STATE OF NEW YORK

PUBLIC SERVICE COMMISSION

Proceeding on Motion of the Commission as to the Rates, Charges, Rules and Regulations of Consolidated Edison Company of New York, Inc. for Electric Service.	CASE 16-E-0060

DIRECT TESTIMONY
OF
ELISHEBA SPILLER

Dated: May 27, 2016

New York, New York

Environmental Defense Fund
257 Park Avenue South, 17th Floor
New York, NY 10010

I. Introduction and Qualifications

Q. Please state your name, title, and business address.

A. My name is Elisheba "Beia" Spiller, and my title is Senior Economist. My business address is Environmental Defense Fund, 257 Park Avenue South, New York, New York.

Q. Would you please summarize your educational background, professional experience, and current general responsibilities?

A. I have almost a decade of experience in researching energy economic topics ranging from gasoline and vehicle demand to impacts of natural gas extraction. However, over the past three years, I have primarily focused on electricity pricing and regulation. I received a Bachelor of Arts in Environmental Policy from University of California, San Diego in 2003; a Master's in Public Policy from Universidad Torcuato di Tella in Buenos Aires, Argentina in 2005; a Master's in Economics from Duke University in 2008; and a PhD in Economics from Duke University in 2011. During my tenure at Environmental Defense Fund ("EDF"), I have engaged in the topic of electricity pricing through multiple efforts. For example, I have participated in rate cases in both New York and California, helping to provide analysis for testimony and comments describing the benefits and impacts of time-variant pricing. In March 2015, I led a forum on time-variant pricing in New York City developed in coordination with Department of Public Service Staff and the Institute for Policy Integrity at NYU School of Law ("IPI") to inform the New York Public Service Commission (the "Commission")'s Reforming the Energy Vision ("REV") proceeding, inviting a number of practitioners from across the country to describe their experiences with time-variant pricing and how it could be implemented in New York. Throughout this time, I have followed many studies that demonstrate the effects and outcomes of time-variant pricing and other types of advanced pricing mechanisms, and, jointly with IPI, developed a conceptual proposal for a time-variant pricing pilot for Consolidated Edison Company of New York, Inc. ("Con Edison" or the "Company"). Although the Company did not act upon that proposal at that time, citing funding constraints and lack of advanced metering infrastructure ("AMI"), our understanding was that the Company planned to pursue a study at a later time, when AMI was available.

Q. Are you sponsoring any exhibits to your testimony?

A. Yes. I offer Exhibit__ (ES-1), which consists of the Company's response to EDF Interrogatory Request ("IR") 14; Exhibit ___ (ES-2), which consists of U.S. Department of Energy Smart Grid Investment Grant Technical Advisory Group Guidance Document #12; Exhibit __ (ES-3), which consists of U.S. Department of Energy Technical Advisory Group Guidance Document #7; and Exhibit ___ (ES-4), which consists of the SmartPricing Options Final Evaluation developed by Sacramento Municipal Utility District ("SMUD").

II. Overview

Q. What is the purpose of your testimony?

A. The purpose of my testimony is to enhance the record as it relates to time-variant pricing pilots or demonstration projects that may be approved in connection with this rate case or simultaneously with this rate case in a parallel proceeding. The need for such pilots or demonstration projects and the Company's intentions with respect to them is raised in the Company's testimony. For example, the Policy Panel states that "Phase Three" of the CONnectED Homes Platform project (the "CONnectED Homes Demo Project") "has, as a feature, the evaluation of alternative rate design. Specifically, the Company will examine customer responses to alternative rate designs, such as TOU rates, technology, and tools focused on education and awareness." (Policy Panel Testimony at page 37, lines 8-12.) The expectation that time-variant pricing pilots will be undertaken by the Company in the near future is further reinforced by provisions of the Commission order (Case 13-E-0030, *Proceeding on Motion of the Commission as to the Rates, Charges, Rules and Regulations of Consolidated Edison Company of New York, Inc. for Electric Service*, Case 13-G-0031, *Proceeding on Motion of the Commission as to the Rates, Charges, Rules and Regulations of Consolidated Edison Company of New York, Inc. for Gas Service*, and Case 15-E-0050, *Proceeding on Motion of the Commission as to the Rates, Charges, Rules and Regulations of Consolidated Edison Company of New York, Inc. for Electric Service supra*, Order Approving Advanced Metering Infrastructure Business Plan Subject To Conditions (March 17, 2016) (the "AMI Order")) concerning the Company's AMI business plan (Case 13-E-0030, *supra*, Case 13-G-0031, *supra*, and Case 15-E-0050, *supra*, Con Edison Advanced Metering Infrastructure Business Plan (Oct. 16, 2015) (the "AMI Business Plan")). The AMI Order states that the customer engagement plan, which is to be filed on July 29, 2016, "is to include innovative pricing proposals, which should include one or more pilot programs, developed in consultation with stakeholders." (AMI Order at page 38.)

