

# Using Experiments to Foster Innovation and Improve the Effectiveness of Energy Efficiency Programs

Presented by Michael Sullivan  
Freeman, Sullivan & Co.  
May 20, 2009

Sponsored by the California Public Utilities Commission  
Prepared for CIEE Behavior and Energy Program  
Edward Vine, Program Manager



FREEMAN, SULLIVAN & CO.

A MEMBER OF THE FSC GROUP

# Acknowledgement and Disclaimer

- ◆ Thanks to Ed Vine for his patience and careful review of the work
- ◆ Thanks also to the anonymous reviewers who took the time to read and comment on the effort in the paper and to offer advice concerning how to remedy the many defects that were present in the first draft
- ◆ Any errors that remain in the work are mine alone
- ◆ Thanks to the CPUC for their generous sponsorship of this work
- ◆ The findings, opinions and conclusions expressed in this presentation are mine alone and do not represent the opinions or policies of the CIEE or CPUC



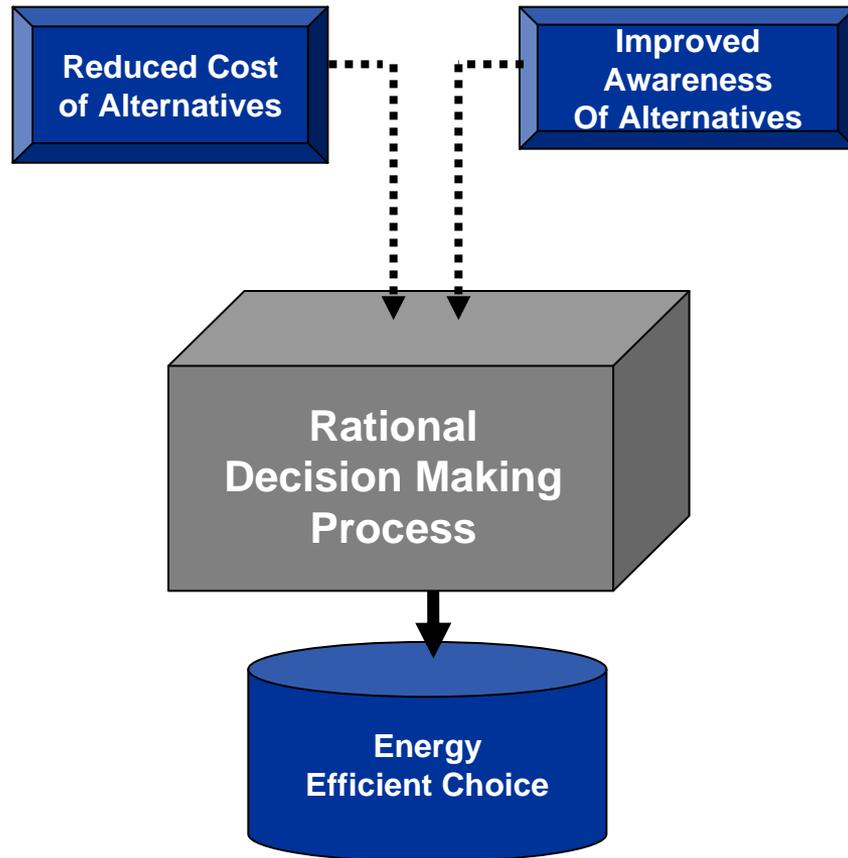
# The Bottom Line

1. Realization of much of remaining economically justifiable energy efficiency potential requires changing behavior of energy users
2. However, significant research and development will be required to improve the performance of energy efficiency programs using behavior change strategies
3. Fostering innovation in the development of energy efficiency programs based on behavior change is key
4. Innovation requires highly managed and focused research and development effort based on experimentation
5. Significant institutional barriers stand in the way of progress and these must be overcome

# Assumptions Underlying “Conventional” Energy Efficiency Programs



# Historical policy paradigm (PTEM) downplays role of consumer behavior in improving efficiency of energy use



- ◆ Efficiency improvement comes from changing technology not behavior
- ◆ Adoption decision is the exception
- ◆ Other consumer behaviors are of secondary concern
- ◆ Decision-maker is assumed to be economically rational
- ◆ Programs concentrate on providing information to improve awareness and incentives to offset first cost, risk and other costs

Physical Technical Economic Model (PTEM)



# Examples of Classic PTEM programs

- ◆ Tax credits for making building energy efficiency investments in buildings
- ◆ Direct subsidies to targeted market segments to make energy efficiency improvements
- ◆ Appliance rebates to consumers who purchase energy efficiency equipment
- ◆ Information programs designed to increase consumer awareness and knowledge of energy efficient alternatives
- ◆ Stimulus package contains massive infusion of capital directed at PTEM-based programs

