Introduction

Many utilities and jurisdictions have been testing innovative rate designs and/or technologies over the past two decades

- These innovative/alternative rate designs address one or more deficiencies of the current flat and volumetric rate construct
- Customer response and experience with these rates should be understood before offering these rates to the broader population

Brattle has been maintaining a database of the pricing and technology pilots conducted in the U.S. to track customer responsiveness to these rates, as well as to improve the state of the art for the next generation of pilots
Customer response to time varying rates is well understood

Based on the experimentation over the past decade, we have conclusive evidence that customers respond to time varying rates (TVRs)

Source: Results from 349 pricing experiments in Arcturus database.
Brattle developed tools to estimate the impact of TVRs using data from previous pilots.

We have built a model called “PRISM” to estimate the impacts using price elasticities and also developed an impact curve that approximate the peak impact as a function of peak/off-peak ratio.
Given the existing evidence on customer responsiveness, do we need more pilots?

There are high level takeaways one can glean from other pilots, however it is rarely the case that the deployment and design scenarios are identical across different jurisdictions, For instance:

- Opt-in vs. Opt-out
- RCT vs. randomly selected control group

In the event that you decide to do your own pilot, there are critical questions to answer:

- How would you articulate the objective of the pilot?
- What types of rates should you test in the pilot?
- What is the likely approach to offering these rates to the broader population (i.e. opt-in, opt-out, mandatory)?
- How should you design the experiment given the likely deployment approach?
- Should you also bundle some enabling technologies and information treatments along with the rates?
- Are you interested in measuring impact for sub-populations of interest (i.e. low income, NEM customers, etc.)
How should you proceed with the pilot?

1. Plan to run it **at least for a year** and plan on spending real money on it but no more than the value of information you hope to gain from the pilot.

2. If the objective is to estimate customer behavior to dynamic pricing in addition to understanding customer acceptance, you will need to do an experiment that follows the **scientific principles of pilot design**.

3. Prepare a **comprehensive pilot proposal** that should address the following:
   - Rate design details
   - Pilot design details (i.e. design approach, sample size calculations)
   - Marketing, customer education and recruitment plan
   - Evaluation, measurement and verification plan
   - Budget and cost recovery
   - Pilot timeline

4. Incorporate **stakeholder input** to the pilot proposal.
# Checklist for a Scientifically Valid Pilot Design

1. Clear articulation of pilot objectives
2. Internal validity, meaning a cause and effect relationship can be established between the treatment being tested (the TOU rate) in the pilot and the outcome of interest (change in peak usage)  
   - requires a robust control group and pre-treatment data
3. External validity, meaning that the results from the pilot program can be extrapolated to the population of interest  
   - requires pilot recruitment to mimic potential wide scale deployment; can be ensured by selecting appropriate design approach
4. Determine sampling frame/eligible population for the pilot
5. Undertake “statistical power calculations” to determine minimum size requirement for treatment and control groups to detect statistically significant impacts
6. Incorporate attrition assumptions in the final sample sizes
Scientifically Valid Pilot Design Approaches (and control group strategy)

There are three widely accepted pilot design approaches:

<table>
<thead>
<tr>
<th>Possible Pilot Design Approaches</th>
<th>Description and Pros/Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized Controlled Trial (&quot;RCT&quot;)</td>
<td>Involves a random assignment of the recruited customers into the treatment and control groups. While this is the most rigorous approach from a measurement perspective, it is rarely used by electric utilities due to a potentially adverse impact on customer satisfaction (as it would involve either “recruit-and-denry” or “recruit-and-delay” approaches for some portion of the recruited customers).</td>
</tr>
<tr>
<td>Randomized Encouragement Design (&quot;RED&quot;)</td>
<td>Allows the researcher to construct a valid control group, maintaining the benefits of an RCT design by not negatively affecting the customer experience. However, it requires much larger sample sizes, relative to the RCT approach, in order to be able to detect a statistically significant impact. Large sample sizes increase pilot implementation costs.</td>
</tr>
<tr>
<td>Random Sampling with Matched Control Group</td>
<td>Involves recruiting treated customers from a randomly selected sample, and using regression analysis to identify and match customers from the rest of the population that are most similar to the treatment customers. This matched control group approach strikes a good balance between achieving statistically valid results and requiring a manageable level of pilot participants.</td>
</tr>
</tbody>
</table>

A statistically valid pilot design yields comparable treatment and control groups. This is an essential requirement in order to be able to attribute the difference between the two groups to the treatment impact.