More recently, based on the Company's response to an interrogatory request submitted by EDF, we understand that the Company believes that the AMI Order's requirement for innovative pricing pilot proposal "supersedes previous plans to test alternative rate designs in the demonstration projects." The Company stated in its IR response that "Based upon the outcome of that effort, the Company will evaluate the need to examine customer responses to alternative rate designs, such as TOU rates, technology, and tools focused on education and awareness in Phase 3 of the Connected Homes demonstration project." (Exhibit ___(ES-1)).

Even more recently, the Commission's Order Adopting a Ratemaking and Utility Revenue Model Policy Framework, issued on May 19, 2016 in the REV proceeding, requires the Company to file one or more Smart Home Rate ("SHR") demonstration proposals by February 1, 2017. (Case 14-M-0101, *Proceeding on Motion of the Commission in Regard to Reforming the Energy Vision*, Order Adopting a Ratemaking and Utility Revenue Model Policy Framework (May 19, 2016) (the "Track Two Order") at page 157.) As the Track Two Order explains, "[a]n SHR combines time-variable rates with the full value LMP+D compensation that is being developed in Case 15-E-0751." (Track Two Order at page 135.) The Track Two Order also established an expectation that the Company would examine its existing time-of-use rates with reference to the design characteristics and practices used by utilities with substantially higher customer adoption and include in its next filing (or file on June 1, 2017, if the Company has a rate plan that expires after January 1, 2018) a proposal to revise its voluntary time-of-use rates for mass market customers, including an analysis of how the proposed rate compares with rates in other jurisdictions. (*See* Track Two Order at pages 134, 155-156.)

As a result of all this, it is clear that the Company should be expected to undertake some number of pilots and/or demonstration projects addressing time-variant pricing in the near future. However, as expectations for each of these pilots and/or demonstration projects have not yet been fully developed, it is unclear whether the Commission will be able to ensure that the studies undertaken are optimally designed to advance the public interests that time-variant pricing can serve.

The lack of detail about any such proposed pilots or demonstration projects contemplated in the testimony, the AMI Order, and the Track Two Order presents significant risks, which are compounded by the cloudy procedural outlook for developing such pilots or demonstration projects. Specifically, there is a material risk that pilots or demonstration projects will be undertaken by the Company in the near future (pursuant to the rate order in this case or otherwise) which cannot be reasonably expected to achieve what the Commission and Con Edison's ratepayers should demand from such pilots.

Q. Can you please provide a brief summary of your testimony?

A. My testimony addresses how to design a pilot or a demonstration project in a way that would allow the Company to accurately estimate the impacts of one or several alternative pricing mechanisms. Much of the thought behind this discussion (in particular, the opt-in Randomized Control Trial which is discussed in length below) emerged from the development of the demonstration project concept that EDF and IPI developed for Con Edison in early 2015. The benefit of conducting a carefully designed experiment is that it can provide sufficient information to inform future rate design that would support "the achievement of environmentally and economically efficiency electric power through self-sustaining markets and market regulation," which is "the ultimate goal of REV." (Case 14-M-0101, *supra*, Order Adopting Distributed System Implementation plan Guidance (April 20, 2016) at page 23.)

III. Time-Variant Pricing Pilot Design

Q. What objectives should the Company seek to achieve through time-variant pricing pilot projects?

A. The Company's Policy Panel states, with respect to the evaluation of "alternative rate designs" that it contemplates as part of phase three of the CONnectED Homes Demo Project, that "[t]hrough the knowledge and insights gained as this demonstration is executed, the Company will have a better understanding of customer motivations and response to their energy usage profiles. This will translate into an ability to design informed targeted campaigns and effective engagement through the established HER campaigns and customer portal." (Policy Panel Testimony at page 37.)

In light of the Commission's focus on time-variant pricing as a tool for accomplishing key REV goals, the goals of studying alternative rate designs can and should be articulated with greater ambition and greater specificity. We would suggest that the following three objectives be incorporated in the design of any time-variant pricing pilot or demonstration project:

1. To determine how the consumption and pattern of energy use is influenced by varying rate designs and accompanying enabling technologies. Time-variant

pricing pilots should study the effectiveness of time-variant price signals on the level of overall energy consumption, on peak demand reduction, and on shifting the energy usage from peak periods to off-peak periods.

2. To determine the extent of environmental benefits, such as the reduction in carbon dioxide emissions and other air pollution, that can be achieved by the use of such time variant prices and enabling technologies.
3. To evaluate the potential economic benefits of implementing time variant prices.

Q. What purposes should alternative rate designs serve and how can pilots and/or demonstration projects be developed to serve those purposes?

