Measuring the Impact of Time-of-Use Pricing on Electricity Consumption: Evidence from Spain

Jacint Enrich Ruoyi Li Alejandro Mizrahi Mar Reguant*

February 2023

Abstract

We evaluate the effect of a time-of-use pricing program introduced in Spain on residential electricity consumption. Using a Difference-in-Difference approach, we find that households responded by reducing consumption during peak hours, although we do not find significant evidence of load-shifting. We then use machine learning for variable selection and show that it can help to obtain more precise estimates. We find that the program could have reduced consumption up to 12% during peak periods, partly offset by a 1.7% increase during off-peak hours. We also find evidence of habit formation during periods of uniform pricing, accompanied by an adaptation process that ends with a permanent change in consumption behavior. The results suggest that a predetermined pricing program can enhance consumer awareness and increase household price elasticity, thus making it an effective policy tool to reduce peak electricity demand and improve market efficiency.

Keywords: demand response, dynamic pricing, electricity.

JEL: H23, L94, Q41, Q48

*Jacint Enrich: Barcelona School of Economics, jacint.enrich@bse.eu. Ruoyi Li: PhD, Paris School of Economics, ruoyililin@gmail.com. Alejandro Mizrahi: PhD, Toulouse School of Economics, alejandro.mizrahi@tse-fr.eu. Mar Reguant: Department of Economics, Northwestern University, Barcelona School of Economics, CEPR, and NBER, mar.reguant@northwestern.edu. Reguant acknowledges the support of the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101001732-ENECML).

1 Introduction

The costs of generating electricity vary significantly between sources and technologies. Simultaneously, electricity demand fluctuates considerably over time, yielding periods of peak demand. In those periods, the optimal mix includes generating capacity with high marginal costs to balance supply and demand. If prices faced by consumers do not reflect these cost variations, consumption will be distorted and generate market inefficiencies (Joskow and Wolfram, 2012). Thus, several governments have tried pushing regulatory reforms from time-invariant to dynamic or marginal cost pricing. These reforms are more and more urgent as the share of intermittent renewable generation increases.[1]

The efficiency gains from dynamic pricing will depend on the price mechanism and the presence of potential market failures and frictions. For instance, a day-ahead real-time pricing (RTP) scheme with constant hourly price changes may fail in the presence of imperfect information or adjustment costs. In that case, introducing automated load-shifting technologies might be necessary to induce a significant demand response (Bollinger and Hartmann, 2020; Fabra et al., 2021). Alternatively, policies with predetermined retail pricing could ease information constraints and short-term costs of shifting consumption. Time-of-use (TOU) pricing with predetermined tariffs varying between different periods within a day is an example of such a mechanism.

We study the introduction of a TOU pricing program in Spain on residential electricity consumption. In June 2021, the Spanish government introduced a new regulation where system and network charges (which traditionally amount to 50% of the total electricity bill) would be charged at three different marginal prices depending on the hour of the day and the day of the week. According to the new regulation, peak hours range from 10 am to 2 pm and from 6 pm to 10 pm, and off-peak hours cover hours from 12 am to 8 am during working days. The remaining hours in working days are considered mid-peak, while all hours on weekends and national holidays are off-peak.

To estimate the causal impact of the policy, we compare the evolution of electricity demand in Spain with its nearby country, Portugal. Portugal is a natural control group for several reasons. First, geographically both countries have similar weather conditions, a determinant factor of electricity demand. Perhaps more importantly, Spain and Portugal trade in the same electricity wholesale market. This attenuates the possibility that our results are driven by differences in supply curves, consequently affecting equilibrium prices and quantities.

In the estimation, we focus on the responses of Spanish residential consumers who were under the regulated tariff during the period spanning from 2018 to September 2021.[2] The regulated tariff is offered by five main distribution groups, each of them being the only retailer entitled to offer the regulated tariff in a given distribution area. After the policy implementation, regulated prices increased by more than 60% in average, partly driven by rising prices in the wholesale market and partly by the policy change. Importantly, increases in energy costs impacted prices at all hours of the day, while the TOU policy has large price jumps within the day. This allows us to identify the effect of changes in TOU prices. Indeed, the reform provided large incentives to consume during cheap hours of the day: after the policy, prices at off-peak hours were 86% lower, while

[1] According to Wolak (2019), setting day-ahead and real-time prices that reflect transmission network constraints will reduce the costs of integrating intermittent renewable generation capacity.

[2] We focus on this subset of consumers due to data availability and the fact that consumers on commercial tariffs do not necessarily have a time-of-use pricing scheme mimicking the regulated tariff, although many do. Households on the regulated tariff represent around 39% of residential consumers in our sample period.

consuming at peak hours was 200% more expensive.

