

Best Practices for Behavior-Related Programs

BEHAVIOR RELATED PROGRAMS:

Best Practices for Program Design and Evaluation

ACEEE-MT Conference, 2009

Project oversight: Ed Vine, Linda Schuck CIEE

Lisa A. Skumatz, Ph.D.

Skumatz Economic Research Associates, Inc. (SERA)

762 Eldorado Drive, Superior, CO 80027

303/494-1178 email: skumatz@serainc.com

Team: SERA, Cadmus, bKi, NERA, Yale

TOPICS

- Project scope & definitions
- Review of findings
 - Impact
 - Attribution
 - EULs
 - NEBs
- Summary

PROJECT SCOPE

- CIEE – Scope
 - Measuring / evaluating energy savings from behaviors
 - Measuring non-energy impacts from behaviors
 - Attribution of savings, free riders, takeback, etc.
 - Persistence of savings from behaviors
- Overall → Evaluation of impacts and attribution from programs
- Considerations / complexities
 - Transition to non-measure-based programs
 - “Chatter” in marketplace – multiple influencers

PROJECT ISSUES

- Goals, applications
 - What are we trying to measure; what is necessary to measure?
 - Measure, program, or portfolio level
- Degree of effort appropriate?
 - Methods practical? Defensible? Learn from other fields?
- Want approaches that:
 - Assure public dollars responsibly spent
 - Apportion dollars and efforts between strategies
 - Help identify appropriate exit time
- State of art – best practices, applications, methods
- Areas for additional research

PROJECT DEFINITIONS & APPROACH

- Definitions / scope
 - Behavioral: Participation decision, equipment acquisition, use / operation / maintenance of equipment, energy use behaviors.
 - Broad consumer; broad energy
 - Broad programs including measures, interventions, education, advertising, other.
- Interviews, literature, analysis

IMPACT EVALUATION

IMPACT EVALUATION

- Interviews / literature review shows varied approaches
 - Gross – widget counting
 - Verified gross – inspection vs. agreed / reasonable
 - Market effects – sales
 - Education – survey based / self-reporting
 - Problems linking to direct savings; some simply don't count / try

IMPACT EVALUATION

- Findings / Issues / Questions:
 - Key part is knowing what program is trying to accomplish – then clearer to measure impacts
 - Some say big sophisticated models not working, especially for behavioral programs, technologies, noise – need to build up from rudiments
 - Capture close to point of making the decision
 - Need more market assessment
 - Some deemed – others say deemed savings ridiculous
 - Some say more important to get widgets out there than to bean count – and should things on the shelf get counted (CFL)
 - Harping on best practices, but if rubber stamp same programs every year, and not reward creativity... no innovation rewards – and need non-participant surveys
 - Best practices should include clear study of what is out there in market – critical to program planning- what to incent, when to get in, when to leave.