A. Currently, most electricity customers face a very simple volumetric rate that does not vary either by time of customer demand or customer location. Historically, these flat volumetric rates were a valid approximation in an era of expanding demand and growing electrification. In today's world, where customers are participating more than ever in the electricity market and are becoming much more actively engaged in the dynamic operation of the electric system –through generating their own electricity, shifting towards efficient purchases, participating in demand response programs, and so on – these flat volumetric rates are an artifact that neither reflects cost causation nor incentivizes the right type of customer behavior which could lead to a more vibrant electricity market and reduced long run costs. In the recent Track Two Order, the Commission recognized the need to rethink the most basic building blocks of ratemaking – for example, modifying the principle of cost causation to include embedded costs as well as long-run marginal and future costs (while specifically clarifying that fixed charges should only be used to recover costs that do not vary with demand or energy usage). (Track Two Order at page 122 and Appendix A.) Consistent with its revised and updated rate design principles, the Commission, in the Track Two Order, adopted "the policy direction that more granular rate design must be made available to engage customers efficiently in multi-sided DER markets. This policy direction is contingent on the availability of market opportunities and enabling technologies for customers to respond to price and value signals. In addition, improvements in rate design must be coordinated with a plan to reduce the overall energy burden on low and middle-income households…." (Track Two Order at page 123.)

Understanding how to implement rates in order to achieve a particular outcome and avoid certain other outcomes is key. The Commission noted this in the Track Two Order, stating that "[w]hile our policy direction is clear, further demonstrations and analyses of bill impacts are needed before generalized demand charges, or default time of use rates, are adopted for mass-market customers." (Track Two Order at page 123.) Essentially, demonstration projects and pilots can help the Company better understand how to construct an alternative pricing tariff in order to achieve these outcomes.

Q. Would time-variant pricing pilots and demonstration projects in Con Edison's service territory provide different insights compared with time-variant pricing pilots and demonstration pilots that have been undertaken elsewhere in the past?

A. Though many time-variant pricing pilots have been implemented across the country, there are several reasons why piloting time-variant pricing in New York City is so important. First, New York's climate is different from the climate in the locations where some of the most significant time-variant pricing pilots have been undertaken – which means that the heating and cooling needs of all New Yorkers, urban and non-urban, and their willingness to change their consumption of those services based on price signals, may be distinct from what has been found in studies of customers in, for example, California. Second, New York City has a unique urban fabric and residential housing stock – which means that customers in New York City may not react exactly the same way to a rate as customers in other cities or to non-urban customers even in the same climate zone. Third, many of the most well-designed pilots have been implemented in states with vertically integrated utilities (*see* the discussion about Sacramento Municipal District's pricing pilot later in this testimony); New York State has a significantly different regulatory construct from those jurisdictions, and now that REV has begun to change what is expected of utilities, New York's regulatory construct is somewhat different from those in *any* other jurisdictions. Thus, in New York, the emerging regulatory construct appears to include somewhat familiar elements (utilities own and operate the distribution grid but generally not generation), less familiar elements that have been coming into focus over the past two years (utilities act as distributed system platform providers, a role that obligates them to optimize how portfolios of distributed energy resources are dispatched), and quite novel elements that have only just been introduced (utilities can expect opportunities to increase their earnings by contributing to decarbonization). Against that novel backdrop, the Company could test out various price structures that recover costs in somewhat different manners and send price signals that can be expected to yield market outcomes that will affect different parts of the system in different ways – e.g., rates designed to avoid the need for distribution upgrades, and/or to minimize supply costs, and/or to improve efficiency of dispatching renewables.

Q. Would it make sense, in a well-designed study of alternative price structures, to test the effectiveness of a single treatment or intervention (that is, a single pricing structure and/or customer-enabling technology)?

A. As a general matter, no. It is important to have multiple treatments in a single study.

Q. Why is that?

A. Having multiple treatments or interventions allows the utility to answer a slew of questions that otherwise they would be unable to ask. At the most basic level, having only one treatment means that the utility can only answer the following question: "What is the effect on consumption/load shape from this particular rate?" That question, for a number of reasons, is very limited in its ability to help the utility ultimately develop a tariff that can be marketed to customers across the service territory. First, a different type of pricing structure other than the one that was tested may have been more effective at reducing costs, but if the pilot or demonstration project only tests one pricing structure, then this will never be known. Furthermore, given the amount of heterogeneity across customers and locations, it would be important to test out different types of pricing structures to understand the effects across these various groups/locations.

Fundamentally, the more treatments a pilot or demonstration project includes, the more questions it can answer. For example, imagine a pilot or demonstration project that tests out a particular pricing mechanism on two different groups: a group of individuals who participate in a demand response program, and another group that does not. With these two treatments, the utility can measure whether the pricing mechanism under consideration is more or less effective at reducing peak demand when the customer who is subject to that pricing mechanism also participates in a demand response program. If there is an additional treatment group that includes customers who only face the demand response program (but whose pricing mechanism is conventional, i.e., the flat volumetric pricing that customers outside the pilot generally pay), then the utility can test the relative cost-effectiveness of an advanced pricing structure compared to an automated control program for achieving a specified level of reduced peak demand.