# Strengths of the PTEM Paradigm

- ◆ Conventional measure-based PTEM model fits well with policy guidance to achieve energy efficiency as a way to avoid unnecessary and wasteful energy cost
- ◆ It is easily integrated with decisions about energy resource acquisition for utilities
- ◆ Impacts *seem to be* easily measured and verified
- ◆ Persistence of measures *seems* more likely than persistence of behavior change
- ◆ Economic worth of savings easily calculated for purposes of cost-benefit analysis

# Weaknesses of PTEM-Based Energy Efficiency Programs

- ◆ Programs developed under this paradigm are often not very effective
- ◆ Decisions by consumers and businesses are shaped by many considerations beyond perceived costs and benefits
  - Most decisions involve acquisition of something other than the raw output of the purchased device (e.g., automobile provides both transportation and social status)
  - Information about energy efficiency is limited and costly
  - Costs of discovery and implementation can greatly exceed the economic worth of energy savings
  - Economic rationality depends on assumptions about future energy costs
  - Economic rationality requires probabilistic thinking
- ◆ Much of human behavior that affects energy use is not really a choice – exclusive focus on measures leaves a lot on the table
  - Trip to work
  - Daily personal hygiene

# Recent Intensive Efforts Have Focused on Low Hanging Fruit

## ◆ Commercial and Industrial

- Lighting 60-70% of savings in sector
- Lighting 25-30% of all savings
- Process 15-18% of savings in sector
- Process 6-8% of all savings
- Other 7-20% of savings within sector
- Other 3-6% of all savings

## ◆ Residential

- Lighting 75-90% of savings in sector
- Lighting 35-50% of all savings
- Refrigeration 5-20% of savings in sector
- Refrigeration 2-10% of all savings
- Other 2-5% of savings within sector
- Other 1-3% of all savings

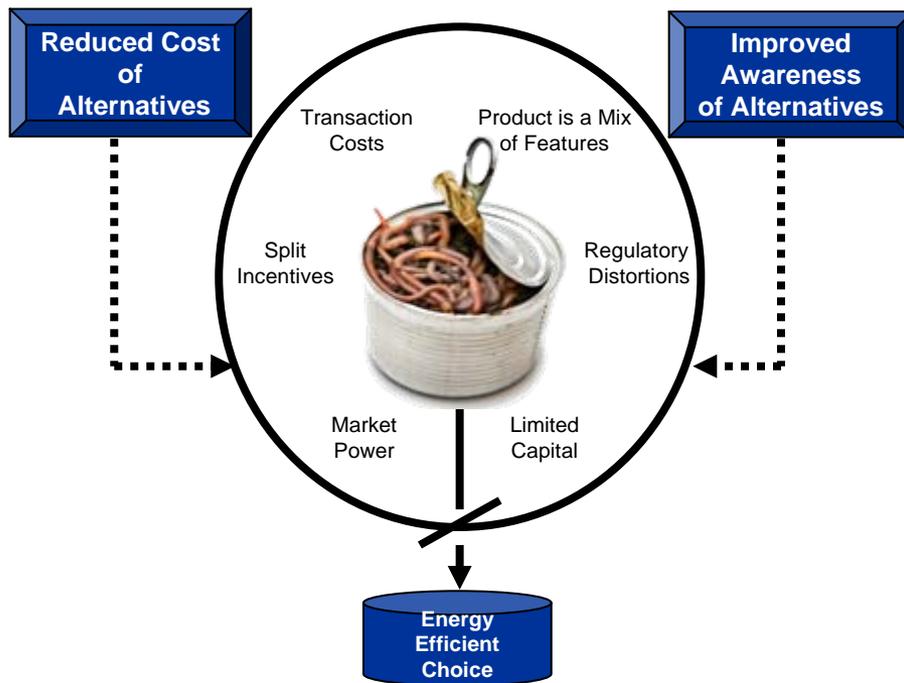
From "Meeting Aggressive New State Energy Goals for Utility Sector Energy Efficiency: Examining Key Factors Associated with High Savings," Martin Kushler, Dan York and Patti Witte, March 2009, WCEEE Report Number U091

# And After the Low Hanging Fruit?

- ◆ It is an uphill battle
- ◆ It is difficult to persuade consumers to make significant capital investments in energy efficiency based solely on costs and benefits
  - Availability of capital
  - Risk of new technology in production processes
  - Perceived risk of obtaining the benefit over time as fuel prices fluctuate
  - Unusually high discount rates applied by consumers
- ◆ Transaction costs often exceed projected benefits in mass markets
- ◆ Most energy efficiency program designers are skeptical of the ability of incentives and information to achieve significant gains beyond those that are achieved in the low hanging fruit
- ◆ There is a growing interest in developing programs that take advantage of other factors influencing behavior