We estimate two empirical models. The first is a Differences-in-Differences (DID) fixed effects model, comparing hourly differences in electricity consumption before and after the policy in Spain's distribution areas and Portugal, controlling for a rich set of fixed effects. However, this richness of the data makes it difficult to choose the correct specification among many possible candidate controls. We thus turn to machine learning techniques (ML), particularly LASSO and forest estimators, for variable selection. In a nutshell, our empirical strategy consists of two steps. First, we use pre-treatment data to create a firm-specific electricity consumption model. We then use these estimates to create out-of-sample predictions for the post-treatment period following Burlig et al. (2020). While the difference between the actual outcome and the prediction in treated units already gives us an idea of the treatment effect, it does not control for time trends that would otherwise be accounted for in a DID setting. In a second step, we regress these prediction errors on a treatment dummy and the same set of fixed effects used in the DID strategy previously presented. The identifying assumption requires treated and control units to trend similarly in prediction errors, relaxing the standard parallel trend assumption in DID settings.

Our results indicate an overall reduction in electricity consumption. In our preferred specification of the standard panel fixed effect regression, results are insignificant for off-peak and mid-peak hours, while we observe a 10.3% decrease during peak hours. Surprisingly, we do not find evidence that using machine learning techniques helps to reduce the sensitivity of the estimates across different specifications. Nonetheless, the algorithm helps to reduce the noise in the dependent variable, leading to more precise estimates. In particular, consumption decreased by 1.6%, 7.8%, and 12.1% during off-, mid-, and peak hours, respectively.

We then split off-peak hours between weekdays and weekends. The goal of this analysis is threefold. First, it allows us to identify the within-day shift in consumption. Second, we define "fake" mid- and peak hours during the weekend, thus identifying possible effects of the policy that could be related to habit formation. Third, methodologically, if unobserved confounders threatening the parallel trend assumption are similar between weekdays and weekends (for instance, due to aggregate shocks affecting one of the two countries), we may be able to reduce the bias of our estimates by computing the additional effect of the policy during weekdays when compared to weekends.

Notably, we observe a significant demand response during the weekends for all three periods. Indeed, the magnitude of the coefficients for off-peak and mid-peak hours during weekends is comparable to weekdays. In particular, looking at our ML model, weekend consumption increased by 1.5% during off-peak hours, while it decreased by 6.3% and 7.9% for mid- and peak hours. Moreover, once we control for weekend responses, weekday consumption during off-peak hours increased by 1.7%. With that, we define a lower bound of policy effects, this is, estimating the difference between weekdays and weekends. In that case, we still find a decrease in consumption during peak hours of 4.5%, although the coefficient is not statistically significant anymore. However, we think that this approach is to conservative, given that tests for parallel trends capturing differences in the evolution of the dependent variable before the policy was introduced fail to find significant effects during most hours of the day. All in all, it does not seem that households shifted consumption from weekdays to weekends, but from peak to off-peak within a day.

We complement our main analysis with two additional extensions. First, using price data and the exogeneity of the policy, we can compute household price elasticities. We find significant price elasticities ranging between -0.03 and -0.04. These estimates align with the literature on electricity demand estimation, where consumers are usually found to be relatively insensitive to price variations. With that, we compute back-of-the-envelope demand responses given the price changes, and we find that they fall short of the estimated effects of the policy. One possible explanation for these differences is that the response to variations in TOU tariffs is stronger than the response to changes in energy costs captured in our demand estimation analysis. Second, we present an exploratory analysis of the relationship between Google searches and the effect of the policy. Searching behavior is consistent with an adaption process, with households reducing their overall consumption during the first weeks, followed by a subsequent stabilization of policy effects and Google searches.

The paper is organized as follows: Section 2 reviews the literature on dynamic pricing in the electricity market. Section 3 describes the empirical setting and data. Section 4.1 presents the empirical strategy to identify the causal effect of the policy and discusses the results. Section 4.2 repeats the same analysis using the machine learning methodology and compares results with the standard panel fixed effect model. Section 5 provides a framework to compute price elasticities and compare them to the identified policy effects, as well as it relates these estimates with consumer searching behavior. Finally, Section 6 concludes.

2 Literature review

This paper contributes to the literature on time-varying pricing in the electricity market.[3] The rationalization of making electricity consumption more closely tied to the variations of the marginal cost of generation appeared already in the late 1950s (Steiner, 1957; Boiteux, 1960; Williamson, 1966). In the last two decades, there has been a push to reconsider dynamic pricing given the recent developments in the electricity market, such as the deployment of smart meters and the increasing share of intermittent renewable in the generation mix (Joskow and Wolfram, 2012).

Potential efficiency gains from dynamic pricing are studied, among others, by Borenstein (2005) and Borenstein and Holland (2005) in the context of real-time pricing RTP. They show that, under ideal market conditions, adopting RTP would reduce peak electricity production capacity and would lead to significant welfare gains in the long run (up to 11% for the California electricity market). Holland and Mansur (2006) show that benefits are relatively modest in the short run, suggesting that a large portion of welfare benefits originates from a reduction in the construction of new generation capacity. Moreover, RTP proponents argue that real-time pricing could alleviate market power since a more elastic demand would reduce firms' incentives to curb their output to increase prices (Borenstein, 2002). [4]

In addition, Holland and Mansur (2008) show that dynamic pricing does not always bring down emissions, at least in the short run. By using exogenous changes in temperature and economic activity, they link variations in load with variations in emissions, arguing that the conclusions could be helpful to link real-time pricing (affecting load variance) with environmental benefits. Unfortunately, the effect of a reduction in load variance

[3]Dynamic pricing methods are consistently used in other sectors where the capacity is limited in the short-term, such as airlines, hotels, and car rental firms (Elmaghraby and Keskinocak, 2003; McAfee and Te Velde, 2006; Gibbs et al., 2017).

