

# Comparison Groups for Whole Building Program Impact Evaluations:

They are Harder and Easier Than you Think

Katherine V. Randazzo, Ph.D., Opinion Dynamics; Richard Ridge, Ph.D., Ridge & Associates; Stef Wayland, Opinion Dynamics; Brian A. Smith, Pacific Gas and Electric

2017 IEPEC Conference — Baltimore, Maryland

# So, What's the Problem?

- RCTs for net savings = gold standard, but often not possible
  - Researchers don't discuss comparison group (CG) issues and choices enough
  - We need to think and write more about whole building CGs
- Confusion out there on what we get with CGs: Net or Gross
- CGs hard w/single-measure programs
- For whole building programs even more complex

# Comparison Groups & The Counterfactual

- The right CG can yield net effects
- Represents the naturally-occurring rate *IF*:
- Equivalent to participants on these factors:
  - History,
  - Selection,
  - Maturation,
  - Statistical Regression to the Mean,
  - Testing, &
  - Instrumentation (Shadish, Cook, & Campbell 2002)

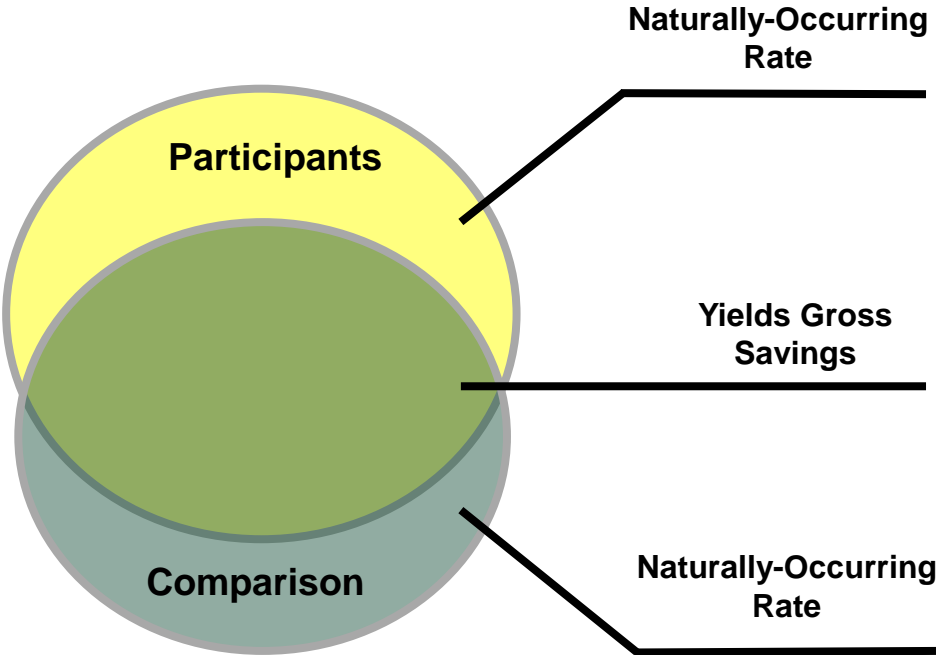
# Comparison Groups & The Counterfactual (Cont)

- Factors relevant to our industry:
  - Weather (History & Selection),
  - Economic Conditions (History),
  - Building Characteristics (Selection)
  - Occupancy Characteristics (Selection),
  - Geographic Location (Selection),
  - Motivations & Attitudes & Behavior (Selection),
  - Changes in those (Maturation, Statistical Regression to the Mean), **as well as**,
  - **Naturally-occurring** relevant installations & motivations to install ([Self-] Selection, maybe Maturation)



# Comparison Groups & The Counterfactual (Cont)

- Controlling for all but the **naturally-occurring** installations & motivations will produce **gross savings**.
- Not net.




# Comparison Groups & The Counterfactual (Cont)

- The naturally-occurring rate of installing program-promoted measures includes:
  - *What was installed, and*
  - *The motivation for the installation*

## Single-Measure Programs

What do we look for to represent the counterfactual or the naturally-occurring rate of installation?


Considering only **what was installed** 

1 Participant-Installed Program Measure	2 Installed Measure that Would Qualify a Non-Participant CG Member for Estimating Net
SEER 17 AC	Any AC
SEER 19 AC	Any AC
Tankless Water Heater	Any Water Heater
Duct Sealing	Any bldg. w/working HVAC system?
R30 Insulation	Any Insulation? None?
Envelope Sealing	Any bldg. w/working HVAC & no sealing?

# Comparison Groups & The Counterfactual (Cont)

## Whole-Building Programs

What do we look for to represent the counterfactual or the naturally-occurring rate of EE upgrades?

Considering *only what was installed* 

- Must consider these measures as a group
- Any combination could appear in prog if enough savings
- But what combinations necessary for naturally-occurring EE upgrades?