# Beyond Economic Rationality

# Beyond Economic Rationality



- ◆ Behavioral school of thought about consumer decisions concerning energy efficiency principally dominated by:
  - Psychology
  - Sociology
  - Anthropology
  - Organizational Behavior
  - Behavioral economics
- ◆ Energy use (and energy efficiency) are viewed as byproducts of human actions
  - Mobility
  - Sustenance
  - Security
  - Household maintenance
- ◆ Decision to purchase energy using equipment is just one of many behaviors that humans exhibit that affect their energy consumption

# Beyond Economic Rationality – observations about energy efficiency decisions from behavioral science

- ◆ Consumers weigh a number of considerations in making purchasing decisions beyond costs and benefits
- ◆ Rationality is only one of several decision-making heuristics
- ◆ Motivations other than economic gain often dominate decision-making
- ◆ Social influences can dramatically influence decision outcomes
- ◆ Feedback on consequences of behavior can change it
- ◆ In organizations, opinions, beliefs, desires and objectives of the management matter
- ◆ Transaction costs strongly influence decision outcomes

# Beyond Economic Rationality – interesting possibilities

- ◆ More effective advertising messages and marketing channels based not on minimizing energy cost/benefit ratio but on other consumer needs
- ◆ Improving efficiency of program delivery by identifying and marketing to “susceptible parties”
- ◆ Using well-known social mechanisms such as norms, social networks and social influence hierarchies to improve receptiveness of new technological alternatives
- ◆ Marketing through alternative information channels (e.g., community organizations, service groups, etc.)
- ◆ Providing feedback to consumers regarding the consequences of their choices
- ◆ Undertaking serious long-term efforts to change attitudes, opinions and social norms to improve efficiency of energy use (ala anti-smoking campaigns)

# Possibilities – Unfortunately are not Programs

- ◆ Significant research and development will be required to develop improvements in energy efficiency programs taking advantage of knowledge from the behavioral sciences
- ◆ What is needed is innovation – revolutionary ideas translated into new products and services

# Management of Innovation

# Innovation

- ◆ Innovation is a kind of market transformation activity
- ◆ Revolutionary changes in products and practices seldom happen by accident
- ◆ They result from a painstaking process of trial and error – in essence experimentation
  - To perfect the product or practice
  - and to assimilate it into the market
- ◆ Most innovations today are the result of carefully managed efforts to produce new products that ensure competitive position and respond to changing consumer tastes and needs
- ◆ Innovation management techniques can be used to develop more effective energy efficiency programs

# Management of Innovation – an example, the Prius



# Development of the Prius

- ◆ **1993 – Toyota’s Chairman expresses concern about future of the automobile market**
  - VP of R&D establishes a project to develop a new small car (the G21)
  - Objective – develop a car with 50% better fuel efficiency than Corolla (its best selling small car)
- ◆ **1994 – Basic Concept submitted to management**
  - Team instructed to produce concept car by 1995
  - 80 alternative designs considered; narrowed down to 4 with hybrid in the lead
- ◆ **1995 – Work begins on hybrid prototype**
  - Target for production for Japan in 1997 – just two years from start to finish
  - Marketing skeptical about customer acceptance but designers press on
- ◆ **1995-1997 – Many serious engineering problems in addition to normal product development problems**
  - 1,000 engineers working full time
  - Transition between gasoline and electric modes
  - Electronic controls
  - Battery safety and performance
  - One by one problems were eliminated

# Development of the Prius

- ◆ **1997 – First models introduced in Japan**
  - Cost to develop and produce first 2,000 cars \$1 Billion (about average for new model)
  - Immediate success in the market – demand far outpaces supply
- ◆ **1999 – First Prius is delivered in California**
  - Toyota US still skeptical of market acceptance when vehicle first introduced
  - There were issues – US consumers didn't like the feel of the brakes, there wasn't room in the car for baby strollers and other big stuff
  - Focus groups produced disappointing results
  - It was a Japanese car
  - Back to the drawing board
  - Honda Insight comes to market – tepid response
- ◆ **July 2000 – Revised “California” Prius Arrives**
  - As in Japan, huge unexpected success
  - Met a variety of consumer needs (inexpensive, hip, high status, low-cost operation)
  - Demand exceeded supply for months
- ◆ **Today –**
  - Strong demand for Prius continues despite downturn in automobile industry
  - In some California counties, hybrid sales exceed 20% of the market