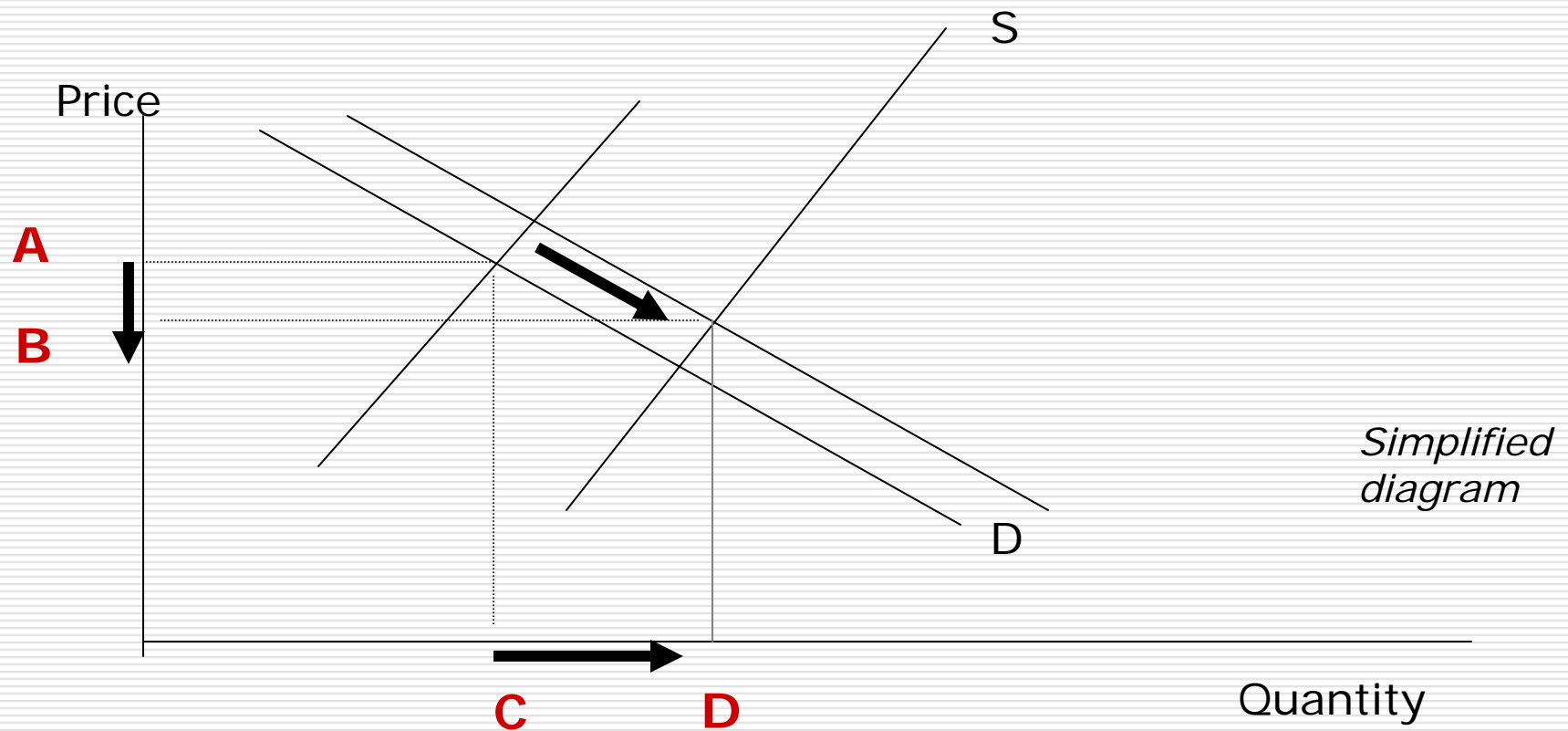
IMPACT EVALUATION

- Findings / Issues / Questions:
 - Require programs to be cost-effective – but maybe not 1 year horizon! Helps education and behavioral
 - Capturing behavioral energy savings tough if all behaviors aren't measurable, and measureable relative to each other
 - Baseline / program record problem...
 - Regular surveys on market situation and saturations
 - Need research like lighting logger studies
 - Evaluations over time can provide information essential to improved program design – and when to get out of market
 - Sales data declining (e.g. fell apart when big box stores stopped reporting)
 - Innovation – price approach