[4]Poletti and Wright (2020) finds that efficiency gains from RTP are 41% larger in the presence of market power in the New Zealand market.

is pollutant and location dependant. In particular, the shift of consumption towards off-peak hours drives SO_2, NO_X, and CO_2 emissions down in regions where peak demand is supplied with oil-fired power stations. Still, the effect becomes the opposite when hydroelectric power is more generally used to cover peak demand.

The benefits of dynamic pricing materialize only if households react to prices by adjusting their demand. Nevertheless, research has found that consumers generally exhibit inelastic demand. Allcott (2011), using data from an RTP pilot program, finds that enrolled households have an average elasticity of -0.1, with no load shifting from peak to low price hours. Similar results are obtained by Fabra et al. (2021) studying the first large-scale deployment of RTP in Spain in 2015. Both papers point out that low price variation is one of the possible reasons that explain the lack of demand response. They also highlight the importance of information and adaptation costs.

Jessoe and Rapson (2014) and Bollinger and Hartmann (2020) analyze more explicitly the impact of these costs and the role of assisting technology. Jessoe and Rapson (2014) analyze the effect of providing in-home devices that display households' electricity usage in real-time. They show that households provided with information were three standard deviations more responsive to prices than those without. Bollinger and Hartmann (2020) show that consumers incur adjustment costs and thus cannot shift consumption in response to short-term price changes unless they use automation technologies - such as programmable communicating thermostats.

Even if TOU prices adjust more accurately to high demand than flat rates, they fail to acknowledge all the differences in marginal costs within a day. As a result, efficiency gains were first estimated to be just 20% relative to RTP, according to Borenstein (2005). Woo et al. (2013) study the effect of TOU pricing and conclude that households respond by reducing electricity usage during peak hours (while off-peak either remains the same or increases only slightly) and lower overall electricity consumption. However, new estimates on the efficiency of TOU pricing relative to spot pricing suggest that well-designed pricing schemes perform relatively well in indicating relative price differences within days and provide relatively effective load-shifting incentives (Schittekatte et al., 2022).

Faruqui et al. (2020) report that while nearly four hundred TOU rates have been tested in pilots around the globe, full-scale deployment of TOU rates is quite limited, with only 4% of residential consumers being on TOU rates. Moreover, the evidence drawn from these field experiments faces some important limitations. Sample sizes are usually small, and the design of the programs, in which signing up for dynamic rates is optional, increases the potential for selection bias. Indeed, To overcome these concerns, Fowlie et al. (2021) partnered with a utility in California and implemented a large-scale randomized control trial to study the effects of opt-in vs. opt-out of dynamic pricing schemes. They find that households in the opt-in are more responsive than households in the opt-out group. Our case is a unique opportunity in that all households belonging to the regulated segment are put into the TOU tariff by default.

A potential concern of time-varying tariffs is their possible distributional impacts. With fixed tariff rates, households with a flatter consumption profile would cross-subsidize those with higher consumption at peak hours. The effect of transitioning to dynamic prices on low-income households is not clear ex-ante. Borenstein (2013) studies the effects of switching to critical peak pricing (CPP) and finds that, even assuming no demand response, low-income households face no significant changes in their electricity bills.[5] By combining substation

[5] Critical Peak Pricing, or CPP, combines flat rates with significant price increases, typically in a limited number of hours annually in which the electricity grid is under high stress, such as extremely hot summer days

data with demographics characteristics for a utility in Victoria (Australia), Leslie et al. (2021) find that areas with low housing prices, a high share of renters and elderly people are better off with RTP.

Nevertheless, Cahana et al. (2022) find that while low-income households react to hourly variation in prices within a day, regressive impacts arise from monthly variation, since low-income households have a larger share of electric heating during winter. Given that price differences are greater across months, they find a slightly negative overall effect. In the end, efficient tariffs compatible with distributionally equitable rates are a matter of design. Burger et al. (2019) propose a two-part tariff with a fixed charge based on income (or other correlated measures) that would mitigate the adverse effects of dynamic pricing while maintaining most of the efficiency gains.

Our methodological approach is based on machine learning techniques. These techniques have emerged as a powerful tool to build post-treatment counterfactuals, a key element in causal inference (Varian, 2016). In this paper, we closely follow the method used in Burlig et al. (2020) in which first, they use pre-treatment data and LASSO regressions to estimate electricity consumption, and second, generate in- and out-of-sample predictions to construct prediction errors that are used to identify the policy treatment effects.