# Management of Innovation – Lessons from Prius

- ◆ Innovation happens by trial and error – a painstaking process of experimentation in which better ways are found of doing things by trial and error
- ◆ It takes time
- ◆ It can be very costly
- ◆ It takes a lot of effort and commitment on the part of the development team to make something new that the market wants
- ◆ It isn't just solving technical problems, it is solving market acceptance problems
- ◆ It can and must be managed
- ◆ It can be massively worth the effort

# Managing the process of innovation

## Normal Product/Service Development Process



### Applies to all kinds of development problems

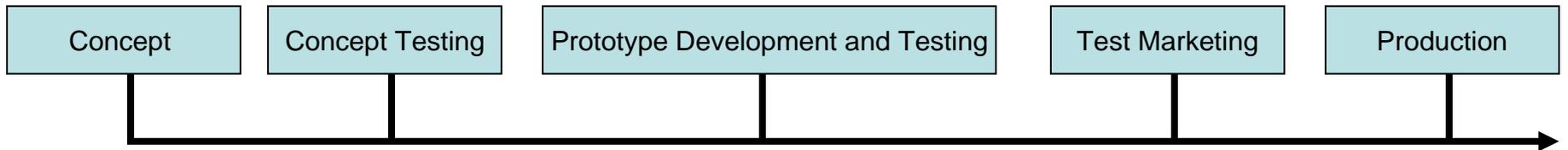
- ◆ Automobiles
- ◆ Consumer electronics
- ◆ Banking
- ◆ Pharmaceuticals
- ◆ Food and beverage
- ◆ Energy Efficiency Programs?

# Principles for Managing Innovation

1. Innovation requires project champions (sounds hokey but that is what they are) who are personally invested in development and responsible for knowing all aspects of the development problem
2. Maximum utility must be obtained from information early on in the development process
3. Experiments must be done frequently to test design alternatives
4. Experiments must be done rapidly to not hold up development
5. Failure should happen early and often, while avoiding mistakes

# Implications for Energy Efficiency Program Development

## Normal Product/Service Development Process



## Typical Energy Efficiency Program Development Process



Development process for EE programs frequently shortcuts all important research, development and demonstration steps – leading to faster implementation of often ineffective programs that are hard to terminate

# Experiments or Pilots – what's the difference?

## Pilots can be experiments, but usually they are not

### Pilots

- ◆ More or less full-scale test of program within a limited geographical area or for a short time
- ◆ Test of effectiveness of fully-developed prototype
- ◆ Integrated within utility production organization
- ◆ Headed for full-scale implementation unless fatal problem is discovered
- ◆ Evaluation of ex-post impacts and process evaluation

### Experiments

- ◆ Small-scale tests designed to conclusively determine whether a given program design alternative works better than another
- ◆ Multiple program design alternatives tested
- ◆ Careful attention to research design to allow conclusive decision about how to proceed
- ◆ Often not integrated with production organization
- ◆ Full-scale implementation may or may not be contemplated - depends on outcome of test

# Example of DSM Program Development Problem

## Problem

- ◆ Thousands of customers on emergency load control program
- ◆ Program not cost-effective at current annual incentive levels to customers
- ◆ Could be converted from emergency (dispatched only in system emergency) to economic (dispatched when wholesale prices are extremely high) dispatch design to improve economic performance
- ◆ Considerable sunk cost in program implementation could be lost if significant attrition occurs during conversion
- ◆ Several options for moving customers

# Example of DSM Program Development Problem

## Conversion Options

1. Offer customers the choice of continuing the program at significantly reduced annual incentive or signing up for a new program based on pay for performance
2. Offer customers the choice of continuing program at significantly reduced annual incentive or critical peak pricing in combination with TOU rate
3. Offer customers a menu of reliability choices at prices varying with their willingness to experience different levels of operational severity
4. Assign existing customers to economic performance-based program with the ability to opt out

# Example of DSM Program Development Problem

## Not Obvious How to Proceed

- ◆ Focus groups and customer surveys will suggest how customers might respond but usually the results are not conclusive
- ◆ Very significant losses may be experienced if a reasonable way of converting customers is not found – it might even be better to do nothing
- ◆ Experiment or Pilot?
- ◆ Should favor an experiment but probably ends up a pilot because of the difficulty in integrating test into normal billing operations and requirement to obtain regulatory approval