Similarly, if the pilot includes one or more technology treatments, the utility can test the extent to which particular technologies enhance the customer's ability to respond to pricing, and can also test the cost-effectiveness of these technologies at reducing peak demand when paired with pricing. The more treatments in the pilot, the more questions that can be answered; this will help lead to a more effective deployment of advanced pricing systems.

While the list that follows is not exhaustive, a well-developed pricing pilot or demonstration project with multiple treatments could potentially answer the following research questions, which are crucial ones to answer in order to achieve REV objectives:

1. What effect does each pricing structure under consideration have on consumption, load shapes, and peak demand?

2. What is the marginal effect of various pricing structures when implemented jointly with demand response programs?

3. In the absence or presence of any enabling technology, which pricing structure has the greatest effect on local peak reductions, whole-system peak reductions, and conservation?

4. Which results in more cost-effective reductions in critical peaks: time-variant pricing with technology or demand response?

5. What is the impact of information on behavior, either when combined with time-variant pricing or alone?

6. What is the effect of a particular rate on demand for customers with distributed generation?

7. What is the effect of a particular rate on the adoption of distributed energy resources ("DERs")?

Q. How does a pilot test customer responsiveness to price signals?

A. Interpreting customers' responsiveness to the treatment (i.e., price signals) in a pilot requires the analyst to quantify the unobservable counterfactual: What would have been the customer's demand had he not been subjected to the treatment? Because it is impossible to observe the counterfactual at the individual level (i.e., a participant cannot simultaneously be exposed to time variant prices and not be exposed to time-variant prices), the experiment should be set up in a way that another group of individuals not exposed to the alternative price mechanism act as proxies for the counterfactual to the outcomes of individuals in the treated group. Thus, it is important that there be a control group (i.e., a group of customers not exposed to the treatment of an alternative pricing structure) made up of customers who are as similar as possible to those in the treatment group (i.e., the customers who are exposed to the treatment). "Similarity" has two aspects: observable similarity (e.g., similar housing type, same types of appliances, similar demographics, etc.) and unobservable similarities (e.g., similar preferences and willingness to change behavior). Both of these aspects are important, although the second is much more difficult to identify. If the control group is not similar to the treatment group, this will cause the estimated results to be biased. This means that, for example, the estimated responsiveness of customers to price signals could be over- or understated, leading to an incorrect understanding of the effects of a certain price structure.

Q. Which experimental designs should be implemented in order to ensure that the control groups are valid and therefore that the resulting outcomes are correctly estimated?

A. There are two highly credible experimental designs, that generally allow the experimenter to produce valid and unbiased estimates of the effects of the treatment (i.e., advanced pricing structure) on the relevant outcomes (i.e., peak demand reductions, load shape changes, etc.). The first, the "gold standard" in science and economics, is a Randomized Control Trial ("RCT"), and the second, which can closely approximate a RCT, is a Randomized Encouragement Design ("RED"). Both of these experimental designs would allow the Company to ensure that the results of the pilot can be extended to a broader setting: one in which the tariff is a mass-market rate available to all customers.

However, the utility's decision of which of these two experimental designs to choose from will depend on the following: 1) whether the utility is seeking to test opt-in, opt-out, or mandatory rates; 2) whether the utility would ultimately want to provide a menu of tariffs to customers from which they can choose; and 3) the recruitment budget.

Below, I discuss both of these two designs in detail, and how the aspects as described above would affect the choice of design. A further description of these two experimental designs can be found in two guidance documents that the United States Department of Energy developed for utilities participating in the Smart Grid Investment Grant; these documents describe how to implement RED (*see* Exhibit ___ (ES-3)), and how to measure the effects of the treatments under both RED and RCT (*see* Exhibit ___ (ES-2)).

Q. **How does a RED work in the context of a pricing pilot/demonstration project and under what circumstances would this design be the optimal choice?**

A. A RED experiment works in the following way: customers are chosen at random from the population and encouraged to participate in an experiment or pilot. The key difference between a RED and a RCT is that, in a RED, "treatment" is induced through encouraging the customer to participate in the treatment group, rather than being assigned to it. The customer can either accept or refuse, but either way, that customer was exposed to the treatment by being encouraged to participate. This implies that the control group is made up of those customers who were not encouraged to participate in the pilot. If the encouraged group is a randomly selected sub-group of the overall population, the control group is also a randomly selected sub-group of the overall population who were not encouraged to participate in the pilot. The treatment group in a RED experiment therefore consists of the randomly chosen group of people who received encouragement to accept an alternative price structure, regardless of whether they accepted that price structure or not, and the control group consists of a group of people selected at random in the same manner as the treatment group (*see* Exhibit ___ (ES-3) at page 44 for a graphical representation of the treatment and control groups). This allows the utility to accurately measure the impact of the treatment on customers who are encouraged to adopt a particular rate; an adjustment is made at a later stage to tease out the effect of the treatment from the effect of the alternative pricing mechanism (*see* Exhibit ___ (ES-3) at pages 45-47 for a discussion of the statistical validity of this method, and how to estimate outcomes post-pilot). This design is a good choice if the utility wants to test the impact of one opt-in rate, where customers are actively pursued and encouraged to opt in to that particular rate.