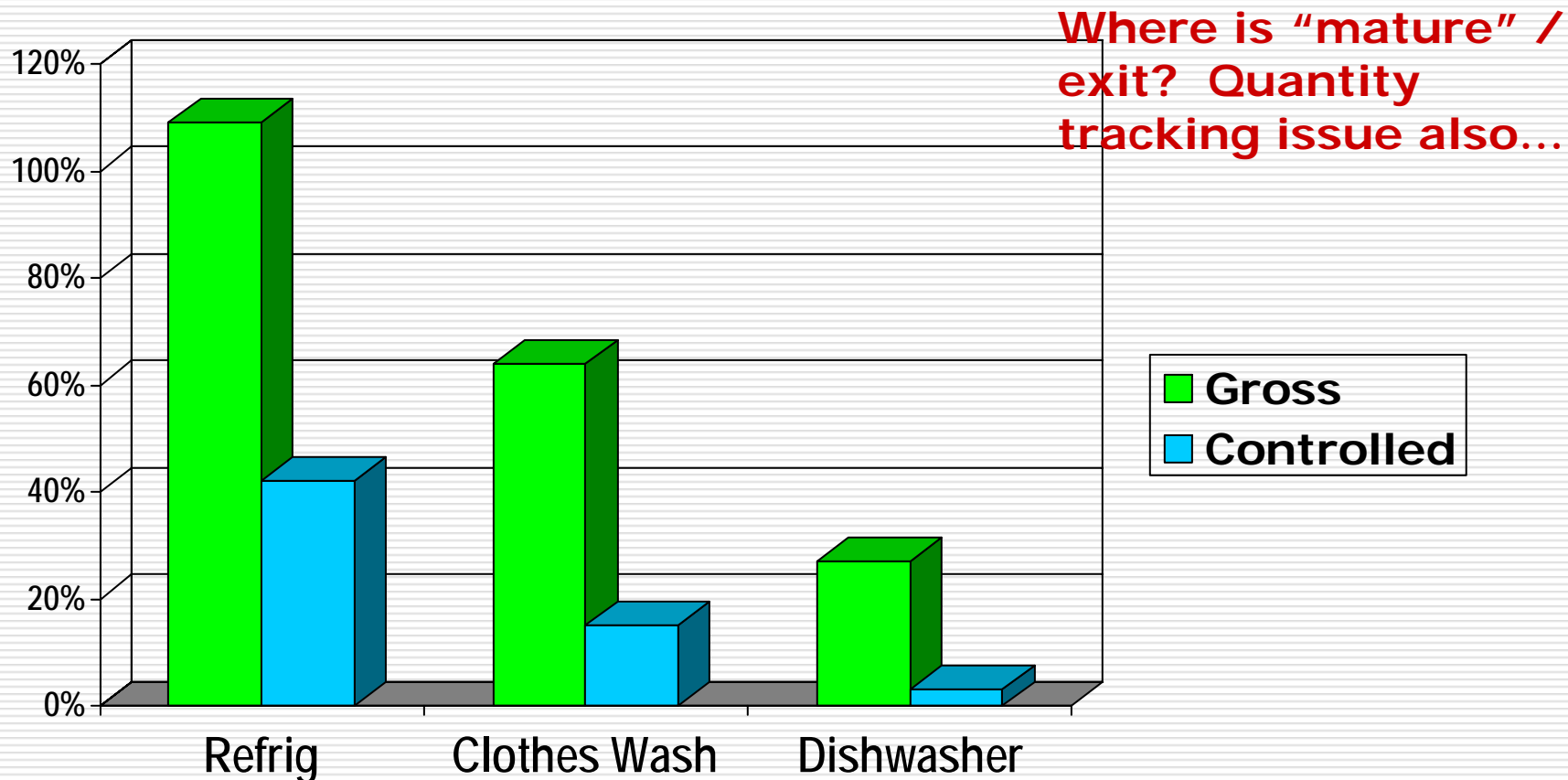
EVALUATION OF TRACKING OPTIONS

Method	Assessment	\$	Q
Periodic sales data from dealers, mfg	Balky/little incentive, business sensitive, labor intensive / goodwill, representativeness issues, data consistency	VH	VH
Shipments	Inconsistent, reselling, state boundaries, available for some commodities	M	H
In-store models tracking, shelfspace	Inaccurate for quantities; not very practical for comm'l.	M	M
Dedicated or Saturation surveys; purchases	Lack of knowledge, incentive to respond, sample on new purchasers hard; sample size comm'l	H	ML
On-site inspection / audit	Better knowledge, sample size issues; finding sample	H	VH
Industry / assoc. sales data	Some sales data / limited equipment; much equipment not available	M?	H
Warranty cards / rebate	Promise; depends on product; control sales missing	M	M

REFLECTING PROGRESS IN MARKET – 2 WAYS...



GROSS VS. CONTROLLED ENERGY STAR PRICE PREMIUM – REFLECT DIFFERENT “STAGES”



Where is “mature” / exit? Quantity tracking issue also...

