Extension of the analysis from application to a single **applied** research project to a **cluster** of technologies or portfolio of projects has the advantage of providing a more useful, scaled-up measure without a similar scale-up in evaluation costs.

Partial Bs

Total Cs

Quantitative Bs
Of selected projects

versus

Qualitative Effects of other projects in cluster

Investment costs of projects for detailed study

Investment costs of entire cluster/portfolio

Benefit-Cost Analysis Extended to Multiple Categories of Benefits (DOE/EERE Example)

❑ Retrospective for greater credibility

❑ Followed best-practice methodology

❑ Comprehensive – DOE's treatment of 4 types of benefits:

- economic

- environmental

- security

- knowledge

❑ Consistent & uniform across studies (as appropriate)

- use of unifying framework

- consistent set of B-C conventions

- consistent set of Economic Performance Measures

- consistent report features

❑ Guide for contractors on DOE/EERE B-C Methodology followed

❑ [Ruegg & Jordan, 2010, draft; revision in 2011]

16

Areas of Special Focus in Estimating Benefits

Specifying the Next Best Alternative (the "Defender Technology"), i.e., what would have been used in lieu of the subject technology?

Accounting for "Additionality", i.e. what was different as a result of the Program/Subprogram? Are there "rival explanations" that must be eliminated?

A Matrix for Assessing Attribution by Technology Stage

Categories of Information Needed for Additionality Assessment	Technology Timeline (Stage of Research, Development, and Commercialization)→					
	Preliminary & detailed investigation	Develop components	Develop system	Validate/ demonstrate	Commer- cialize	Market Adoption
History of the technology						
What DOE Did						
What Others Did (Rival Explanations—Private Sector and Other Nations)						
What Others Did (Rival Explanations –US & State Government)						
The DOE Effect						
Description of DOE Influence And its strength						
Basis of evidence of influence						

Example: DOE/EERE B-C Study of Vehicle Combustion Engine R&D

Cluster of Technologies in Vehicle Combustion Sub-program:

- **laser diagnostic and optical engine technologies**

- **combustion modeling**

- emission control technologies

- solid state energy conversion

(Selected for Detailed Analysis: "red" -- both focused on heavy duty diesel engines)

Example: DOE/EERE B-C Study of Vehicle Combustion Engine R&D

➢ Economic Benefits:

 w/o EERE, fuel efficiency 4.5% lower, and fuel savings of17.6 billion gallons of diesel fuel from 1995 through 2007

Monetary value ($2008, undiscounted): $34.5 billion

➢ Environmental Benefits, reduction in air emissions:

177.3 million metric tons of CO_2; 0.063 tons NO_x; 3.080 tons PM; 0.096 tons SO_x

Monetary value of heath impacts avoided in $2008, undiscounted: $35.7 billion

➢ Security Benefits: equivalent of 417.9 million barrels of imported crude oil = 1% reduction of total crude oil imported by US 1995 thru 2007

➢ Knowledge Benefits: Foundation for more than 12 important technologies in combustion, plus advances in ion mobility spectrometry

Example: DOE/EERE B-C Study of Vehicle Combustion Engine R&D

Performance Measures for EERE's investment in Vehicle Combustion Subprogram:

NPV benefits: $23.1 billion

BCR: 53:1

IRR: 63%

(discount rate @ 7%; base year = 1986)

Issues RE Extension of B-C Analysis

- Extending monetized benefits to new categories of benefits w/o reducing data quality.

- Reducing focus from program mission.

- Tendency towards inappropriate comparisons across programs.

