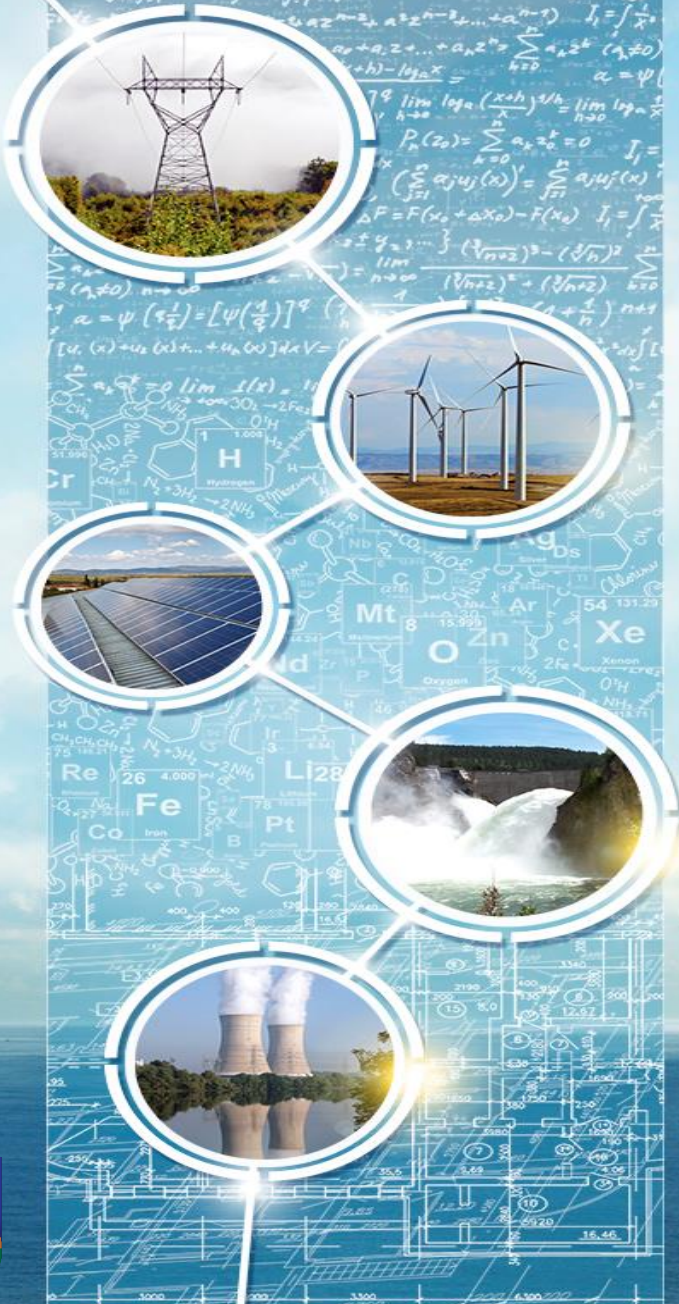**TETRA TECH**

Quasi Experimental Dating: You Matched Me with Who? Opportunities and Limitations of Matching Techniques

2019 IEPEC – Denver, CO

Jonathan Hoechst, Principal Consultant
Ethan Young, Principal Consultant

IEPEC

The Need for Accurate Matching Methods

Identical in all ways but one…

- Control group needed to provide counterfactual argument
    - What would have happened in lieu of the program's treatment?

- Ideally, control group/members will mirror the treatment group/members in all observable and non-observable characteristics
    - Allows evaluators to isolate effect of program

- Problem: evaluators often have only pre-intervention energy consumption data to use for matching control groups

- Yes!
  - Well maybe. Stay tuned.

- When evaluators have additional parameters to use for matching, the matched control group is more likely to accurately reflect the treatment group

- There are likely characteristics and demographics that are correlated with energy usage that are not directly captured by pre-intervention energy use

- Incorporating more of these characteristics should produce more accurate savings estimates
  - Read: not necessarily *higher* savings estimates

Behavioral Program Analysis

- Client: utility in the Southern United States with 900,000 customers

- Program type: behavioral
  - Email, bill inserts, or home energy reports in different waves

- This analysis focuses on customers that began receiving home energy reports via mail in April 2016
  - Treatment group: approximately 40,000 customers
  - Potential control group: approximately 300,000 customers

- Important:
  - Since the program's inception, utility obtained age and income data for all customers

- Using these data, could a better control group be built, and how would that affect savings?