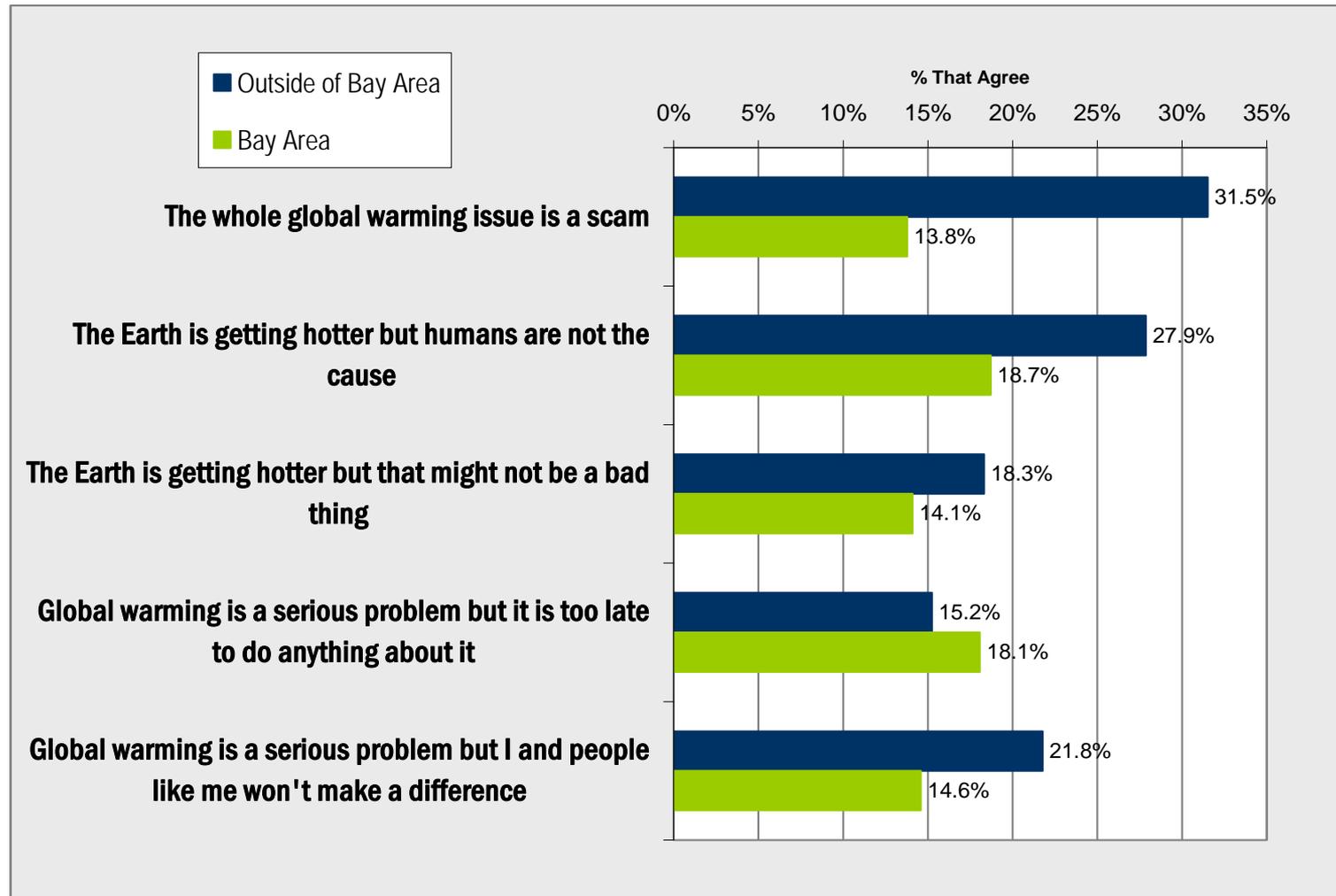
# Experimental Design



# What's the big deal about experiments?

- ◆ Innovation is accomplished through a trial and error process in which things that work are separated from things that don't
- ◆ Trial and error studies of humans and human systems generally require the use of rigorous scientific investigation techniques (i.e., formal experiments)
  - Studies of ad hoc groups and studies where behavioral interventions are not controlled are generally inconclusive and provide little guidance in program design
  - Not all humans and human systems are alike (i.e., differences in emotions, intelligence, cultural milieu, values, business models, etc.)
  - Not all environments in which humans are operating are alike (i.e., affiliation and family groups, communities, etc.)
  - Interactions between different kinds of humans and their environments strongly influence observed behaviors
  - Necessary to control these factors to conclusively determine what the effect of a given change really is
- ◆ Experiments provide a basis for conclusively determining whether a change in behavior has occurred

# Example of Differences among Population Members



# Issues in the Design of Experiments

- ◆ **Basic Principles (John Stuart Mill)**
  - Causes precede effects
  - Causes are correlated with effects
  - There are no other plausible explanations for observed effects
- ◆ **Internal Validity (other plausible explanations?)**
  - History
  - Maturation
  - Testing
  - Instrumentation
  - Statistical Regression
  - Mortality
  - Selection
- ◆ **External Validity (is it legitimate to generalize result?)**
  - Experimental subjects are different from population of interest
  - Experimental situation is different from the situation of interest
  - Experimental treatment is different from real world treatment