Methods of R&D Evaluation

Expert review
Survey
Indictor metrics
Interim Performance Rating Schemes
Case Study
Benchmarking
Social Network Analysis
Systems Dynamics Modeling
Bibliometrics
Historical tracing
Econometrics and Statistical Analysis (used variously)
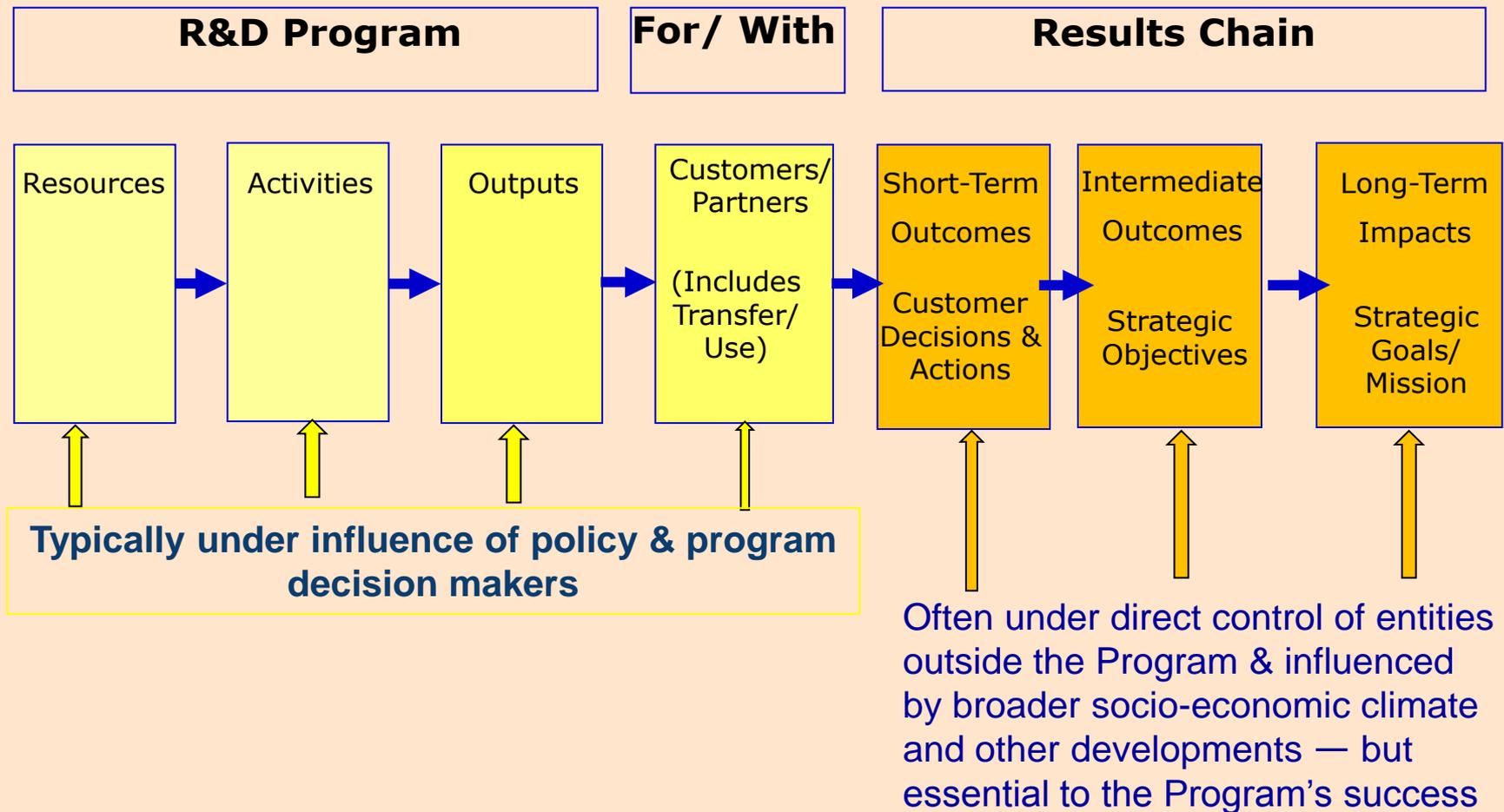Benefit-Cost Analysis
Others—it is an emerging field

methods illustrated: in red

3. Evaluation Framework

- Mapping evaluation needs to a program "Logic Model"