If all customers who are encouraged to participate in the pilot accept the pilot price offering, then RED and RCT experiments are identical, and causal estimates of the average treatment effect ("ATE") may be derived by comparing the difference in mean outcomes between the treatment and control groups.

The limitations of a RED can arise when some portion of the encouraged participants refuse to adopt the pilot price offering. In these cases, comparing outcomes between those who accept and those who refuse the alternative pricing mechanism could yield biased estimates, if the decision to accept the price mechanism is associated with some unobservable characteristic that drives outcomes. That is, if compliance with the encouragement is imperfect (i.e., where a portion of those encouraged choose not to accept the new pricing mechanism), the analyst must take into account how acceptance of the pricing option among the encouraged individuals could affect the results. Fortunately, RED offers an elegant solution to this potential selection issue. Instead of comparing the outcomes of those who accept the pricing offering and those who do not, the analyst compares the outcomes of the whole treated group (i.e., those who were encouraged) with the control group, regardless of whether all of the encouraged participants accept the pricing offering. Because the analyst is comparing randomly selected groups, selection bias is not a concern, though the interpretation of the results is somewhat changed when some members of the treatment group decline the price offering. Instead of estimating the ATE, a RED experiment with imperfect compliance yields an estimate of the average intent to treat ("ITT") effect.

Essentially, a RED experiment with imperfect compliance would yield an estimate of the average effect of encouraging a group to adopt a particular tariff. In many cases (such as if, for example, regulations require that individuals be allowed to opt out of a particular rate structure), the ITT will be more relevant than the ATE. In any case, the ITT can be scaled up to the Local Average Treatment Effect ("LATE") – which is the impact of the intervention on those who were encouraged to participate – by dividing the ITT by the percentage of customers who accepted the pricing intervention (*see* Exhibit __ (ES_3) at page 29 for further discussion).

A properly designed RED experiment would ensure that customers are only encouraged to adopt one particular rate. Rather than encouraging the customers to participate in the pilot and then offer them a menu of options, the customer is instead encouraged to participate in only one option. So, if the pilot is testing out three different price mechanisms, a customer would receive a letter encouraging her to participate in one of these three options. The pilot participant would not be offered either of the other price options.

If the utility were to allow prospective participants to choose from a menu of different rate options, customers would sort into one rate or another based on their preferences, causing significant differences across the groups, resulting in selection bias. RED helps eliminate this selection bias by encouraging customers to opt in to a particular treatment, without allowing for them to choose the rate that they believe will be the most beneficial one for them out of a menu of options.

Finally, RED allows for a utility to test opt-in vs opt-out scenarios by allowing the encouragement to be either opt-in or opt-out, to the extent feasible under applicable laws and regulations. Under an opt-in scenario, as described above, the customer is encouraged to participate in a treatment and has the option of saying yes. Alternatively, under an opt-out scenario, a randomly selected group of customers would be defaulted into the pilot and would be encouraged to not opt-out; in such a scenario, they would still be allowed to actively choose the default flat volumetric tariff to which they are currently exposed, but in such an event, they would still be considered a part of the treatment group. As mentioned earlier, a corresponding control group of customers would be chosen at random as well from the non-encouraged population. In such an opt-out experimental design, encouragement consists of defaulting customers onto a rate; the results of that experiment would show how customers would react to being defaulted onto a rate. This experimental design is a good choice if the utility would like to implement a particular tariff in an opt-out or default manner after the pilot is done.

One drawback of a RED pilot is that the utility will not be able to measure the impact of providing a menu of alternatives to customers, where they are allowed to opt in to whichever rate they feel is best for them. Having a menu of options can be very beneficial in terms of ensuring that a large amount of customers will adopt time-variant rates. For example, Arizona Public Service provides their customers with the option of choosing from five residential rate plans that are time-variant. This, coupled with extensive outreach and education, has helped them reach a record amount of customer opt in to these time variant rates (at times, this percentage has exceeded 50%; *see* Faruqui, Ahmad. The Principles and Practices of Time-Variant Pricing. Presented to CPUC, 2014). Given the merits of providing customers with options when it comes to rates, it is important to understand what will be the effect (on peak demand, or any other relevant outcome) of allowing customers to choose which rate is best for them (without encouraging them to adopt a particular one). The RED experiment would likely not estimate correctly the effect of this menu of options, and for this reason it may be necessary to pursue a different experimental design. Below I describe how a RCT can be designed in order to test out this particular question.

