## Description of CER Experiment

\footnote{The detail about the CER experiment presented hereinbelow is a summary of A DOCUMENT.}

The Commission for Energy Regulation (CER), the regulator for Ireland's electricity and natural gas sectors, conducted the Smart Metering Electricity Consumer Behavior Trial (hereafter, the``trial'') between July 2009 and December 2010. As part of the Smart Metering Project initiated in 2007, the trial's purpose was to assess the impact of various TOU tariff structures, along with different Demand-Side Management (DSM) stimuli, on residential electricity consumption. The CER carefully recruited households to construct a representative sample of the national population. Opt-in to the trial was voluntary. Participants received balancing credits not to incur any extra costs than if they were on the regular electric tariff (i.e., the flat rate of 14.1 cents per kWh). Also, they received a thank-you payment of 25 cents after pre- and post-trial surveys. All credits were distributed outside the treatment period to avoid unintended effects on participants' electricity consumption.\footnote{While the first balancing credit was paid at the end of the base period (i.e., in December 2009), the participants received the second one at the immediate month after the treatment period (i.e., in January 2011). And the after-survey payments were credited to their bill with the balancing credits.}

The households who voluntarily opt-in to the experiment were randomly assigned to control and treatment groups.\footnote{The optimal sample size for the trial was determined to be 4,300 participants in the design phase. In the allocation phase, 5,028 households were assigned to the control and treatment groups to consider participant attrition. The published CER experiment data only include electricity consumption data only for 4,225 households.} Baseline electricity consumption data were collected during the second half of 2009 (i.e., July to December 2009), while the treatment period was from January through December 2010. All treated households received two kinds of treatments simultaneously: 1) one of four TOU tariff structures and 2) one of four DSM stimuli. In other words, there were 16 distinct treatment subgroups. The CER provided the treated with a fridge magnet and stickers to facilitate accustoming them to the TOU pricing schemes.\footnote{The fridge magnet and stickers outlined the timebands during which different prices were applied. Moreover, they were tailored for each tariff group.} On the contrary, the households allocated to the control group remained on the normal flat tariff.

The four TOU tariff structures had different prices during each of the three rate periods in a day. The day in the treatment period was divided into three periods: 1) peak rate period from 5:00 p.m to 7:00 p.m., 2) day rate period from 8:00 a.m. to 5:00 p.m. and from 7:00 p.m. to 11:00 p.m., and 3) night rate period 11:00 p.m. to 8:00 a.m. As illustrated in Figure (X), the order of magnitude in rate changes for the peak rate period is the opposite of that for the rest of the rate periods. The reason for designing the tariff structures in such a way is to enable participating households to face similar energy bills on average when maintaining their electricity consumption pattern, regardless of the rate structures to which they were assigned.

The four DSM stimuli differed in the degree or the frequency of feedback on each household's electricity usage information. The control group just received their bills at the same cycle (i.e., bi-monthly). On the contrary, all households assigned to the treatment group received a detailed energy usage statement combined with their bill, including their detailed weekly usage, average weekly costs, tips on reducing electricity use, and comparisons to peer households. The first stimulus subgroup received a bill with a detailed energy statement bi-monthly, while the second subgroup received the documents every month. An electricity monitor, which displays their usage against their pre-set daily budget, was also provided for the households allocated to the third DSM stimulus subgroup. The last stimulus subgroup received an Overall Load Reduction (OLR) incentive. Under the OLR incentive, the households that reached their 10% reduction target over the eight-month period beginning May 2010 were rewarded with 20 Euros.\footnote{A household's reduction target in electricity consumption was set based on the participant's actual usage during the first four months of the treatment period. And the households in the last DSM stimulus subgroup were updated on their progress with each bi-monthly bill.}

## Description of CER Experiment Dataset

The CER experiment dataset disseminated by the Irish Social Science Data Archive (ISSDA) consists of participating households' electricity consumption and survey data.

Throughout the baseline and treatment periods, meter reads for residential participants were recorded in 30-minute intervals. The high granularity of the electricity consumption data generated from a well-designed experiment enables quantifying where the energy savings stem from when transferring to TOU electricity pricing for each of the three rate periods.

The survey data contains participants' responses to more than 300 questions in both pre- and post-trial surveys. The primary purpose of the two surveys was to gather trial-associated experiential and attitudinal feedback from the households. The surveys also included questions intended to collect residential participants' socio-demographic characteristics. In addition, questions about the physical attributes of their house were included in the surveys.