3 Context and Data

There are five main distribution groups in the Spanish electricity market.[6] These five groups compete in each other's territories as retailers alongside new entrants. In each distribution area, only the vertically integrated retailer can offer the regulated tariff. This tariff is prescribed by law based on market conditions, and it is unique throughout the country. In contrast, non-regulated or commercial tariffs can be offered by any retailer across distribution areas, and the contract terms are freely set.

Despite improvements in competition since deregulation, the Spanish retail market is still highly concentrated. There are five main retailing firms, which belong to the same business groups as the five largest distribution companies. In 2019, 39% of consumers were still on the default regulated tariff, and an additional 42% were served by the commercial brand of the distribution company. Thus, over 80% of residential households are still served by mainly the five business groups. In addition, Enrich et al. (2022) find a significant incumbent advantage, with the probability of choosing the commercial brand of the distribution group higher on its own market, as well as strong consumer inertia.[7]

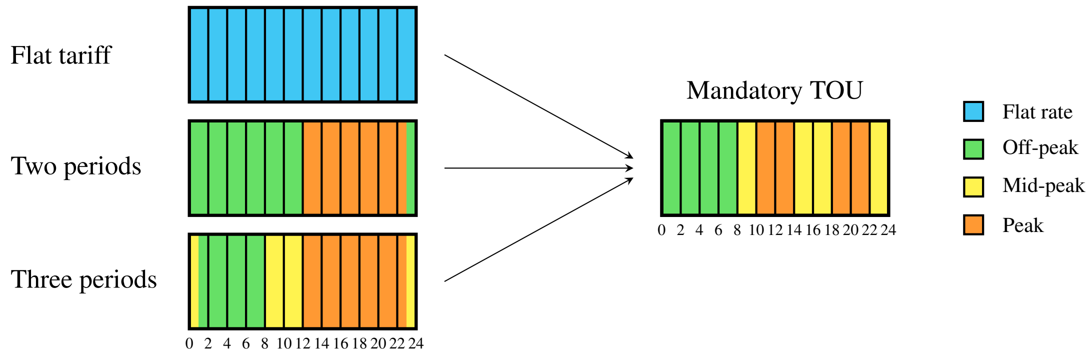
In recent years, the roll-out of smart meters has enabled the introduction of new and more flexible tariffs that can adjust better to the consumption profile of each household. The Spanish electricity bill has four components: energy costs, network charges, system charges, and taxes.[8] On average, system charges and energy costs represent the largest part of the bill (30% each), followed by network charges and taxes (each around 20%). In October 2015, households under the regulated tariff were put into an RTP scheme for the energy charge of their electricity bill. Despite this reform, Fabra et al. (2021) find that households did not respond to RTP, which

[6] The largest distribution groups are Endesa, Iberdrola, Naturgy, EDP, and Repsol. There are also a few small distribution companies that serve much smaller areas.

[7] Byrne et al. (2022) provide experimental evidence of variations in consumer willingness and ability to search and bargain, contributing to the different consumer bases and price distribution between incumbents and entrants.

[8] System charges include subsidies to renewable energies, deficit payments, and compensation of the extra cost of generation in non-peninsular regions.

Figure 1: Tariffs Before and After the Policy



may be attributed to the lack of consumer awareness, costly information acquisition, and small gains of demand response due to low price variation.

A significant change in the design of the tariffs was introduced on June 1, 2021. The policy change introduced mandatory TOU pricing in the regulated tariff, with system and network charges having three different tiers, depending on the hour of the day and the day of the week. According to the new regulation, peak hours ranged from 10 am to 2 pm and from 6 pm to 10 pm, and off-peak hours covered hours from 12 am to 8 am during working days. The remaining hours in the working days were considered mid-peak, while weekends and national holidays were fully off-peak.

In Figure 1, we represent on the left the different pricing schemes that consumers in the regulated tariff could choose from before the reform. Consumers were put into a flat rate and could opt-in to time-of-use charges in their electricity bill, facing 2 or 3 prices within a day. Before the policy, the share of consumers in the regulated tariff under TOU remained below 10%.[9] On the right-hand side, we show the new pricing scheme depending on the time of the day. Although the policy was compulsory for all residential contracts, retailers in the liberalized market could offer flat rates that compensate for the variation of the TOU tariff. Thus, we restrict our analysis to the regulated market.

Table 1 and Figure 2a show how prices in the regulated segment changed after the policy for each TOU period. We take energy costs and system and network charges, which constitute the largest part of the electricity bill. In Table 1, we compare the months affected by the policy (June to September 2021) to the same period between 2018 and 2020. Figure 2a shows the evolution of these components throughout the whole period. Total regulated prices increased by more than 60%, partly driven by rising prices in the wholesale market translated into higher energy costs. Importantly, increases in energy costs are independent of the hour of the day. If energy costs would have affected total price differently between TOU periods, our policy estimates of within-day shifts would have captured price changes external to the policy. Next, the reform increased average charges by nearly 50%. Still, it provided incentives for load shifting: prices during off-peak hours were 86% lower while consuming at peak hours was 200% more expensive after the policy. In Section 5.1, we use the variation induced by the policy to estimate price elasticities.