Q. How does a RCT work in the context of a pricing pilot/demonstration project and under what circumstances would this design be the optimal choice?

A. Generally, a RCT works in the following way: customers are chosen at random from the population to participate in a study, and then, also in a random manner, placed into either the treatment or control group. The random placement of subjects into either the treatment or control group without allowing the subjects to choose which bin they are in is the fundamental aspect of a RCT. In contrast to a RED, the entire treatment group would be subjected to the price being tested. A RCT eliminates selection bias by mandating some customers to participate in the treatment group and others to participate in the control group. Which individuals are mandated to participate in which group is decided randomly; because of this randomization, the control group should be statistically similar to the treatment group (both in observable and unobservable manners) and bias is eliminated. Importantly, this bias is eliminated to a greater extent than under a RED pilot, and is therefore a marginally more robust experimental design. The results of this study would allow the utility to answer the question "How would the entire population's load shape or peak demand be affected by a mandatory change in rates to time-variant pricing?" Of course, mandating that customers adopt a particular rate can create a concern of equity and legal issues. In the NY setting, this type of RCT would likely be infeasible to implement.

However, this does not mean that a RCT cannot play a role in the piloting of rates in New York City. Although a RCT as usually implemented tests the effect of a mandatory regime, it is possible to design a RCT which not only avoids mandating customers onto a rate, but also allows the utility to test out two different alternatives, by 1) providing customers with the ability to choose from a menu of tariffs, and 2) posting a tariff online without actively pursuing/encouraging customers to adopt it. If the utility ultimately wants to pursue either of these alternatives with their mass market tariff, then a particular type of RCT can be implemented. I will call this approach an opt-in RCT and will now describe it in some detail.

An opt-in RCT would essentially allow the Company to answer the following question: “What will be the effect on consumption or load shape of allowing customers to opt into a tariff or menu of tariffs that is posted online?” The fundamental difference between an opt-in RCT and a regular RCT (in which customers are chosen at random to participate in the study) is that an opt-in RCT moves the randomization into a later stage.

In the first stage of an opt-in RCT, customers are allowed to opt into a treatment – for example, the utility may post a treatment or menu of treatments on a webpage and allow customers to choose which treatment they like. The randomization then occurs in the second stage, where those who opt into a particular treatment are randomly selected to be part of either the treatment or control group. Thus, while the treated group is allowed to choose their rate, a portion of those choosing the rate are placed into the control group; importantly, this results in a separate control group for each pricing treatment within the pilot (if the pilot is testing out multiple interventions). This random placement of opt-in customers into the control group eliminates selection bias because those placed into the control group are unobservably similar to those placed into the treatment group: both of these groups decided that opting into the rate would be a good move, and the only difference is that some were randomly selected for the control group. For those that opt in, this approach completely eliminates selection bias, as in a standard RCT.

It is important, at this stage, to point out that a control group made up of a random sample of the customer population *would* cause bias in the results, as those who opt into a rate are likely different from those who chose not to opt in (i.e., the remaining customer population). Thus, it is essential that the control group consist of customers who have made the same pricing choices at the outset as the treatment group. It is also important to note that the estimated outcomes from an opt-in RCT experiment are only valid for those opting in. That is, it would not be appropriate to suppose that the results derived from the opt-in RCT would hold for the general population if a pricing structure that was tested on an opt-in basis was deployed as a default or mandatory pricing structure. This is because it is likely that those that do and do not opt into the pilot are different in unobservable ways that could impact how they respond to the treatment.

The placing of customers into the control group can be done in two ways, known as "recruit and deny" and "recruit and delay". "Recruit and deny" randomizes across the volunteers and places some directly in the control group. "Recruit and delay" allows all volunteers to be in the treatment group, but for a random sample of the volunteers, consumption is measured over a given amount of time prior to the treatment; the pre-treatment measurements are then the control for the treatment group. The benefit of "recruit and delay" is that all customers who volunteer for the pilot get to participate in it; however, if the utility pursues a "recruit and delay" strategy, it is important that the pilot participants understand when they are recruited that they may not be subject to the rate immediately in order to not cause backlash.

The benefit of "recruit and deny" is that there is no need to gather pre-treatment measurements.

Q. What is the difference in costs for a RED vs an opt-in RCT pilot?

A. The main difference in costs associated with a RED vs an opt-in RCT pilot comes down to recruitment. In the RED as described above, customers are randomly selected from the population to participate in the control group (as described earlier, those who opted out or chose not to opt in to the alternative pricing offering continue to be in the "treatment group", and the control group is a sample of customers who were not encouraged to participate). The consumption patterns of the control group are measured over time, but the control group customers are generally not subjected to any treatment (though having this control group participate in a survey would be a best practice; see the discussion of surveys later in the testimony). Thus, there is no need to spend money recruiting the control group.