Refrig, C/W Evolving; D/W mature

PARTIAL APPROACH – PRICE METHOD

- Doesn't attribute to program – total market approach
- Substitute / augment sales method
- Other benefits
 - Easy data collection
 - Indicate when to exit market – set goals
 - Mimics decision-making
 - Understand key decision drivers – help program
 - Fits with process & impact...

IMPLICATIONS BACK TO PROGRAM DESIGN

NET TO GROSS / ATTRIBUTION

NET TO GROSS / ATTRIBUTION

- Attributable savings
 - Market effects / untracked savings
- Interviews / literature review shows varied approaches
 - Deemed (1, 0.8, 0.7, etc.)
 - Models with dynamic baseline
 - Paired comparisons of stores
 - Largely self report batteries with corroborating information / influencers / triangulation
 - Most only Free Rider – not spillover, virtually no non-participant spillover measures

NET TO GROSS / ATTRIBUTION

- Conclusions / Issues / Questions:
 - Some not measuring indirect effects – considered MT
 - Need more market assessment
 - Trade allies critical –
 - Impossible to model / predict - dynamic
 - Changes upstream critical
 - Gather data close to point of decision
 - Focus on understanding behaviors is
 - Baseline / program record problem...

NET TO GROSS / ATTRIBUTION

- Conclusions / Issues / Questions:
 - No reason why can't look at NTG for education / training – just rarely done.
 - Some deemed – others say deemed savings ridiculous –
 - Cannot improve programs with deemed – must feed back on when to exit, how to improve, when to delete or upgrade measures
 - FR affected by NEBs
 - Ignoring spillover a problem...they are MT effects; hard to measure and make control group problem harder - and isn't that some of the education program key influence?
- Best methods- Experimental design – no substitute – measurement essential to improve programs

NET TO GROSS / ATTRIBUTION – BEST METHODS

- Methods well known for 20+ years
- OK: Quasi-experimental – control group state
 - Pure no-program states, etc.
- Best: Experimental Design - no real substitute for experimental design
 - Random assignment into treatment / non-treatment
 - Over time
 - Don't forget vendors
- Use discrete choice models with explanatory variables – unobserved factor issue – with true control group
- Overall Quote: “not measuring is not the answer” (Rufo)

IMPLICATIONS BACK TO PROGRAM DESIGN

MEASURE LIFETIMES

MEASURE LIFETIMES / EULs

- Key part of attributable effects
- Methods:
 - Fairly standard statistical approach
 - Some states look to others (e.g. CA)
 - Some look at 10 years for all industrial
 - A few looking at age in the field (furnaces)
- Virtually unmeasured for behavioral and education programs...
- Compare, and compare over time, and verify
- Variations by climate, usage
 - Need logger studies for some types
- Early removal – remaining useful lifetimes (RULs)
 - Little study / much discussion – some fans, some not

IMPLICATIONS BACK TO PROGRAM DESIGN

NET NON-ENERGY BENEFITS (NEBS)

NON-ENERGY BENEFITS / IMPACTS - NEBS

- Defined: NET positive and negative effects beyond energy / bill savings attributable to program –
 - Omitted program effects (incremental)
 - Naming convention
- Most common treatment –
 - Carbon adder
 - Economic development / jobs
 - Specific “direct” / quantifiable; lists, descriptions
- Issue of revised C/E tests...
- KEY information for improving programs
 - Influencers in participation; influencers in free riders for industrial
 - Negative – IDs concerns about performance, etc. of use to program design

NEBS – THREE PERSPECTIVES

	Utility	Participant (Res & C/I)	Societal
NEBs	Financial, T&D, reliability, etc.	Water / other bills, service quality (comfort, maint, productivity), psychic (enviro, pressure) behavioral/understanding, other	Economic dev'p, Enviro, Health, commun.value, infrastructure
Measurement method(s)	Direct, billing records, engineering	Direct, survey (WTP/A, relative, ranking, etc.); statistical, market based, other	Models (I/O, etc.), epidem,
Reliability	High reliability	Medium to low reliability (tradeoffs; replicability)	Medium to high reliab.
Relative size	Low value	Potentially large (value)	Large (value)