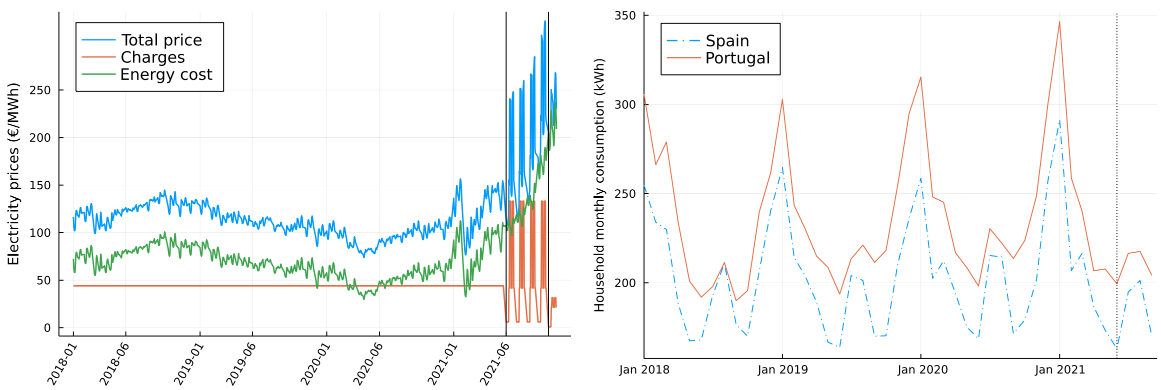
[9]In Portugal, nearly 90% of consumers in the regulated tariff pay flat rates, a share that has remained stable over our period of study, according to the Portugal Regulation Authority. Similar to Spain, consumers can choose between a flat rate, a two-period, and a three-period.

Table 1: Variation of the Energy Components Before and After the Policy

	Average	Off-Peak	Mid-Peak	Peak
Total Price	61.4%	12.2%	39.6%	117.5%
Charges	49.1%	-86.3%	-5.1%	199.4%
Energy Costs	68.6%	71.2%	65.4%	70.2%

Notes: Average prices are weighted by consumption in the regulated segment. We restrict to the months of the policy: June to mid-September since system charges were reduced on 15 September 2021 to cope with rising energy prices. Taxes are not included.

Figure 2: Electricity prices and demand



(a) Decomposition of retail electricity prices

(b) Electricity Demand per Capita

Notes: Figure 2a plots the average monthly price for each hour weighted by consumption in the Spanish regulated segment. Figure 2b plots the average household monthly consumption.

To estimate the causal effect of the policy, we compare the evolution of electricity consumption in Spain and its nearby country, Portugal, between January 2018 and September 14, 2021. In particular, we end our sample on 14 September since the Spanish authorities reduced system charges by 96% as a response to the rising energy prices. As a result, TOU tariffs were essentially canceled. Restricting the sample to September 2021 also avoids the most heated episodes of the energy crisis, which given how Spanish regulated prices are set, introduces substantial volatility.[10]

Consumption data for Spain come from the archives of the Spanish System Operator, *Red Eléctrica de España* (REE). These data measure the hourly demand of all the Programming Units that provide electricity into the national grid and are differentiated by regulated and non-regulated demand. The former only includes demand from residential households, whereas the latter includes domestic consumers, small- and medium-sized enterprises, and industrial consumers. This provides an additional reason for limiting our sample to the regulated segment, as we are only interested in households' demand response to the policy.

For the case of Portugal, electricity consumption data come from the Iberian electricity market operator, *Operador del Mercado Ibérico de Energía* (OMIE). We also restrict the sample to domestic consumers under the regulated tariff. Unfortunately, there is only one leading distribution company in Portugal, and therefore the

[10] As explained above, the energy component of the regulated price reflects day-ahead wholesale market prices.

data are reported at the aggregate national level.[11]

Since aggregate consumption under the regulated tariff in Spain is roughly ten times larger than in Portugal, we take per capita consumption as our variable of interest. Figure 2b shows the evolution of demand per capita for Spain and Portugal. We observe that electricity consumption follows a clear seasonal pattern over the months of the year. Even if Portuguese households in our sample present a relatively greater consumption, which widens in the winter, an important surge in Spanish consumption occurs yearly in the months where the policy is analyzed, mainly in the summer.