In the opt-in RCT, on the other hand, the control group needs to be recruited from those who choose to opt-in to the treatment. Thus, the recruitment costs could be twice as large under an opt-in RCT as under the RED.

However, the RED also can become expensive in terms of recruitment depending on whether the treatment is default versus opt-in. Because customers can opt out of (or choose not to opt in to) the encouraged treatment under RED, a RED pilot would require that the number of customers contacted and encouraged to participate be large enough to ensure that an adequate number of customers accept the offer and so are exposed to the pricing mechanism. It has been shown that customers tend to stay with defaulted rates at much higher proportions than customers in an opt-in regime select the opt-in rate; in 2013, Lawrence Berkeley National Lab reported that for opt-in rates, the average acceptance was 14%, compared to an average acceptance of 82% for default/opt-out rates. (Todd, Annika, Peter Cappers, and Charles Goldman. "Residential customer enrollment in time-based rate and enabling technology programs." Lawrence Berkeley National Laboratory (2013).). Thus, the more customers are placed into an opt-out treatment rather than an opt-in treatment, the lower the recruitment costs will be for the RED. *Therefore, the experimental design with the lowest recruitment costs is an opt-out RED.*

Q. How large should treatment groups be?

A. Getting the size of the treatment groups correct is essential in order to be able to correctly identify the impacts on changes in consumption. Defining sufficient sample size depends on the Minimum Detectable Effect ("MDE"), the proportion of the sample receiving the treatment, the statistical confidence level, the statistical power of the test, and the skewness of the load curve. The U.S. Department of Energy produced guidelines for sample sizes that demonstrate the interaction of these parameters with sample size (*see* Exhibit ___ (ES-3) at pages 10-11).

The proportion of volunteers placed into the treatment group rather than the control group is also important. The ability of the researcher to correctly identify an effect (i.e., the "power" of the test) is maximized by placing half the volunteers into the treatment group, but only if the variances in the expected outcomes (such as a change in load shape or reduction in peak demand) are the same across the treatment and the control. It may not be possible to know what these variances are before the experiment is conducted; thus, the utility can assume they are the same *a priori* and place half the customers in the treatment and half in the control group. However, if the treatment is more expensive than the control, the optimal number of customers in the treatment and control groups can depend on the relative costs of the treatment and the control groups. The equation used to calculate the optimal number of customers in the treatment vs control group is:

$$\frac{n_1^*}{n_0^*} = \sqrt{\frac{C_0}{C_1}\frac{\sigma_1^2}{\sigma_0^2}}$$

where 0 indexes the control and 1 indexes the treatment, n* is the optimal number of customers in each group, C is the cost of either the control or treatment, and σ^2 is the variance of the respective outcomes of the experiment (as described above). This equation demonstrates that if the cost of the treatment is substantially larger than the cost of the control, the optimal relative number of customers placed into the treatment decreases (List, John A., Sally Sadoff, and Mathis Wagner. "So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design." Experimental Economics 14.4 (2011): 439-457).

Q. What information gathering should occur in order to ensure that the right baseline information will be available for analysis after the pilot is performed?

A. It is important to have information on the customers that participate in both the control and treatment groups. Surveys can be used to track demographic and household characteristics (such as existence of appliances, environmental preferences, number of people in household, income, dwelling type and size, understanding of electricity prices, etc.).

Q. Why is it important to gather this information?

A. This information is needed for several reasons.

1. It can be used to control for observable attributes in the final estimation.

2. It can be used to ensure that the treatment and control groups are balanced (i.e., similar) in observable attributes. This is important due to the fact that while randomization in the experimental design (either RED or RCT) helps to minimize selection bias, randomization does not necessarily ensure, *a priori*, that samples are balanced. Having information on household characteristics and preferences can help balance the sample prior to estimation. Thus, surveying those in the control group will help to ensure that the groups are balanced.

3. It can be used to reveal how elasticities vary depending on personal characteristics and to tailor future programs accordingly. Today, customer segmentation by electric utilities is often limited to identifying customers with particular classes, such as residential and commercial classes. However, more specific segmentation is possible based on customers' temperament and other less obvious attributes; incorporating differences in behavioral responses of different segments will improve future pricing offerings and will have energy and environmental policy ramifications.