NEBS - MEASUREMENT

- Utility – direct measure / estimates
- Societal –
 - GHG models
 - Input-output models
 - Some direct measure
- Participant 8+ methods tested
 - with pros and cons

MEASURING PARTICIPANT NEBs

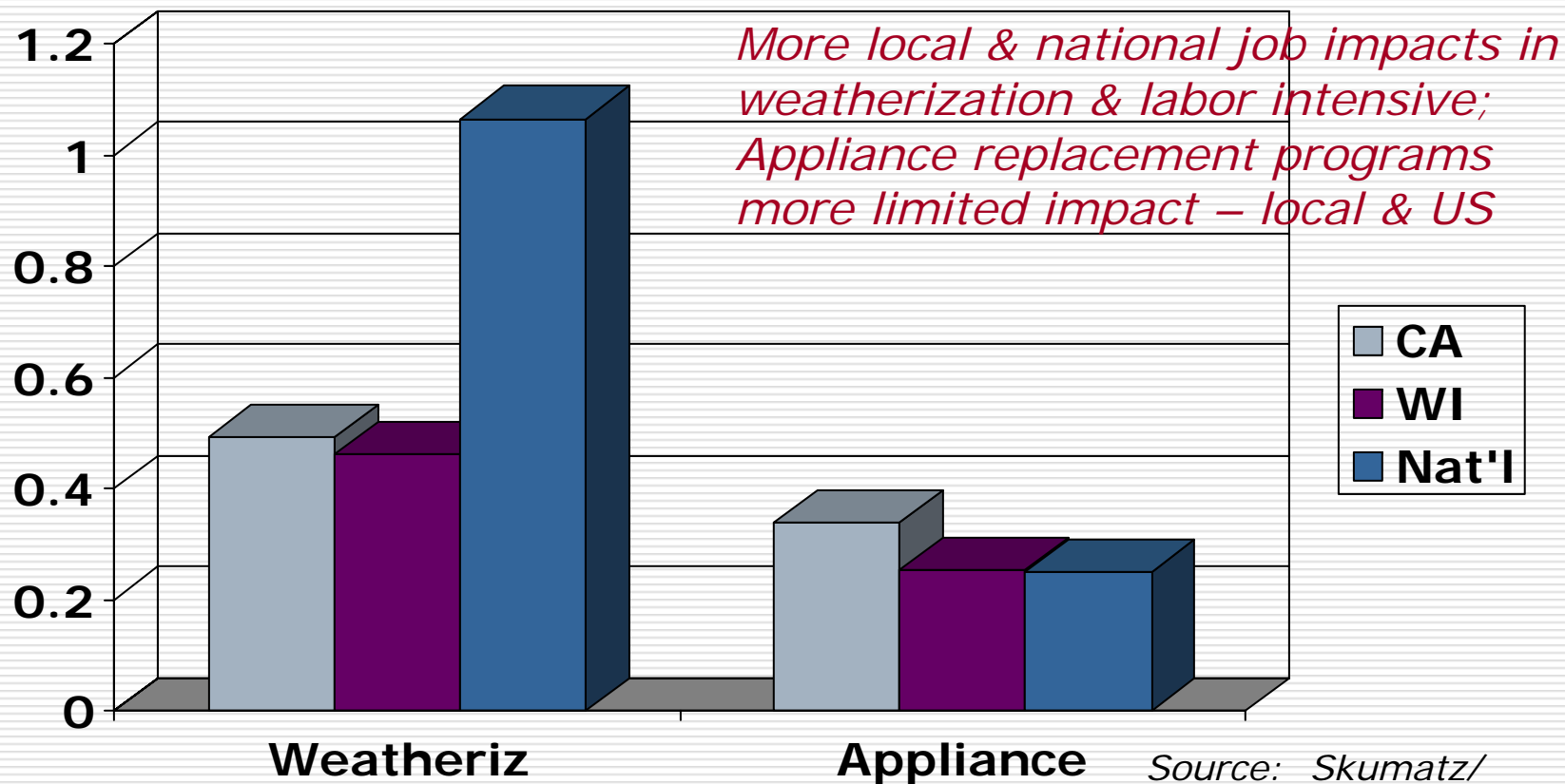
Method	Key Pros	Key Cons
Direct study	Some straightforward	Missing obs/bias
WTP / WTA	Literature	Volatile, uncertainty, hard
Comparable / relative	Fast, large sample, replicable, demonstrated in academic lit; multiple actors	Concerns about base
LMS	Fast, large sample clear, lists, replicable; academic lit; actors	Mistrust
Logit, conjoint, rank; discrete choice	Replicable; academic lit; actors	Expense, difficult to understand, limited cases
Regression	Robust*	Expensive; limited measures
Market value	Reliable when avail.	Rarely available, bias
Other	Explorations...	