The information on the number of consumers in Spain has been kindly provided by the Spanish National Markets and Competition Commission (CNMC, according to the Spanish acronym) based on the Reports on Oversight of the Retail Electricity Market.[12] The number of consumers is presented quarterly for all retailers across all distributing areas in Spain. We interpolate the data to monthly frequency using a monotonic algorithm proposed by Fritsch and Butland (1984). We obtain Portugal's consumer data from the Bulletins of the Liberalized Electricity Market, which provide the number of consumers under the regulated and the liberalized segment by month.[13]

As in many liberalized markets, the number of consumers under the regulated tariff in Spain and Portugal has decreased in recent years. In particular, Portugal has set the end date for the regulated tariff to 2025. As we observe in Figure 3, in January 2018, the regulated tariff included over 11M consumers in Spain and over 1M in Portugal, representing respectively 41% and 20% of residential consumers. In the last year of our sample, the share of consumers in the regulated segment reached a record low at 34% in Spain and 15% in Portugal. However, during the months of the policy, the drop was steeper for Spain, suggesting that part of the consumers might have switched to the liberalized market in search of flat rates. By design, Spanish consumers with a regulated tariff are more affected by changes in wholesale prices, which may be a factor explaining the decline of consumers in the context of high price increases.[14] In contrast, the regulated tariff in Portugal is determined ex-ante, and its prices stayed relatively lower than the commercial prices, thus mitigating the impact of price increases on end-users. This may explain the consumer surge under the regulated tariff in October 2021.[15]

Finally, we obtained hourly temperature data from the Modern-Era Retrospective analysis for Research and Applications (MERRA-2) released by the NASA.[16] To match the temperature data at the 50x50km grid point-level with our consumption data, we weigh the temperature data by population in each distributing area.

In addition, we use Google Trends data to capture consumer awareness of the new tariffs. This public tool uses a largely unfiltered sample of actual search requests occurring in a given location during a certain period of time.[17] To account for all searches related to the policy, we use different wordings and synonyms that are closely related to the topic in Spanish. A limitation of the data is the scale on which search interest is defined because it is not only contingent on the generated sample but also the combination of regions and

[11] In continental Portugal, 99% of low-voltage consumers have their energy distributed by EDP Distribuçao (see here). This large distribution area is comparable in terms of size and consumers to the largest distribution companies in Spain: Endesa and Iberdrola.

[12] See IS Mercado Minorista de Gas y Electricidad.

[13] See Boletim do Mercado Liberalizado de Eletricidade.

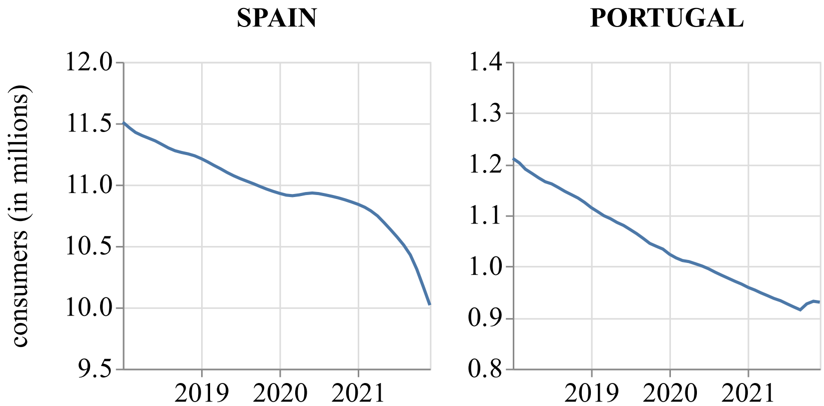
[14] If those consumers switching to a liberalized contract had different consumption patterns, our policy coefficients would be capturing these composition effects. Nonetheless, we can still compare policy effects for different TOU prices.

[15] Find more information on the Portuguese tariff on the regulator website.

[16] See GES DISC.

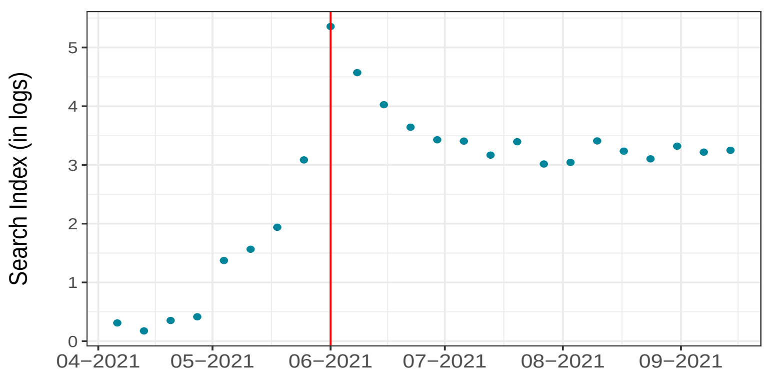
[17] To read more, visit Google Trends' Help Page.

Figure 3: Number of consumers



Notes: Evolution of consumers in the regulated market for Spain and Portugal. We interpolate quarterly data to monthly frequency.