Note: The shaded regions indicate peak hours. Control group was constructed using a matching analysis.
While RCT and RED are the most rigorous design options, implementation considerations may call for matched control groups

**Propensity score matching** is a widely-used statistical matching method in economics and other social sciences

- Uses statistical analysis to identify the variables that are most closely correlated with enrollment in the pilot
- Using the results of that analysis, “predicts” the probability of participation for both enrollees and control group
- Identifies, for each enrollee, a control group member who is “most similar” with respect to the observed covariates
- The ultimate goal is “covariate balance” – we want the control group averages to be as close as possible to the pilot group averages, particularly on the variables that “matter” the most
- Achieving “perfect” balance is rare, but this approach is usually successful, on net, in generating a control group that “looks like” the “treatment” group
Treatment vs. Control groups (Before Matching)

Average Load Profile by Customer, Unmatched
Treatment vs. Control groups (After Matching)

Average Load Profile by Customer, Matched
In order to determine the pilot’s impact in a statistically significant fashion, the sample size should be large enough. There are several parameters that affect the sample size.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description of parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group means</td>
<td>Average amount of electricity consumed by each group</td>
</tr>
<tr>
<td>Standard deviations</td>
<td>Amount that electricity consumption varies across households within each group</td>
</tr>
<tr>
<td>Number of repeat observations</td>
<td>Number of observations per household</td>
</tr>
<tr>
<td>Correlation coefficients</td>
<td>Degree to which electricity consumption is similar over time for a given household in the treatment and/or control group(s)</td>
</tr>
<tr>
<td>Statistical significance</td>
<td>Degree of certainty that the program reduces usage [one-sided test]</td>
</tr>
<tr>
<td>Statistical power</td>
<td>Degree of certainty that the statistical test gives the correct answer</td>
</tr>
</tbody>
</table>
Statistical power calculations are necessary to determine the sample size.

- As the minimum detectable impact (MDI) increases (i.e. due to higher peak to offpeak ratio), sample size requirement decreases.
- As the statistical power and statistical significance requirements increase, the sample size increases.
- As the resolution of the analysis increases (i.e. hourly vs. monthly), sample size requirement decreases.
Pilot Design Approaches Used in Other Pilots

Early pilots typically relied on random sampling with voluntary participation + randomly selected control groups

– California Statewide Pricing Pilot, 2003; Baltimore Gas and Electric Smart Energy Pricing Pilot, 2007)

Some of the more recent pilots used RCT and RED

– SMUD SmartPricing Pilot, 2014; Ontario RPP Pilots, 2018

However, practical considerations (i.e., denying participation to the recruited customers in the RCT or large sample size requirements of RED) were not surmountable for other recent pilots. These pilots opted to use random sampling with matched control group

Checklist for the Recruitment Process

Practices followed in the recruitment process play a key role in maintaining the validity of the pilot design and offer important insights for broader deployments

- Follow best practices in developing customer education and outreach materials (including samples of effective vs. ineffective marketing materials)
- Consider different recruitment strategies through different channels based on the type of treatment offered and recruitment for special interest groups
- Identify approaches to minimize marketing costs while maximizing the number of recruited customers;
- Develop strategies to improve retention rates
- Be aware of correct and incorrect ways to introduce incentives to the recruitment process
- Incorporate new information that becomes available during the recruitment process to improve the success of recruitment
- Provide robust training to the marketing team to ensure that they don’t inadvertently compromise
- Design pre- and post-treatment customer experience surveys aligned with pilot objectives
Common Mistakes during Recruitment

- Recruitment team deviating from the pilot design plan to meet the sample size targets
- Nonexistent or infrequent communication between the recruitment and design teams that might introduce inefficiencies to the overall pilot management
  - Loss of marketing cost savings
  - Loss of valuable course correction opportunities
- Misuse of incentive payments
- Recruitment starting around the holiday times
- Recruitment process that necessitates too many touch points with the customers before sign up
- Not capturing useful customer interactions/communications that might inform future program deployment strategies
Bill impact analyses are useful to understand the distributional impacts of the new rates.

It is useful to undertake a bill impact analysis of the eligible customer population under the new rates.

The bills would be calculated twice: before load response and after load response.

Some utilities/jurisdictions choose to include this information for individual customers in the recruitment materials.

Under this rate, 51% of customers experience lower bills without DR compared to 86% with demand response.
BGE, Pepco and DPL ("JUs") are implementing the pilots

- For all three utilities, the TOU rate is applied to both energy and delivery charges
- Two year pilot which started in June 1, 2019.
- The peak to off-peak ratio is very pronounced and varies from ~5-to-1 to 6-to-1 across the JUs
- The peak hours vary by season
  - HE 15-19 on summer weekdays (June 1 – September 30)
  - HE 7-9 on winter weekdays (October 1 – May 31)
- The treatment customers also get behavioral messaging to reinforce the pricing signal
- The pilots were designed to allow impacts to differ between low- and medium-income ("LMI") and non-LMI customers
- Interim impact evaluation (using Summer 2019 data) yielded promising results; first year analysis will be completed in the summer of 2020
Checklist for the Impact Evaluation