My empirical analysis throughout this paper uses the sample constructed by including observations only for non-holiday weekdays in the published electricity consumption data because the TOU rates were active just on those days.\footnote{The sample is a panel data of households with reliable meter reads only. Specifically, the residential participants who had no consumption for eight days or more are excluded from the sample. In addition, I drop the meter reads for the days when several participating households' consumption data were missed. \par

Although I utilize the sample satisfying the following criteria too for the empirical analysis, applying the criteria does not change the results: 1) Exclude the day immediately following the end of daylight-saving time due to noticeably different consumption levels in the same hours of the day; 2) Drop the observations for the last five days of the baseline and treatment periods because of extraordinarily high electricity demand on those days.} This process results in 4,096 households.

The control and treatment groups in the sample are largely balanced, as shown in TABLE. Such indifferences between the two groups over many observables are consistent with previous studies that also examined the CER experiment dataset.\footnote{To check the balance between the control and treatment groups, Prest (2020) employs a linear probability model, while a probit model is used in Pon (2017).}

## Description of Weather Data

In this research, weather data are an essential element. The main interest of the majority of TOU papers has been to measure how residential consumers respond to TOU prices or the heterogeneity in the responsiveness of households across different information stimuli. Hence, those studies usually do not control for temperature variations directly. For example, Pon (2017) and Prest (2020), which also exploited the CER experiment dataset, added weak-of-sample and month-by-year fixed effects (FEs) to their specifications, respectively, in order to control for variations in usage due to seasonal changes. On the other hand, the primary objective of this paper is to decompose the TOU-price-inducing demand reductions into two parts--reductions in temperature-control use and those in non-temperature-control uses. Since the electricity consumption for temperature-control use is driven by weather, especially temperature, it is necessary to link the 30-minute interval consumption data and reliable weather data with an appropriate level of resolution.

I utilize average daily temperatures to quantify the energy savings of each of the two different sources after introducing TOU prices. More granular temperatures, like hourly temperatures, are not a dominant determinant of electricity demand for temperature-control use at a point in time. It is not easy to believe that residential customers adjust their electricity consumption according to ever-changing outside temperatures elaborately and instantly. Furthermore, as shown in FIGURE, their electricity demand is the lowest in the early morning, the coldest time of the day. Considering those two points, I measure the TOU-tariff-inducing reductions in electricity consumption conditional on the average heating need in a given day.

I exploit hourly temperature data for the Dublin airport weather station, provided by Met \'{E}ireann, Ireland's National Meteorological Service, to compute average daily temperatures. There is no available location information in the published CER experiment dataset for privacy and security reasons. Therefore, it is not possible to match a participant's consumption data with weather data of the closest weather monitoring station to him. But fortunately, in Ireland, temperatures do not vary much across areas for a given day. As demonstrated in the TABLE, the temperature correlations between the Dublin station and stations near densely populated cities are high. Because of this reason, I use the mean daily temperatures obtained by averaging the Dublin airport station's hourly temperatures as the representative temperatures in the following analysis.

Using the average daily temperatures, I calculate daily HDDs. Instead of 65 degrees of Fahrenheit, which is a normal base temperature in the United States, 60 degrees of Fahrenheit is utilized to compute daily HDDs, according to Liu and Sweeney (2012). The upper part of the FIGURE shows that many days in the treatment period had lower average daily temperatures than the lowest one during the baseline period. The evolving pattern of heating-purpose demand for electricity on days with extreme--at least in Ireland--temperatures could be significantly different under distinct rate structures--flat rate and TOU rates. If this is true, the lack of counterfactual consumption observations will cause bias in the measured impact of introducing TOU rates on household electricity consumption. So, I drop observations for those days in the treatment period when constructing the sample to address the possibility.