Figure 4: Google trends' weekly search index



Notes: This figure shows the logarithm of the search index based on keywords related to the policy in Spain. The search index is constructed on a scale from 0 to 100, where 100 represents the maximum number of topic searches for the period of interest.

periods selected by the researcher. In particular, search interest is given on a scale from 0 to 100, where 100 represents the maximum number of topic searches for given regions and period selection. For our analysis, we have collected data spanning from January 2018 to October 2021. Figure 4 shows the logarithm of the weekly searches of keywords related to the policy. While we observe a modest increase in searching during the weeks previous to the introduction of the policy (marked with a red line), most searches are concentrated during the first week after the implementation. This increase in consumer awareness is in line with the results derived from a Household Panel survey conducted by the Spanish regulatory body (CMNC),[18] where the share of consumers declaring not knowing its own electricity contract decreased from 35% in the first semester of 2021 to 17% in the second semester, after several years remaining relatively constant.

In sum, to analyze the effect of the policy, we use demand data for the domestic regulated segment at the distribution firm level for Spain and the national level for Portugal. We drop the observations for 2020 due to the

[18] See Estadísticas panel de hogares.

disruption of COVID-19 to electricity demand. In total, we end up with a panel of 142,000 hourly observations from the five Spanish firms and Portugal, spanning from January 2018 to December 2019 and January 2021 to September 14, 2021. Summary statistics of the main variables are shown in Appendix Table 7.

4 Empirical strategy and Results

In this section, we describe the empirical strategy used to recover the effect of the policy on electricity consumption. Throughout the analysis, we will be comparing two different approaches: fixed effects vs machine learning. While machine learning techniques are not built to produce good parameter estimates, they can discover complex structures of the data that were not specified in advance (Mullainathan and Spiess, 2017), for example, by enabling us to select among a large set of covariates. In particular, we will closely follow Burlig et al. (2020), who use machine learning algorithms to estimate robust causal effects.

The main difference between the two approaches is that using machine learning tools and pre-treatment data, we create firm-specific models to predict electricity consumption in the post-treatment period. In the second step, we use these out-of-sample predictions to generate prediction errors that we incorporate as a dependent variable in the panel fixed effect model. On the other hand, the panel fixed effects model directly includes the observed electricity consumption as the dependent variable.

4.1 Panel fixed effects: Differences-in-Differences

To identify the potential demand response to the policy, we start by estimating the following Differences-in-Differences (DiD) regression:

$$y_{ith} = \beta_k D_{itk} + \delta_k P_{itk} + \gamma X_{ith} + \alpha_{ith} + \epsilon_{ith}, \tag{1}$$

where y_{ith} is the log of the demand per capita in the distribution area of firm i at day t and hour h. Note that we treat Portugal as a control distribution area. D_{itk} is a dummy variable that equals one for all Spanish distribution areas after the policy was implemented for each TOU tariff $k = \{\text{Off-Peak, Mid-Peak, Peak}\}$. Therefore, β_k captures the effect of the policy at different tariffs. P_{itk} stands for Placebo, and it takes one for all Spanish distribution areas one month before introducing the policy, thus capturing possible pre-trends unrelated to the policy or anticipatory effects that could cast doubts on the validity of the parallel trend assumption. Ultimately, even if the policy was officially announced in early 2021, the media coverage broadly started a week in advance, as evidenced by the Google Trends searches. Control variables X_{ith} include hourly temperature and an interaction for whether the temperature is above or below 20ºC. Finally, α_{ith} includes different combinations of fixed effects. We weigh observations by the number of consumers to make the sample representative and thus identify an average effect.

Table 2 reports the results of equation (1) under different fixed effects specifications. We only show the policy coefficients, but placebo coefficients can be found in Table 8 in the Appendix. Specification (1) includes firm-month-hour fixed effects. Specification (2) adds firm-year-hour fixed effects. In specification (3), we try an alternative approach to control for temporal patterns: month of sample fixed effects. Finally, in specification (4),

Table 2: Panel fixed effects: Differences-in-Differences

	ln(demand per capita)			
	(1)	(2)	(3)	(4)
Off-Peak	0.012	0.034	-0.025	0.000
	(0.063)	(0.035)	(0.042)	(0.031)
Mid-Peak	-0.016	-0.032	-0.055	-0.056
	(0.026)	(0.167)	(0.066)	(0.155)
Peak	-0.037	-0.083**	-0.083	-0.103**
	(0.150)	(0.029)	(0.132)	(0.032)
Firm-Month-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-TOU-Hour		Yes		Yes
Month of sample-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.943	0.948	0.951	0.953

Notes: This table presents estimates of Equation (1). Controls include hourly temperature and the interaction for whether the temperature is above or below 20ºC. Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

we present our preferred and most stringent specification, which includes all mentioned fixed effects. Basically, for each hour and TOU period, we control for differential patterns across distribution areas, months of the year, and different years, as well as for common shocks or time trends at the monthly level. We cluster standard errors at the firm-month level.[19] We observe that the sign of the coefficients for off-peak hours change across specifications, but none of them is significant. Similarly, coefficients for mid-peak hours, while negative, fail to pass the significance test.