The experimental design of each pilot dictates the optimal evaluation method: differences-in-differences (ANOVA or ANCOVA); panel regressions (fixed-effects or random-effects); individual customer regressions

- Decide on the **evaluation approach** based on the experimental design
- Identify **load impact metrics** to be quantified (i.e. peak, mid-peak, off-peak impacts, average daily conservation impact, etc.)
- Estimate alternative models and select the one that leads to most accurate predictions
- Decide whether quantifying customers’ overall price responsiveness would be useful in the form of **price elasticities**, beyond the ex-post load impacts quantified in the pilot
  - **Own/daily price elasticity** (captures the change in the level of overall consumption due to the changes in the average daily price)
  - **Substitution price elasticity** (captures customer’s ability to substitute inexpensive off-peak consumption for more expensive peak consumption)
Is price elasticity estimation necessary?

Most pilot studies test a single price level for a given rate design

- As a result, impact evaluation quantifies the impact associated with that particular rate

If the Company is likely to offer other rate designs, or different price levels for the same rate design, it is very useful to estimate the own price and substitution price elasticities

Estimating elasticities using the pilot data allow computation of the load impacts from other rate designs, and have the benefit of reflecting utility’s own customers’ price responsiveness
Checklist for the Process Evaluation

A process evaluation consists of an assessment of the implementation of the program, with the goal of producing better and more cost-effective programs in the future

- Typically be conducted by surveying or soliciting feedback from the various groups involved in the pilot program, including both participants, implementers and administrators of the program

Data collection efforts include but are not limited to:

- Customer recruitment and outreach (pre-treatment survey)
- Customer acceptance and interest in treatment (post-treatment survey)
- Understanding the reasons for non-participation and attrition
- Quality control practices
- Time, schedule and budget management
- Lessons learned
- Project resource constraints and staff training
- In-field and back-office challenges with implementation
Recap I

**Upfront investment in pilot planning** is absolutely critical for the success of the pilot

- Well-developed EM&V plans, customer education and recruitment plans increase the likely success of the pilot

**Seeking stakeholder input** during the pilot design process and incorporating this input to the design increase the acceptability of the pilot results

**Resist designing overly complex pilots** that could easily interfere with meeting the essential objectives of the pilot

It is advisable to test treatments and functionality **only if they are likely to be offered in full scale deployments** (i.e., bill impacts, shadow bills, etc.)
Avoid siloing the pilot design and marketing teams during the recruitment stage, as deviations from the recruitment plan may compromise the validity of the pilot.

Estimation of price elasticities is desirable as part of an impact evaluation study to allow estimation of the impacts from alternative rates.

It is important to calculate sample sizes consistent with the pilot design approach that will yield statistically significant results.

An interim impact evaluation after the first season of the pilot is useful to gauge initial results and allow course-correction if needed.
References I

Sanem Sergici, Ahmad Faruqui, and Nicholas Powers, EM&V Plan for the PC44 TOU Rate Pilots, prepared for the PC44 Rate Design Workgroup, June 2018.


Sanem Sergici and Ahmad Faruqui, Measurement and Verification Principles for Behavior-Based Efficiency Programs, prepared for Opower, May 2011.

References II


Dr. Sanem Sergici is a Principal in The Brattle Group’s Boston, MA office specializing in program design, evaluation, and big data analytics in the areas of energy efficiency, demand response, smart grid and innovative pricing. She regularly supports electric utilities, regulators, law firms, and technology firms in their strategic and regulatory questions related to retail rate design and grid modernization investments.

Dr. Sergici has been at the forefront of the design and impact analysis of innovative retail pricing, enabling technology, and behavior-based energy efficiency pilots and programs in many states and regions including District of Columbia, Connecticut, Florida, Illinois, Maryland, Michigan, Ontario, CA and New Zealand. She has led numerous studies in these areas that were instrumental in regulatory approvals of Advanced Metering Infrastructure (AMI) investments and smart rate offerings for electricity customers. She has significant expertise in resource planning; economic analysis of distributed energy resources (DERs); their impact on the distribution system operations and assessment of emerging utility business models and regulatory frameworks.

Dr. Sergici is a frequent presenter on these matters and regularly publishes in academic and industry journals. She was recently featured in Public Utility Fortnightly Magazine’s “Fortnightly Under 40 2019” list. She received her Ph.D. in Applied Economics from Northeastern University in the fields of applied econometrics and industrial organization. She received her M.A. in Economics from Northeastern University, and B.S. in Economics from Middle East Technical University (METU), Ankara, Turkey.

The views expressed in this presentation are strictly those of the presenter(s) and do not necessarily state or reflect the views of The Brattle Group, Inc. or its clients.