# Example – Impact of targeting on marketing efficiency

## Objective

- ◆ Evaluate effectiveness of targeting algorithm in increasing likelihood of acceptance of energy efficiency offer

## Approach

- ◆ Apply targeting algorithm to customers selected to receive offers and select 5,000
- ◆ Compare customer response from targeted customers with response received from the prior year when targeting wasn't used

## Result

- ◆ 5 fold increase in likelihood of adoption of energy efficiency alternative year on year

## Internal Validity Problems

- ◆ Not terrible, but troubling nonetheless
  - **History** – from year to year something other than the targeting that is responsible for the observed improvement in performance (e.g., fulfillment contractor may have changed practices)
  - **Maturation** – population may have become more “susceptible” to program
  - **Testing** – maybe, not probable

## External Validity Problems

- ◆ None

# Classic Experimental Design

	Pre-Test	Post-Test
Treatment Group	$T_{pre}$	$T_{post}$
Control Group	$C_{pre}$	$C_{post}$

$$Effect = (T_{post} - T_{pre}) - (C_{post} - C_{pre})$$

# Classic Experimental Design

## Strengths

- ◆ Treatment is randomly assigned resulting in two statistically identical groups
- ◆ Outcome variable of interest is known prior to and after exposure to the treatment
- ◆ Difference in outcome between treatment and control groups can be attributed solely to causal mechanism of interest
- ◆ Test results can be conclusive

## Weaknesses

- ◆ Sometimes impossible to randomly assign observations to experimental conditions
- ◆ Statistical noise may mask subtle experimental effects

# Targeting Example Revisited

## Experimental Design Alternative

- ◆ Randomly select 1,000 customers who are eligible for marketing
- ◆ Randomly assign half to an experimental group and half to a control group
- ◆ Apply the targeting algorithm to the experimental group, sort in order of the estimated propensity to act and record the order of each group member in list
- ◆ Sort the control group according to a random variable and record the order of each group member in the list
- ◆ Contact and market to both groups in sequence according to the order variable taking care to make sure the marketing organization is not aware of the test
- ◆ Compare the average value of the order variable for the two groups
  - Average propensity to subscribe should be the same in both groups
  - If there is no difference in the average value of the order variable between the two groups, the targeting algorithm has no effect
- ◆ Use log-linear modeling to estimate the impact of the order variable on the likelihood of participation

# Targeting Example Revisited

## Improvements over Demonstration Design

- ◆ Impact of targeting program is unambiguously measured – no plausible explanations exist for the test outcome beyond the impact of the targeting algorithm – test is completely conclusive
- ◆ Provides more information about the relationship between predicted propensity to participate and likelihood of actual participation
- ◆ Smaller number of marketing contacts required to complete the test

