## 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 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, described in detail later. 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 from 11:00 p.m. to 8:00 a.m. As illustrated in Figure \ref{Figure:Time-Of-Use-Pricing-Structures}, the order of magnitude in rate changes during 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 received their bills in 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 the energy savings 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 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 a longitudinal sample that consists of observations satisfying certain conditions only. First of all, the sample is constructed by including observations only for the second half of each experiment period.\footnote{I exclude the observations for the first half of the treatment period because there is no counterpart observation in the baseline period.} From this sample, I drop observations for non-holiday weekdays in the published electricity consumption data because the TOU rates were active just on those days. And then, only households that continuously exploited non-electric fuels for their space and water heating during the experiment periods (i.e., the baseline and the treatment periods) are preserved in the sample.\footnote{From the survey data, it is possible to find out what type of fuel each responding household used for each heating purpose during each period. \par

There are two reasons why only non-electric-heating households are exploited in the following empirical analysis. First, in Ireland, non-electric fuels, such as oil, gas, and solid fuels, fulfill most of the residential heating demand. Specifically, according to \cite{Heating-and-Cooling-in-Ireland-Today\_SEAI\_2022}, only 4\% of Irish households utilize electricity to heat their space and water. Therefore, with respect to fuels for heating in Ireland, the sample consisting of non-electric heating households only is representative. Second, as Figure \ref{Figure:Pre-and-Post-Treatment-Household-Average-Daily-Electricity-Consumption} demonstrates, even non-electric-heating households consumed more electricity as temperatures decreased. In other words, electricity is still essential for non-electric-heating households to warm their space or water. Hence, the sample, including non-electric-heating households only, is well aligned with one of the primary purposes of this research: to evaluate the impact of TOU pricing on for-heating-purpose residential electricity consumption separately.} Moreover, among the non-electric-heating households, those with unreliable meter reads are excluded from the sample.\footnote{To be specific, 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 examining the CER experiment dataset.\footnote{To check the balance between the control and treatment groups, \cite{Peaking-Interest:How-Awareness-Drives-the-Effectiveness-of-Time-of-Use-Electricity-Pricing\_Prest\_2020} employs a linear probability model, while a probit model is used in \cite{The-Effect-of-Information-on-TOU-Electricity-Use:An-Irish-Residential-Study\_Pon\_2017}. \par

Both papers point out that voluntary opt-in might cause bias in the estimated treatment effect. Refer to \textit{5.5.3 External Validity} in \cite{Peaking-Interest:How-Awareness-Drives-the-Effectiveness-of-Time-of-Use-Electricity-Pricing\_Prest\_2020} and \textit{5.1 Addressing Self-Selection} in \cite{The-Effect-of-Information-on-TOU-Electricity-Use:An-Irish-Residential-Study\_Pon\_2017}. }

## Description of Weather Data

In this research, weather data are an essential element. The main interest of most TOU papers has been to measure how consumers respond to TOU prices or the heterogeneity in their responsiveness across different information stimuli. Hence, those studies usually do not control temperature variations directly. For example, \cite{The-Effect-of-Information-on-TOU-Electricity-Use:An-Irish-Residential-Study\_Pon\_2017} and \cite{Peaking-Interest:How-Awareness-Drives-the-Effectiveness-of-Time-of-Use-Electricity-Pricing\_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 electricity 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 and non-temperature-control uses. Since the electricity consumption for temperature-control uses is driven by weather, especially temperatures, it is necessary to link the 30-minute interval consumption data and reliable weather data with an appropriate level of resolution.