Q. **How should surveys be performed?**

A. Initial surveys should be completed at the outset of the pilot, and a small monetary payment for survey completion can be provided. Having a payment for survey completion has been shown to increase the number of responses to the survey; though the size of the payment need not be large, increasing the payment can help to ensure that certain parts of the population participate in the survey (*see, e.g.,* James, Jeannine M., and Richard Bolstein. "Large monetary incentives and their effect on mail survey response rates." Public Opinion Quarterly 56.4 (1992): 442-453; Singer, Eleanor, and Richard A. Kulka. "Paying respondents for survey participation." Studies of welfare populations: Data collection and research issues (2002): 105-128; and Coughlin, Steven S., et al. "The effectiveness of a monetary incentive on response rates in a survey of recent US veterans." Survey Practice 4.1 (2013).) A second survey should be completed at the end of the pilot in order to gather information about customer satisfaction with different pricing structures. Quantifying the acceptance levels and identifying the level of understanding of prices is essential for moving toward a mass market tariff. Post-pilot surveys are very important, and should also be requested from those who choose to exit the program prior to completion. Payment for completion of the post-pilot survey is also beneficial, in order to ensure increased survey responsiveness as described above.

Q. Con Edison will begin its AMI deployment in 2017 and some customers will have access to AMI functionality in 2018. Does this mean that they can begin their pilots at that time?

A. Yes. However, it is important for the Company to begin preparing sufficiently ahead of time. For example, SMUD conducted a pricing pilot in 2012 and began to work towards setting up the pilot one year ahead of time. This pre-pilot time was spent conducting focus groups and surveys to ensure that their marketing strategy for the pilot would be most effective when deployed (*see* Exhibit __ (ES-3) at page 14). Recruitment and initial surveys of the treatment group can also take time and so should begin prior to the deployment of AMI, if the Company would like to launch the pilot in 2018.

Q. How should the study design address the fact that customers move?

A. If a significant portion of the residential population moves each year, this must be added to the sample size to ensure that the remaining size of each treatment group remains in the appropriate size range for the duration of the pilot. For example, if 25% of the population can be expected to move, the initial sample size should be 25% larger than the targeted minimum number at the outset, to account for natural attrition.

Q. How should study design account for customers wanting to exit the pilot?

A. Ensuring that sample sizes remain sufficiently large at the end of the pilot or demonstration project will require consistent outreach and engagement of the customers. Allowing customers to exit the program is important as exiting behavior provides information about satisfaction with the alternative rate structure under consideration; high levels of attrition in one price treatment shows that the rate under consideration would likely not be a successful price structure in the overall population. Locking customers in for the entire course of a pilot can cause resentment among participants and hurt future attempts at engaging customers with time-variant pricing. Furthermore, having a lock-in clause may reduce the willingness of customers to enter into the pilot, thereby increasing recruitment costs. However, given the costs associated with bringing customers into a pilot, it is desirable to minimize attrition. Requiring a 60-day notice period for exiting the project can provide time for the Company to reach out to the unsatisfied customer, with the purpose of understanding the reason why the customer wants to leave and identifying any possible remediation to avoid attrition.

Q. Has a pilot incorporating the level of statistical rigor and complexity that you recommend ever been implemented to test out pricing structures, and was it effective?

A. Yes. For example, SMUD conducted a carefully structured pilot in 2012-2013 to test out three different time-variant pricing structures (critical peak pricing, time-of-use, and a combination of the two), a technology intervention (in-home display), and the impact of two different recruitment options (default versus opt-in). SMUD also employed both an opt-in RCT and a RED, the latter being employed to test how the different recruitment options (default vs opt-in) affected consumption patterns. Furthermore, they employed a number of different educational and marketing strategies to understand how outreach can impact behavior and lead to more attractive rate structures. Importantly, one of the largest expenses of the pilot was attributed to the outreach program, which included two parts: a marketing campaign in the year prior to the pilot implementation, and a shift towards customer retention after the pilot had begun. SMUD's marketing campaign involved conducting 25 focus groups and 4 surveys (with up-front cash incentive payments of $5, resulting in 38% response rates; *see* Exhibit ___ (ES-4) at pages 89 and 92) targeted to over 2,000 customers. This helped them to develop successful materials and communication tools used during the pilot.

Once the pilot was in motion, they employed a number of different tactics to retain customers and continue to educate them on how to respond to pricing. These included different forms of communication such as direct mail, emails, mass media, as well as communication through several social media outlets. The pilot was successful overall, resulting in high levels of acceptance of the different types of rates and beneficial outcomes in terms of reduced demand and strain on the system. These high levels of acceptance were revealed through the post-pilot surveys, which were paid $2, and had response rates of 40% (*see* Exhibit ___ (ES-4) at page 119). Because of their thoughtfully designed pilot, SMUD was able to identify 1) which types of pricing mechanisms had the largest impact on demand reductions, and 2) whether technology was helpful. As a result, they were able to conduct a cost-benefit analysis of deploying time-variant prices across their service territory (depending on the pricing mechanism, technology deployed, and method of recruitment). SMUD published a report for the U.S. Department of Energy describing the pilot, including the design, set up, and results of their analysis (*see* Exhibit ___ (ES-4)).

Q. Does this conclude your testimony?

A. Yes, it does.