Pre-Intervention Average Daily Energy Use (kWh)

- Unmatched control group v treatment

- Difference in average daily use statistically significant for each month

- Something systemically different (and not accounted for in energy consumption) between treatment and control group

- Result
  - Control group showed lower average daily consumption during every month

- Analysis created four separate specifications for matching:
  - (1) pre-intervention energy use
  - (2) pre-intervention energy use and age of head of household
  - (3) pre-intervention energy use and income of head of household
  - (4) pre-intervention energy use, age, and income of head of household
- For each of the matched control groups produced from 1-4
  - Compare the difference in pre-intervention energy use, by month
    - Ideally this is as small as possible
  - Estimate energy savings using lagged dependent variable model
- How do savings estimates differ when the matched control group more accurately reflects the treatment group?

Matching Results (Average Daily kWh)*

Month	Spec 1: avg. daily energy use		Spec. 2: avg. daily energy use & age		Spec. 3: avg. daily energy use & income		Spec. 4: avg. daily energy use, age & income	
	Control	Treat	Control	Treat	Control	Treat	Control	Treat
May 2015	31.67	31.83	31.70	31.83	31.69	31.83	31.58	31.83
June 2015	41.87	41.92	41.69	41.92	41.80	41.92	41.67	41.92
July 2015	59.21	59.48	59.07	59.48	59.37	59.48	59.25	59.48
Aug 2015	*67.16*	*67.86*	67.40	67.86	67.57	67.86	67.51	67.86
Sept 2015	*53.39*	*54.01*	53.62	54.01	53.70	54.01	53.68	54.01
Oct 2015	*40.65*	*41.11*	40.86	41.11	40.81	41.11	40.79	41.11
Nov 2015	31.08	30.98	30.90	30.98	30.75	30.98	30.75	30.98
Dec 2015	38.48	38.24	38.14	38.24	37.87	38.24	37.88	38.24
Jan 2016	47.39	47.00	46.71	47.00	*46.43*	*47.00*	46.55	47.00
Feb 2016	51.38	51.26	51.03	51.26	*50.60*	*51.26*	50.65	51.26
Mar 2016	39.35	39.34	39.21	39.34	38.95	39.34	*38.86*	*39.34*
Apr 2016	31.74	31.59	31.55	31.59	31.41	31.59	31.32	31.59

*Cells presented in bold italic typeface denote significant differences between treatment and matched control at 95%.

Pre-Intervention Average Daily Energy Use (kWh) Matching on Energy Use & Age

- Adding age into propensity score matching method created accurate control group

- Matching Result:

- No month has statistically significant difference in daily energy usage

- Age differences between treatment and control no longer statistically significant

- Income difference remains statistically significant between groups

Findings: Matching Specifications Effect on Savings

Data used	Trtmnt effect (kWh)	Std. error	t-statistic	P-value	R^2	Adj. R^2
Unmatched	-0.207	0.028	-7.320	< 0.001	0.948	0.948
Spec. 1	-0.306	0.035	-8.797	< 0.001	0.950	0.950
Spec. 2	-0.308	0.035	-8.710	< 0.001	0.950	0.950
Spec. 3	-0.248	0.035	-7.100	< 0.001	0.950	0.950
Spec. 4	-0.243	0.035	-6.911	< 0.001	0.949	0.949

- The matching specification incorporating energy use and age produced highest savings

- But only marginally more than the model using only energy use

- Models using income (Specs. 3 & 4) produced lower savings than energy-only model (Spec. 1)
  - Likely due to collinearity of energy usage and income

- Inclusion of demographic data for matching can produce improved control groups, but effect on savings is varied

- Expand demographic data by geography

- Conduct analysis with different program intervention

- Employ other matching techniques

- Ultimately, evaluators, implementers, and utilities must weigh high cost of obtaining data (both capital costs and cybersecurity risk) against gains in representativeness (and resulting affect on savings estimates)