The electricity savings associated with for-heating electricity consumption are disaggregated using average daily temperatures from the total savings resulting from the introduction of TOU prices. More granular temperatures, like hourly temperatures, are not a dominant determinant of electricity demand for temperature-control uses at a point in time. It is not easy to believe that households adjust their electricity consumption according to ever-changing outside temperatures elaborately and instantly. Furthermore, as shown in Figure \ref{Figure:Average-Consumption-by-Hour-of-Day}, 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 impossible to match a participant's consumption data with the 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 Table \ref{Table:Correlations-in-Temperature-for-Major-Cities-in-Ireland}, the temperature correlations between the Dublin station and stations near densely populated cities are evident. Because of the positive correlations, 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 Heating Degree Days (HDDs). Instead of 65 degrees of Fahrenheit ($^{\circ}F$), a normal base temperature in the United States, 60$^{\circ}F$ is utilized to compute daily HDDs, according to \cite{The-Impacts-of-Climate-Change-on-Domestic-Natural-Gas-Consumption-in-the-Greater-Dublin-Region\_Liu-and-Sweeney\_2012}. The upper part of Figure \ref{Figure:Breakdown-of-Hourly-ATEs-in-the-Peak-Rate-Period} 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 temperatures could be significantly different under distinct rate structures--flat and TOU rates. If this is true, the lack of counterfactual consumption observations will cause bias in the measured impact of introducing TOU pricing on household electricity consumption. So, I drop observations for those days in the treatment period when constructing the sample to address the potential threat to the identification.

# Empirical Strategy

Figure \ref{Figure:Pre-and-Post-Treatment-Household-Average-Daily-Electricity-Consumption}, showing not only household average daily electricity consumption over temperature (in Panel A) but also percentage changes in electricity consumption (in Panel B), clearly demonstrates the motivation of this research.\footnote{An important feature also stands out from the figure: the minimum household electricity consumption occurred at around 60$^{\circ}F$. 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 gradually as the temperature decreased. That is, for Irish households, in addition to temperature-insensitive electricity demand (i.e., for non-temperature-control uses), there was a sizeable electricity demand for heating (i.e., for temperature-control uses), which seems to be highly responsive to temperature variations. In this research, I determine not only how much variations in household electricity consumption occur, on average, in response to the deployment of the TOU tariffs but also how their impact varies according to daily HDDs. In other words, the dynamic-pricing-causing effects on for-heating and non-for-heating electricity uses are separately estimated to figure out the primary source of electricity savings. As shown in the figure, households in the control group consumed less electricity during the treatment period, especially on days with low temperatures, although their percentage reductions are smaller than those 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 for a given temperature bin, the total daily HDDs during the baseline period were generally greater than those during the treatment period is a plausible explanation for the phenomenon.} In light of this, it is necessary to employ an identification strategy that accounts for the before and after differences in household electricity consumption under the traditional tariff structure (i.e., a flat rate of 14.1 cents per kWh for all hours).

I employ a Difference-In-Differences (DID) approach to estimate the electricity savings caused by the TOU price program. The CER experiment dataset primarily utilized in the following empirical analysis was generated from a carefully developed Randomized Controlled Trial (RCT). So, 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, as discussed, there exist non-trivial differences in electricity demand between the control and treatment groups during the baseline period. Following the previous studies exploiting the same data, I utilize a DID estimator to address the possible source of bias.

I include daily HDDs 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 influencing 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 a typical panel DID specification and allow the treatment effect to vary as a function of daily HDDs.\footnote{Under three identifying assumptions, applying a DID strategy to measure electricity 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 during the trial, the assumption implies that the pre-treatment-period load profile for the treated households should be very similar to that for the non-treated households. Figure XYZ-1, 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 XYZ-2, which smoothly evolved over the entire experiment period although heavily fluctuated daily, suggests its high reliability as a counterfactual under the assumption.

The assumption of common temporal shocks is the second identifying assumption necessary for the plausibility of the identification strategy employed. 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 XYZ-2 support the assumption required for the DID approach.

Third, the stable unit treatment value assumption (SUTVA) must hold too. The SUTVA requires that introducing the 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 rate 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 such imperfect compliance, the estimates must be interpreted as local average treatment effects (LATEs). However, according to \cite{Electricity-Smart-Metering-Customer-Behaviour-Trials-Findings-Report\_CER\_2011}, attritions were unlikely to be associated with the RCT. Furthermore, the level of attritions varied only marginally across treatment status.}