1 Participant-Installed Program Measure	2 Installed Measure that Would Qualify a Non-Participant CG Member for Estimating Net
SEER 17 AC	Any AC
SEER 19 AC	Any AC
Tankless Water Heater	Any Water Heater
Duct Sealing	Any bldg. w/working HVAC system?
R30 Insulation	Any Insulation? None?
Envelope Sealing	Any bldg. w/working HVAC & no sealing?

# Comparison Groups & The Counterfactual (Cont)

## Whole-Building Programs

- The good news: The pool of potential CG members large, because:
  - Any upgrade of 2+ installations could qualify, because:
  - Most building upgrades could incorporate EE, so
  - Deciding yes or no represents the naturally-occurring rate, and
- In a contractor-driven program
  - Most contractors are not program-certified, so
  - Most upgraders will not hear about the program, so
  - Pool is large
- Except for self-selection



# Future Participants as Comparison Groups

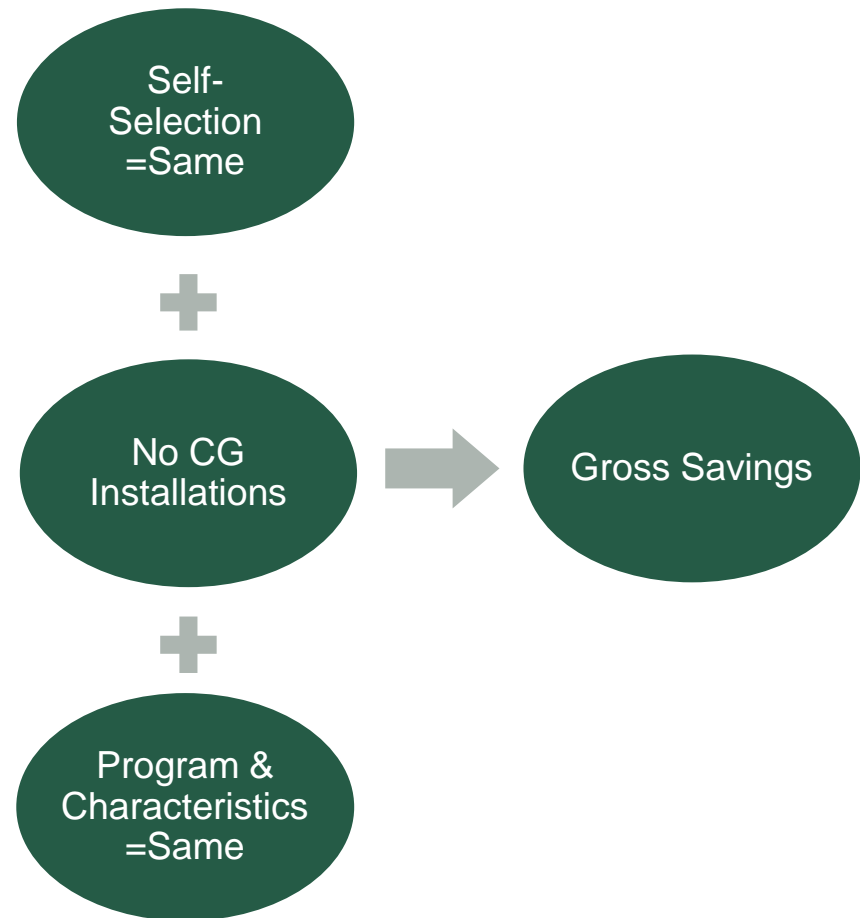
- Future participants have some advantages as a source of CGs

	Program Cycle 1									Program Cycle 2								
Month	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Current Participants	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0
Future Participants	0	0	0	0	0	0	0	0	0	0	0	0	0	X	0	0	0	0

- The biggest advantage is that self-selection is likely to be controlled
- Also should control for exogenous variables like political & economic conditions

# Future Participants as Comparison Groups (Cont)

- Does control of self-selection mean this design produces net impacts? **No.**
- Does it lead to good estimates of gross savings? Not necessarily.



# Future Participants as Comparison Groups (Cont)

- Estimating gross savings requires:
  - Program & Participants unchanged
  - Future participants haven't installed program-qualified measures in their pre-period
- Violations of the last assumption move the estimate toward net
- Violations of the other assumptions means you have nothing and
  - It's essential to know this and make adjustments
- Often, assumptions can be met

# Conclusions

- Identifying a good CG is more complex with whole-building programs
- But CG pool bigger than single-measure programs
- Most non-random CGs will not control sufficiently for free-ridership
  - See Train, Goldberg, and Agnew 2017—next paper
  - They address the non-observable aspects of self-selection/free-ridership
- Future participants are promising pool for CGs
  - But meeting the assumptions is critical
  - Really need to be checked
  - Researchers should be ready to change designs or make other adjustments if assumptions not met

# Katherine Randazzo Principal Data Scientist

krandazzo@opiniondynamics.com  
510.214.0194



***And don't forget my co-authors!***

Rick Ridge, Ridge & Associates;

Stef Wayland, Opinion Dynamics

Brian A. Smith, PG&E