- Questions arising at different phases of a program performance cycle

- Choosing the right evaluation tool(s) to meet each need

- Summary steps in conducting evaluation

- Questions & discussion

Logic Models communicate WHAT is to be done, How, with Whom, and Why

R&D Program			For/ With	Results Chain		
Resources	Activities	Outputs	Customers/ Partners (Includes Transfer/ Use)	Short-Term Outcomes Customer Decisions & Actions	Intermediate Outcomes Strategic Objectives	Long-Term Impacts Strategic Goals/ Mission

Typically under influence of policy & program decision makers

Often under direct control of entities outside the Program & influenced by broader socio-economic climate and other developments — but essential to the Program's success
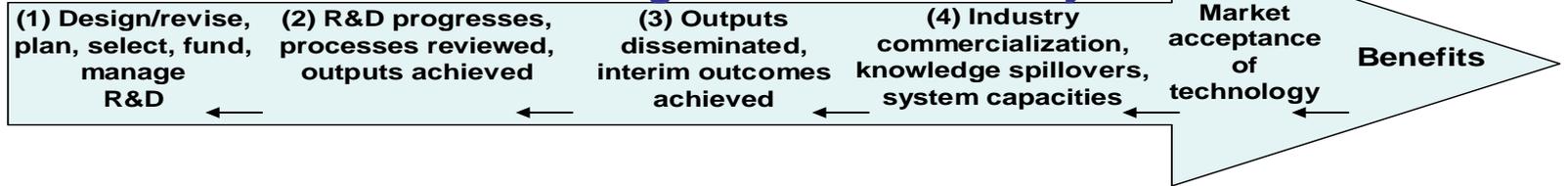
Logic Model Drives Evaluation Questions

INPUTS ACTIVITIES OUTPUTS OUTCOMES IMPACTS

Key Metrics

| Quality, Relevance Management | Technical Progress, R&D Infrastructure | Technology Output Goal, Hand off | Interim/ Diffusion Outcomes | Ultimate Outcomes |

Phases of Program Performance Cycle

(1) Design/revise, plan, select, fund, manage R&D ← (2) R&D progresses, processes reviewed, outputs achieved ← (3) Outputs disseminated, interim outcomes achieved ← (4) Industry commercialization, knowledge spillovers, system capacities ← Market acceptance of technology ← Benefits
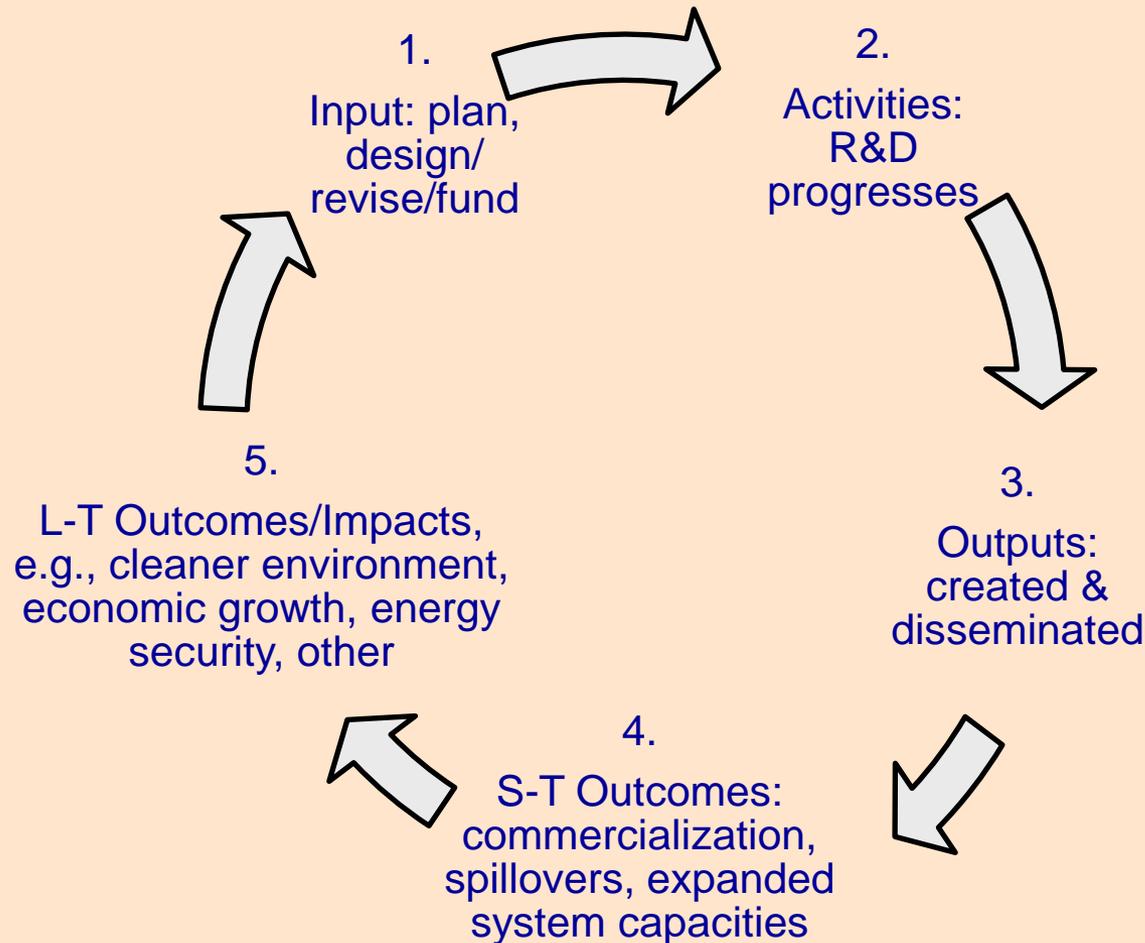
Relevant Questions at Each Phase

