IEPEC 2015
Long Beach, California

Practical Guidance for Selecting Opt-In Research Designs: Addressing Methodological Trade-offs and Avoiding Common Pitfalls

Lucy Morris and Brian A. Smith
Pacific Gas and Electric Company

In Pursuit of the Counterfactual…

"Good experimental design is separable from the use of statistical tests of significance. It is the art of achieving interpretable comparisons…Use of significance tests presumes but does not prove or supply the comparability of the comparison groups or the interpretability of the difference found."

Campbell and Stanley, 1963 (underlining added)

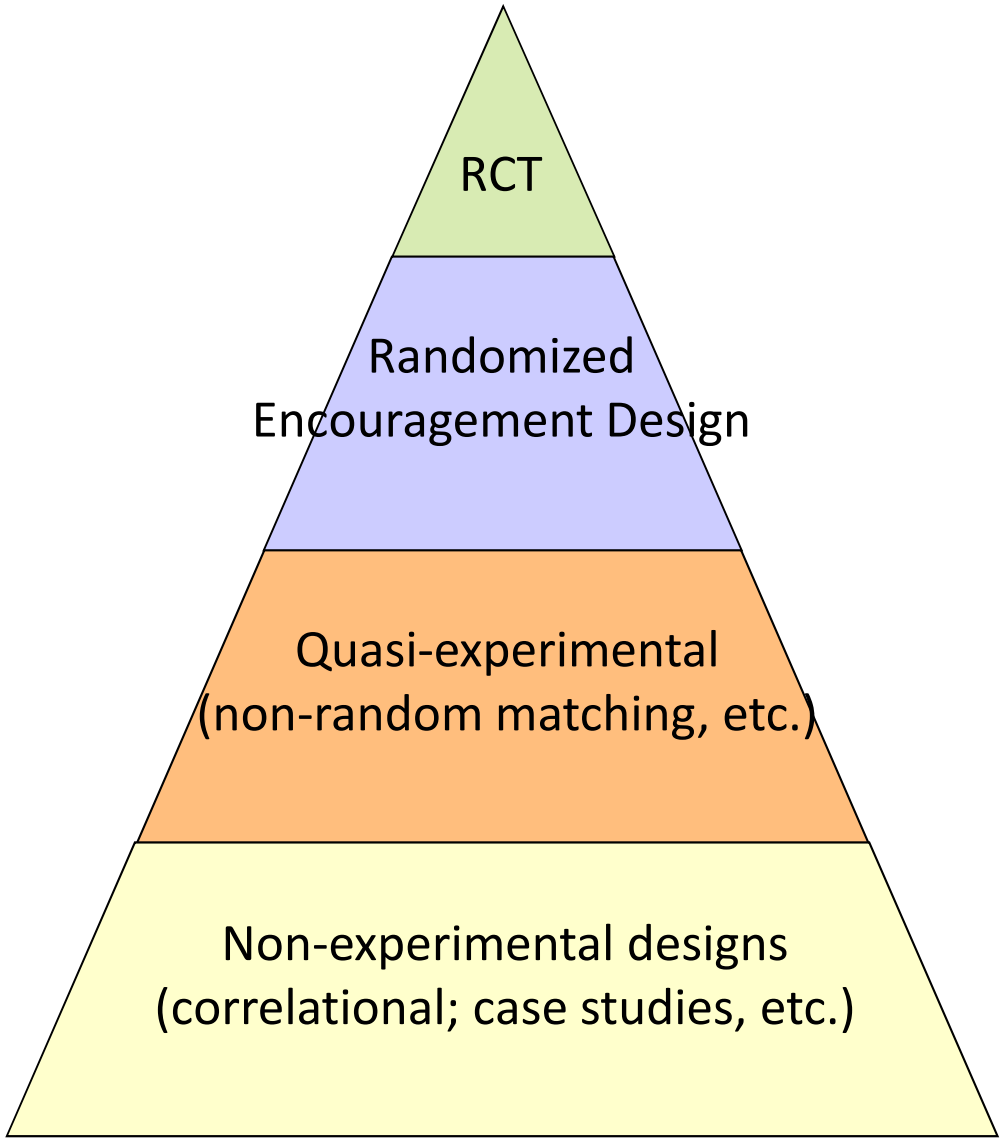
Opt-in EE Behavior Programs = a unique challenge!

- **EE Behavior** = typically small effect sizes that aren't time/day specific
  → Hard to spot in whole home billing data

- **Solutions** = Powerful analysis (i.e. experimental design) and large sample sizes!

But what if a program doesn't lend itself to these solutions??

Ideal Methodology vs. Real World Constraints



Real world considerations

- Budget? Time?
- Expected effect size?
- Rigor needs/requirements?
- Customer experience?
  - Deny/delay treatment
- Data for strategic sampling?

Can I do an RCT?

Yes, if:

- ❑ PA/Utility is okay with denial/delay of treatment to some
- ❑ Can avoid denial/delay by randomizing another factor
- ❑ No equity obligation/requirement
- ❑ Can maintain design integrity

If you can do an RCT, then do it

If you can't, there are other experimental options

- **Randomized Encouragement Design (RED):**
  - "RCT with encouragement" design
  - Randomization precedes customer choice/opt-in
  - Random assignment into "**encouraged**" group or "**unencouraged**" group
  - No treatment denial or delay!

Assumes 3 types of people:

- Never Takers (NT)

No way!