Many variations to these basic approaches used in literature (Note) also used to disaggregate impacts*

SOCIETAL NEBs – MODELING FACTORS

- Environmental
 - Strong modeling work, strong value (significant share to multiples of direct energy savings)
 - Resource mix
 - Peak / off peak
 - Infrastructure impacts (minimal work)
- Economic
 - I/O models
 - Multiplier compared to baseline expenditures (improvements)
 - Potential for 2 impacts
 - Significant share to multipliers of direct energy savings

ECONOMIC MULTIPLIER – PATTERNS



Source: Skumatz/
SERA ACEEE paper 2006

Multiplier for indirect portion of economic output from NET transfer of funds FROM generation TO SIC/NAICS codes relevant to program's design.



NEBS APPLICATIONS

	Utility	Participant	Societal
Portfolio dev'p	Yes	Yes	Yes
Program refinem't	Yes	Yes	Yes
Marketing		Yes	*
B/C internal cust		Yes	*
B/C Tests	Potential (sm)	Potential	Potential (hi)

(*) these reflected in participant indirectly
 Multiple actor interviews provide robust inferences

TREATMENT OF NEBS ELSEWHERE

- Readily measured considerations vary – examples:
 - Future evaluation maybe; sooner possibly for high performance programs (BCHydro)
 - If defensible, include (carbon, water, infra) (OR)

	BCHydro	Oregon
Yes	Maintenance, GHG, lifetime, product loss / waste, equipment productivity, floor space	Carbon value on societal test; PV deferred plant extension, water/sewer examples
No	Sales, property value, satisfaction, worker / student productivity, H&S, comfort, noise, aesthetics, convenience, pride/prestige, environmental responsibility	

TREATMENT OF NEBS ELSEWHERE

NEBs Approach Conservative to Aggressive...	Program Screen	Examples
Program marketing only - Conservative	TRC	Ontario, Manitoba, Quebec
Scenario analysis	TRC	NY (many NEBs for scenario; programs must pass without NEBs)
Project screen	TRC	WI (caveats)
Program screen – readily measurable	Modified TRC PPT	MA (NEBs must be “reliable & with real economic value”); CA (% adder); CA (Low income), VT (maint, eqpt replacement) CO (measurable with current mkt values); NH (15% adder); BCHydro; OR
Program screen – broader NEBs	Modified TRC PPT	None

Source: Updated from original BCHydro work

NEB PATTERNS ACROSS PROGRAMS – SCENARIO RUNS

<i>Highest Benefits</i>	<input type="checkbox"/> Economic <input type="checkbox"/> Environmental <input type="checkbox"/> Property Value	<input type="checkbox"/> Water <input type="checkbox"/> Participant (comfort, etc.) <input type="checkbox"/> T&D, other
<i>Highest % Variations</i>	<input type="checkbox"/> Environmental <input type="checkbox"/> Water <input type="checkbox"/> Economics	<input type="checkbox"/> Arrears <input type="checkbox"/> H&S <input type="checkbox"/> Participant
<i>Highest \$ Variations</i>	<input type="checkbox"/> Environmental <input type="checkbox"/> Water <input type="checkbox"/> Economics	<input type="checkbox"/> Participant <input type="checkbox"/> H&S
<i>Program-dependent</i>	<input type="checkbox"/> Property value <input type="checkbox"/> Rate subsidy <input type="checkbox"/> Economic	<input type="checkbox"/> Water savings <input type="checkbox"/> Measure-based, H&S <input type="checkbox"/> Productivity