Relevancy?
Timeliness?
Partners?
Technologies?
Why these?
Alignment?
Risk
Why?
Cost? Adequacy?
Past cost?
Past benefits?
Expected benefits?
Processes?

Progress?
Quality?
Participants?
Processes?
Knowledge outputs?
Other outputs?
Vs. targets?
Program productivity?

Users?
Importance of?
Relationships?
Commercialized?
Influencing factors?
Details of progress?
Spillover indicators?

Further commercial progress?
Realized benefits and costs?
Attributed program effects?
Links from noteworthy innovations to R&D?
Spillover effects?
Was it worth doing?

How will you know—what measures?

Source: Ruegg & Jordan, 2007

26

Program Performance Cycle

1.
Input: plan,
design/
revise/fund

2.
Activities:
R&D
progresses

3.
Outputs:
created &
disseminated

4.
S-T Outcomes:
commercialization,
spillovers, expanded
system capacities

5.
L-T Outcomes/Impacts,
e.g., cleaner environment,
economic growth, energy
security, other

Exercise: Mapping Evaluation to a Logic Model

| Inputs/Resources | → | Activities | → | Outputs | → | Customers/Partners/users | → | Short-Term Outcomes | → | Intermediate Outcomes | → | Long-Term Impacts |
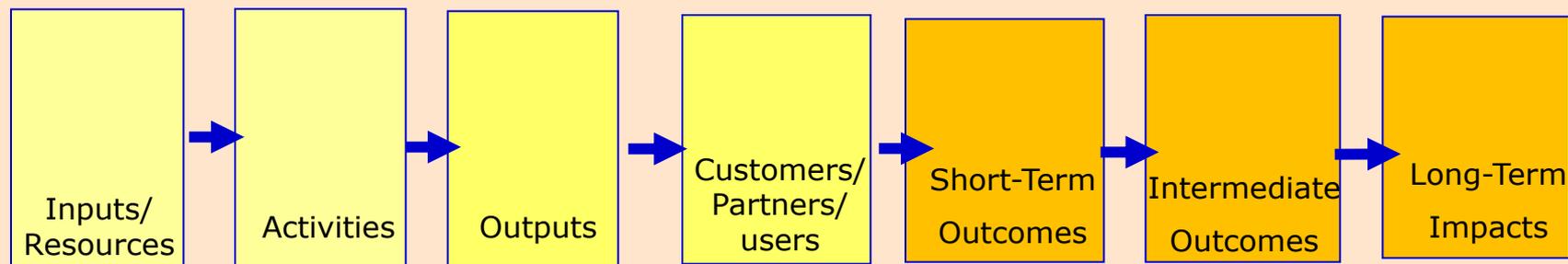
What are I/UCRC expected inputs, activities, outputs, partners/users, outcomes, and impacts?