- Always Takers (AT)

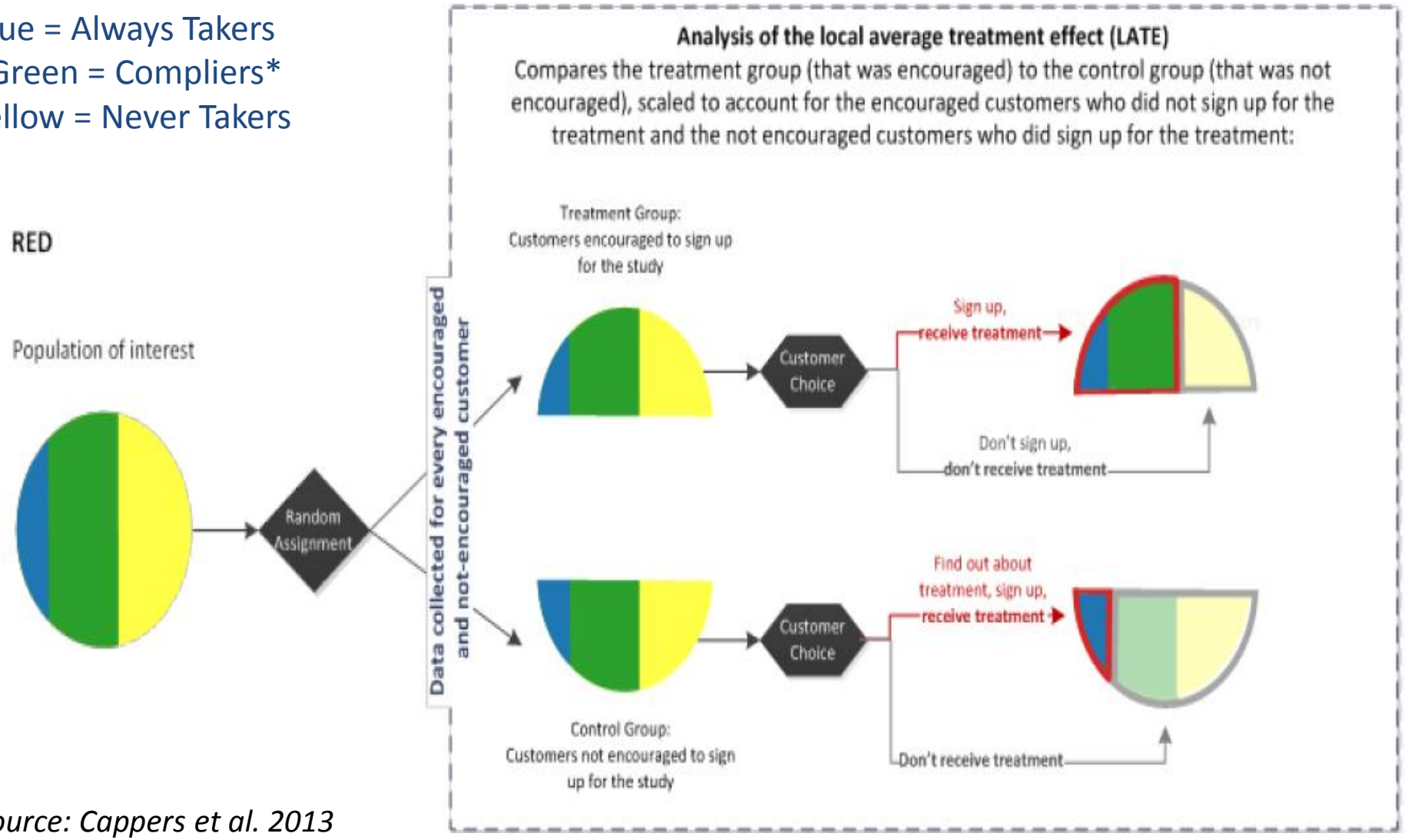
I'll find it and get it!

- Comp

Okay, if you want me to then I'll do it

Randomized Encouragement Design

Blue = Always Takers
Green = Compliers
Yellow = Never Takers



Source: Cappers et al. 2013

RED Considerations

RED challenges:

- Data Availability
- Effective Encouragement
- Marketing control and consistency with program

RED strengths:

- High statistical power of an RCT without denial/delay
- Insight into natural market movement
- Accommodates reality of random assignment integrity
    - Treatment challenges mean an RCT can become an RED

Another option: Quasi-experimental Approaches

Leave this to the experts to weigh the evidence:

- Matching to non-participants on long-term energy usage (ex: Itron, 2013 - sig. 3% of household usage)
  - But is there something "unmatchable" when people opt-in?

- Matching within opt-in participants (ex: Variance-in-Adoption)
  - But a number of requirements make it hard to do this well

Summary and Conclusions

- Evaluation of opt-in behavior programs is challenging: **small effect sizes** & **risk of self-selection bias**
- Need **powerful evaluation designs!**
- If a PA is amenable to recruit-and-deny/delay, then an RCT is the best approach
- If not, assess whether **program design** and **available budget** can **support an RED** with sufficiently disparate uptake to enable a reliable analysis
- If **a true experimental approach is not possible**, use a **quasi-experimental approach** with **two comparison groups**: 1) program participants, and 2)matched non-participants

In all cases, the **best design** will reflect the program's **causal mechanisms, available data, budget, and PA/regulatory tolerance for type I and II errors**

Questions?

Lucy Morris
Measurement & Verification Evaluation Supervisor
Pacific Gas and Electric Company
LLAA@pge.com

Brian Arthur Smith
Expert Strategic Analyst, Energy Efficiency
Pacific Gas and Electric Company
b2sg@pge.com