# Interesting Variations on the Classic Design – dealing with noise and complexity

## Randomized Blocks

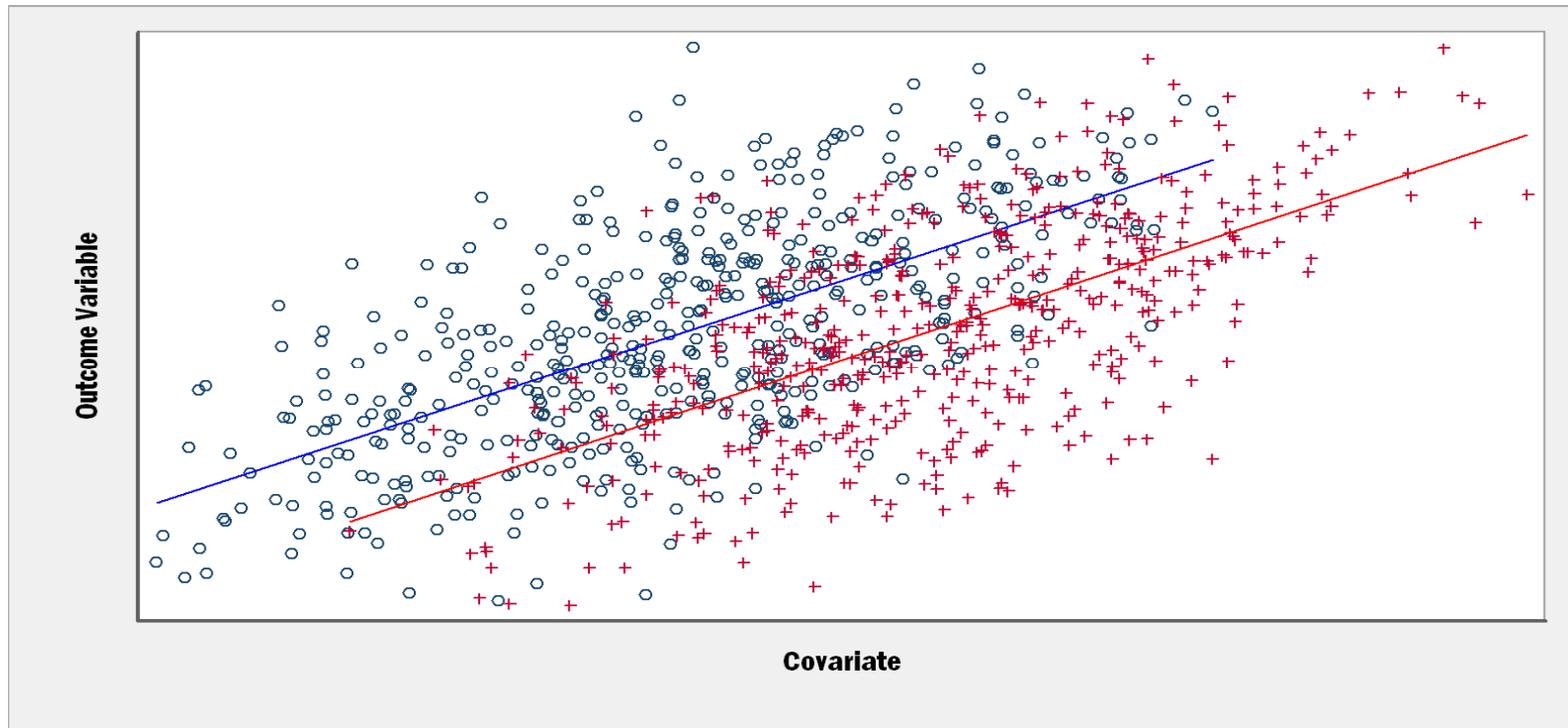
Blocking Factor	Group	Pre-Test	Post-Test
Stratum 1	Treatment Group	$T_{pre}$	$T_{post}$
	Control Group	$C_{pre}$	$C_{post}$
Stratum 2	Treatment Group	$T_{pre}$	$T_{post}$
	Control Group	$C_{pre}$	$C_{post}$
Stratum 3	Treatment Group	$T_{pre}$	$T_{post}$
	Control Group	$C_{pre}$	$C_{post}$

## Factorial

		Group	Factor 2	
			Level 1	Level 2
Factor 1	Level 1	Treatment Group	$T_{11}$	$T_{12}$
		Control Group	$C_{11}$	$C_{12}$
	Level 2	Treatment Group	$T_{21}$	$T_{22}$
		Control Group	$C_{21}$	$C_{22}$
	Level 3	Treatment Group	$T_{31}$	$T_{32}$
		Control Group	$C_{31}$	$C_{32}$

# Interesting Variations on the Classic Design – dealing with noise and complexity

## Covariance Designs



# Quasi-Experimental Designs

- ◆ Discussed in the appendix to the paper
- ◆ When random assignment to experimental and control groups is impossible or impractical, it is still possible to use the logic of experimentation
- ◆ Much effort is required to create measurements that take into account the lack of randomization
- ◆ A number of possibilities
  - **Regression Discontinuity Analysis** – when arbitrary and specific assignment of treatment has occurred (i.e., annual maximum demand exceeding some level) look for discontinuity in the trend line on the outcome variable above and below the assignment criteria. This approach is as powerful as classical experimental design
  - **Non-Equivalent Control Groups** – create a control group, that as much as possible, matches the characteristics of treatment group (e.g., similar individuals in a different community, similar communities for community level interventions)
  - **Interrupted Time Series** – look for differences in the outcome variable of interest before and after onset of the treatment – can be extremely powerful particularly if repeated measures of treatment administration are used (e.g., load impact analysis)

# Short List of Badly Needed Experimental Programs

# Closing Thoughts on Experimentation

- ◆ Experimentation can be complicated, expensive and time consuming
- ◆ Experiments should not be undertaken without considering the benefits and costs in terms of time and resources
- ◆ Everything doesn't have to be subjected to experimental test – but when we want conclusive answers about whether or not a given approach causes behavior change, an experiment is probably required
- ◆ The key to success in innovation lies not in slavishly applying experimentation to test all possible improvements to programs but in strategically applying experimentation to obtain answers to critical questions