What evaluation questions will need to be answered – why, for whom, when?

What evaluation tools will be needed; what measures; who will conduct?

Phase 1: Inputs: plan/design/revise/fund

Sample Questions	Methods/Measures

Phase 1: Inputs: plan/design/revise/fund

Sample Questions	Methods/Measures
Are planned program mechanisms, processes, and activities, and purpose clear and well designed to achieve agency and program goals?	Expert judgment/Reviews and critiques
What evidence is there that required partners/stakeholders will participate?	Survey/Statistical results
How can the planned transformational processes be strengthened?	Analysis of past efforts, e.g., case study; econometric analysis; other
Is there alignment of supporting investments/initiatives?	Expert judgment/Review and critiques
What are expected long-term benefits ?	Benefit-cost analysis (prospective)
Have past efforts been worth the cost?	Benefit-cost analysis (retrospective)

Phase 2: Activities

Sample Questions	Methods/Measures

Phase 2: Activities

Sample Questions	Methods/Measures
What technologies are under development?	Tracking metrics for technologies funded and progress milestones/Comparison against targets
What outreach efforts have been made?	Tracking metrics for outreach efforts/Comparison against targets
Who is participating?	Tracking of participants/Comparison against targets
What is the regional distribution of activities nationwide?	Tracking geographical distribution of activities/Conformance with goals

Phase 3: Outputs: Creation & Early Transfer

Sample Questions	Methods/Measures

Phase 3: Outputs: Creation & Early Transfer

Sample Questions	Methods/Measures
What technologies were successfully developed as laboratory prototypes?	Tracking outputs of laboratory prototypes
What technologies have moved into commercial use?	Tracking first commercial use
How many publications resulted?	Bibliometric counts of publications
How many patents were filed and how many were issued?	Bibliometric counts
How many demonstrations of newly created knowledge have been made and to what client bases?	Tracking information dissemination through, e.g., presentations, meetings, and on-line sites
Is knowledge transfer underway through publication and patent citations?	Bibliometric citation analysis
What barriers are slowing tech transfer and early adoption?	Survey; interview, case study

34

Phase 4: Short-Term Outcomes

Sample Questions	Methods/Measures

Phase 4: Short-Term Outcomes

Sample Questions	Methods/Measures
How many uses are there in the field?	Tracking adoption
What are the early advantages/disadvantages from implementation of the technology?	Case study, survey, interview, statistics
Who is and is not using the technology and why?	Survey or interview of a sample of users and non-users
Are there growing networks of users?	Social network analysis (a snapshot now and later for comparison)
What is the overall performance of the portfolio of projects are this stage?	Interim performance rating schemes to get distribution of projects by performance across the entire portfolio
Are program changes needed?	Process evaluation using the above measures

Phase 5: Long-Term Outcomes/Impacts

Sample Questions	Methods/Measures

Phase 5: Long-Term Outcomes/Impacts

Sample Questions	Methods/Measures
Percentage of potential users who have adopted the technology?	Survey/Statistics
Growth in network of users geographically?	Social network analysis/second look
Comparative influence on knowledge and downstream innovations?	Historical tracing using analysis of citations and licensing; interview; documents & database searches; and comparisons of influence across organizations
Development of an industry/supply chain based on the new technology	Market analysis
Impact on economic growth and the environment?	Impact evaluation, e.g., benefit-cost analysis

Questions & Discussion

Contact info:

Rosalie Ruegg, Managing Director

TIA Consulting, Inc.

ruegg@ec.rr.com

Office Phone: 252-354-9321

Mobile: 252-646-7713