# Empirical Strategy

FIGURE, which shows household average daily electricity consumption over temperature and the pre and post differences in the consumption, clearly demonstrates the motivation of this research project.\footnote{An important feature also stands out from the figure: the minimum household electricity consumption occurred at around 60 degrees of Fahrenheit. This phenomenon supports the setting of the reference temperature for calculating daily HDDs at the very level.} As illustrated in Panel A of the figure, household demand for electricity grew as the temperature decreased. In other words, in addition to temperature-insensitive electricity demand (i.e., for non-temperature-control uses), there was a sizeable demand for electricity for heating (i.e., for temperature-control use) in Irish households, which seems to be highly responsive to temperature variations. In this research, I determine not only how much consumption changes, on average, in response to the time-varying tariffs but also how their impact varies across days with different temperatures. In other words, the dynamic-pricing-causing effects on heating and non-heating electricity uses are separately estimated to figure out the primary source of energy savings. As shown in the figure, households in the control group also consumed less electricity during the treatment period, especially on days with low temperatures, although their percentage reduction is smaller than that of the treated households.\footnote{In Panel A, non-treated households consumed more electricity during the baseline period, especially on days with higher heating needs. The fact that the total HDDs during the baseline period were generally greater than those during the treatment period for a given temperature bin could explain the phenomenon.} This suggests the necessity of employing an identification strategy that deals with the before and after differences in electricity consumption of households remained in the traditional tariff structure (i.e., a flat price of 14.1 cents for all hours).

Because the CER experiment dataset primarily utilized in the following empirical analysis was generated from a carefully developed randomized controlled trial (RCT), in principle, the effect of the TOU tariffs on household electricity consumption can be measured simply through the difference in average usage between the two groups during the treatment period.\footnote{Because random assignment of participating households puts selection bias right, observed differences in electricity consumption between the control and treatment groups after introducing the TOU tariffs are only attributable to their differences in exposure to the time-varying electricity prices.} However, due to the non-trivial difference in electricity demand between the control and treatment groups during the baseline period, I follow the previous studies utilizing the same experiment and employ a DID approach to estimate the electricity savings caused by the TOU pricing program.

I include the temperature as an explanatory variable directly in my econometric models. In the previous papers using the identical dataset, fixed effects (FEs) were utilized to control for time-varying factors that influenced household electricity consumption. Since those studies focused on quantifying how households responded, on average, to the TOU price regimes newly introduced, adding such FEs to their models served their research purpose. In other words, they did not need to explicitly model the relationship between temperature and household electricity consumption to estimate the average treatment effects (ATEs). However, a primary interest of this research is to understand how electricity savings vary with the temperature after shifting to TOU prices. Therefore, more direct controls rather than FEs, not sweeping out temperature variations across days, are required in my empirical analysis. For that reason, I extend the typical panel DID specification and allow the treatment effect to vary as a function of the daily average temperature.\footnote{Under three identifying assumptions, applying the DID strategy to measure energy savings obtained from adopting the TOU prices makes sense. First, the parallel trend assumption is required for the DID estimator. Considering that the 30-minute interval meter reads for participating households were collected from a trial, the assumption means that the pre-treatment-period load profile for the treated households should be very similar to that for the non-treated households. FIGURE A showing average within-day load profiles for the two groups during the baseline period supports the plausibility of the parallel trend assumption. In addition, the electricity consumption profile for the control group illustrated in FIGURE B, which smoothly evolved over the entire experiment period although heavily fluctuated day to day, suggests its high reliability as a counterfactual under the assumption.

The second identifying assumption necessary for the plausibility of the identification strategy employed is the assumption of common temporal shocks. This assumption implies that a treatment-status-irrelevant unexpected event occurring at the same time as or following the deployment of the dynamic prices should have the same impact on both the control and treatment groups. Although the common shocks assumption cannot be tested directly, the similar trends in electricity demand profiles for the control and treatment groups shown in FIGURE B support the assumption required for the DID approach.

Third, the stable unit treatment value assumption (SUTVA) must hold too. The SUTVA requires that introducing TOU prices did not affect the electricity consumption of the untreated households. That is, the SUTVA allows no spillovers. During the recruitment process, the locational distribution of the participating households was aligned with that of the total Irish population to construct a representative sample of the national population. Because only a few thousand households scattered geospatially participated in the nationwide experiment, it is unlikely that the treated households influenced the households allocated to the control group. This again supports the SUTVA required under the DID identification strategy.} That is, I estimate the ATEs of the dynamic prices on household electricity demand by exploiting the within-household electricity consumption changes across not only periods but temperatures.\footnote{The attrition rate during the RCT was about 20%. The main reasons for participant attrition were changes in tenancy and supplier. Due to the imperfect compliance, the estimates must be interpreted as local average treatment effects (LATEs). However, according to CER (2011), attrition was unlikely to be associated with the RCT. Furthermore, the level of attrition varied only marginally across treatment status.}