# Badly Needed Experimental Programs – Feedback

- ◆ Research carried out over the past 20 years indicates that feedback provided in a variety of ways can significantly lower energy consumption – results from prior studies range from 0% to 20%
- ◆ Causes of change in energy use are not well understood
- ◆ Results from these studies are suggestive but far from conclusive because they typically have been carried out as demonstrations rather than experiments and generally do not operate long enough to measure persistence
- ◆ Potential drivers of behavior are:
  - Price
  - Technology and communication channel (e.g., real-time information about energy use, supplemental information provided with billing, supplemental information outside of billing, estimated energy use, etc.)
  - Household characteristics
  - Message format and content
- ◆ These factors undoubtedly interact
- ◆ A focused experimental effort is needed to find and test cost effective approaches to combining information about price, consequences of action and information about decision alternatives that can achieve significant reductions in household energy use
- ◆ Could be the holy grail

# Badly Needed Experimental Programs – Community Level Interventions

- ◆ Community level interventions – full court press at the community level (e.g., schools, community groups, political infrastructure) designed to promote energy efficiency
- ◆ Has been very effective in promoting smoking cessation, reducing health risks from sexually transmitted diseases and lowering incidence of DUI
- ◆ Hundreds of millions of dollars spent on local government partnerships in about 40 communities in California
  - Local governments are in charge
  - Wide range of approaches
  - Difficult to determine what, if anything, is working and what isn't
- ◆ Framework for systematic experimentation is present but not utilized for studying effectiveness and experimenting with new ideas

# Badly Needed Experimental Programs – Development of Alternative Marketing Messages

## Messaging

- ◆ Cost/Benefit is only one of many messages that decision makers might respond to. Others include:
  - National security
  - Preserving the future (i.e., what are you leaving your children)
  - Climate change
  - Responsible citizen
  - Supporting community needs (for business)
- ◆ Messages are not one size fits all and their effects are not stable over time
- ◆ Experimental work to discover effective messages has to be ongoing
- ◆ Experimental designs for evaluating message impacts are well developed and inexpensive to carry out
- ◆ Market segmentation is key to understanding the effectiveness of messages

# Badly Needed Experimental Programs – Development of Targeting Strategies

## Targeting

- ◆ Receptiveness to energy efficiency offers varies within the market
- ◆ Development of effective methods for identifying “receptive” targets in advance of contact may dramatically improve the efficiency of marketing
- ◆ Evidence also suggests that program participants have a tendency to repeat and that the effects of program participation on likelihood of participation can be cumulative
- ◆ So far efforts at targeting have concentrated on development of market segmentation schemes that are predictive of receptiveness to programs
- ◆ The effectiveness of these schemes should be carefully evaluated using experimental methods to determine how much targeting via existing segmentation schemes improves performance
- ◆ Continuing efforts should be made to improve on existing schemes and discard those that do not improve performance

# Institutional Barriers to Innovation

# Discussion of Institutional Barriers to Innovation

- ◆ Some aspects of regulatory environment do not foster innovation in development of energy efficiency programs
  - By design, responsibility for energy efficiency R&D rests with CEC's PIER not with utilities
  - PIER has been focused almost exclusively on R&D related to technology with very little effort focused on behavior change
  - Utilities are paid for achieving energy savings in the short run and R&D penalizes economic performance of utilities in the short run – adds to cost of programs – this discourages investment in R&D
  - Except for relatively small resources dedicated to concept development and testing, utilities set aside few resources for R&D for improving long-run program performance based on behavior change
  - EM&V is principally focused on verifying savings and program management activities and not on program development per se – purpose is not R&D but accounting and management
  - Recent conflicts over realized savings provide strong evidence of the seriousness of these problems

# Discussion of Institutional Barriers to Innovation

- ◆ To overcome aforementioned barriers
  - Overhaul regulatory apparatus so that utilities are provided with funding specifically earmarked for R&D designed to develop energy efficiency program improvements based on behavior change
  - Develop a framework within which utilities and regulatory staff can come to agreement on reasonable short-term and long-term R&D objectives as well as appropriate research protocols to be applied to answer specific types of questions
  - Closely monitor utility performance in achieving R&D objectives including development of organizational capability to efficiently bring new products and services to market
  - Reorganize regulatory staff as necessary to allow them to oversee the work of utilities in designing and bringing new energy efficiency programs, based on behavior change, to market

## Using Experiments to Foster Innovation and Improve the Effectiveness of Energy Efficiency Programs

# For any questions, feel free to contact

Michael J. Sullivan, Ph.D.

Freeman, Sullivan & Co.

101 Montgomery Street 15<sup>th</sup> Floor, San Francisco, CA 94104

[michaelsullivan@fscgroup.com](mailto:michaelsullivan@fscgroup.com)

415.777.0707