Based on NEB runs for variety of programs

TRC TEST APPLICATION – “SOCIAL”

Numerator

- Avoided cost of supplying electricity
- ? Environmental impacts
- ? Economic mult
- ? Market effects
- ? Other societal NEBs

Denominator

- Administrative costs
- Increased supply costs
- Capital cost
- ? Adjust for portion partic. would pay anyway / value
- ? Transaction costs
- ? Value of service lost

Start at TRC=“1” - scenarios

IF enviro=1.5 of savings (existing 10% or less not to scale)

IF econ=0.5 of savings → → → → TRC = 3

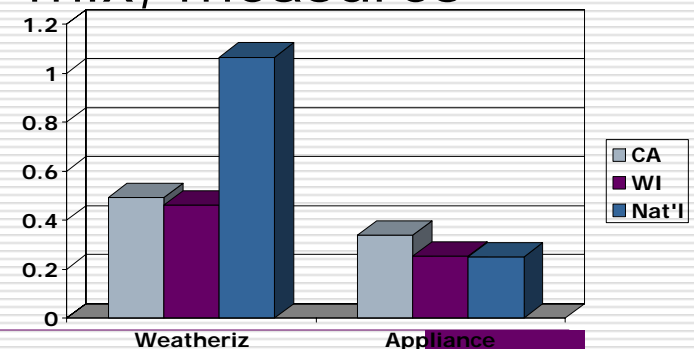
IF partic costs are 50% and would pay 50% → TRC = 4

20% → TRC = 5

Participant test EASILY to 2, depending on program...

TRC IMPACTS

- Goal NOT to increase TRC, but to better select among programs, spend dollars where best impacts.
 - Even directly / reliably measured NEBs significant
- Reliable measurements reflect important program differences, with variations for...
 - Peak / off-peak, fuel, resource mix, measures included, geographic etc.
 - DEEMED doesn't reflect



Start at TRC="1" - scenarios

IF enviro=1.5 of savings

IF econ=0.5 of savings → → → → TRC = 3

IF partic costs are 50% and would pay 50% → TRC = 4

20% → TRC= 5

SERA

OTHER RESULTS...

- Behavioral & educational programs - Applied to:
 - Energy Star (homes, home performance, appliances) in multiple states; repeated results / similarities
 - Weatherization programs with education
 - High performance commercial training
 - Schools programs
 - Other (event, etc. among others)...
- Issues
 - Actions from multiple precipitating influences - uncertainties
 - Impacts for education often not-well measured (exceptions, promise)
 - Need to allow for well-designed control groups; experiments
 - Retention barely touched / strong potential
- Strong NEBs – they are some of the main outcomes of these programs; regulatory tests discourage
 - Suggestion that NEBs maybe key driver for some actions
 - Research on influencing factors

IMPLICATIONS BACK TO PROGRAM DESIGN

SUMMARY

- Issues / transition
- Need information
 - But not at expense of delivery of equipment
 - Key to understand market better – before you jump in and during – key to program design
 - Actors / upstream
- “Deemed” poor approach
 - Never informs programs – good money after bad?
 - For Savings, NTG, or NEBs – Energy efficiency HELPS jobs – but some programs / measures help more than others...
 - Environment – peak / offpeak, mix by program
- Behavioral remain less studied / harder to measure – much more work needed (including EUL); but behavior also represent steps in adoption of ‘measures”
- ***NOT ENOUGH feedback from evaluation into design
- Research
 - Still in progress - Want approaches that meet many goals
 - Areas for additional research

CONTACT INFORMATION

Lisa A. Skumatz, Ph.D.

SERA, Inc.

762 Eldorado Drive, Superior, CO 80027

Phone: 303/494-1178

Email: skumatz@serainc.com

Web www.serainc.com

***Thanks to those that helped –
Weblink for YOUR input – We would LOVE your
input!***

www.serainc.com; click on link ½ way down on left

http://www.surveymonkey.com/s.aspx?sm=QMj8ZXYYsJ2UvHO2XT0gww_3d_3d