In contrast, some of the results for peak hours are significant and negative. In our preferred specification in column (4), we observe that households do not change off-peak nor mid-peak consumption, but reduce peak consumption by 10.3%. These results suggest that households reduce overall consumption in response to TOU prices, without any sign of load-shifting to off-peak hours. Table 8 in the Appendix presents the extended regression results, including the coefficients for the placebo test. Consumption in mid- and peak hours does not show any particular pre-trend before the introduction of the policy, although we do observe an increase in consumption during off-peak hours. It is possible that households started experimenting with new habits on when to consume electricity without shifting consumption from future peak hours.

However, when comparing these different pricing regimes, we combine off-peak hours during the day with all hours during the weekend. Thus, given that part of the consumption cannot be shifted between days, we next split the effect by differentiating between weekdays and weekends by estimating equation (2). The goal of this analysis is threefold. First, as already mentioned, we want to identify changing consumption patterns within a day. Second, we define mid- and peak-hours during the weekend, thus identifying possible effects of the policy that could be related to habit formation. Third, methodologically, if placebo tests exhibit the same

[19] A good rule of thumb is to cluster at the level of treatment assignment. See Abadie et al. (2022) and Roth et al. (2022) for a discussion on the recent developments in the DiD literature.

patterns, we may be able to reduce the bias of our estimates due to a violation of the parallel trend assumption by computing the additional difference of the policy on weekdays with respect to weekends. For the sake of exposition, we define this model as the triple-differences (TD) model. In most cases, we estimate the effect for both groups (weekdays and weekends), instead of the traditional approach of identifying treatment effects as the difference in treatment between groups. Conceptually, this relates to whether we understand the treatment effects on weekends as true policy effects or as a control for unobserved confounders. The estimation equation becomes

$$y_{ith} = \sum_{w \in [0,1]} \beta_{kw} D_{itk} 1[\text{weekday} = w] + \sum_{w \in [0,1]} \delta_{kw} P_{itk} 1[\text{weekday} = w] + \gamma \text{temp}_{ith} + \alpha_{ithw} + \epsilon_{ith}. \tag{2}$$

Table 3 reports the results of the triple-differences model.[20] Even though under the new TOU tariffs, weekend consumption is subject to off-peak prices, we observe a significant demand response for all three periods in some specifications. These findings align with Fowlie et al. (2021), where consumers under CPP reduce their consumption on event and non-event days, pointing to habit formation. Indeed, the null effect on off-peak hours found in Table 2 was an average effect of first, increases during off-peak hours in both, weekdays and weekends, and second, significant consumption cuts for mid- and peak hours during weekends. In fact, during mid-peak hours, results suggest that consumers do not seem to differentiate between weekdays and weekends. Finally, looking again at column (4), results for mid-peak and peak hours during weekdays are similar to the ones found in the DiD model (as they should be given that policy dummies are defined in the same way). However, in our setting, we cannot conclude that neither of those is different from zero.

Figure 5 plots the hourly policy and placebo coefficients for weekends and weekdays. The placebo suggests that we have some unobserved confounding factors during off-peak hours. The results are consistent with a significant policy effect during mid- and peak-hours that is not present in the Placebo, as reported in Table 3. The response is statistically similar across weekdays and weekends, suggesting that households respond to the policy implementation without differentiating weekday and weekend behavior.[21]

4.2 Machine Learning

In this section, we repeat the same analysis but with a machine-learning estimator. This approach allows the researcher to include all possible interactions between controls, making the model selection process transparent and data-driven. To estimate the effect of the policy, we will proceed in two steps. First, we create a firm-specific model of electricity consumption. For that, consider the following standard regression:

$$Y_{iht} = \gamma_{ih} X_{iht} + \epsilon_{iht}, \tag{3}$$

where Y_{iht} is the electricity consumption per capita at distribution area i at hour h in period t, and X_{iht} is a set of control variables. This vector of covariates includes month, weekend, and national holiday dummies,

[20] Again, we only display policy coefficients, but placebos can be found in Table 9 in the Appendix, where we also show the additional effects of the policy during weekdays with respect to weekends.

[21] Table 9 in the Appendix shows the coefficients and standard error for the difference between weekends and weekdays. While the weekday response is larger in magnitude, the two are not statistically different.

Table 3: Panel fixed effects: triple-differences

	ln(demand per capita)			
	(1)	(2)	(3)	(4)
Policy Weekend				
Off-Peak	0.016	0.082	-0.002	0.032*
	(0.039)	(0.075)	(0.030)	(0.019)
Mid-Peak	-0.005	-0.025	-0.067	-0.053***
	(0.017)	(0.027)	(0.078)	(0.009)
Peak	-0.007	-0.045	-0.083*	-0.075***
	(0.090)	(0.031)	(0.043)	(0.009)
Policy Week				
Off-Peak	0.024	0.075**	0.005	0.037
	(0.089)	(0.025)	(0.066)	(0.055)
Mid-Peak	-0.015	-0.033	-0.054***	-0.054
	(0.040)	(0.090)	(0.009)	(0.075)
Peak	-0.037***	-0.086**	-0.082***	-0.101
	(0.004)	(0.033)	(0.011)	(0.084)
Firm-Month-Weekend-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-Weekend-TOU-Hour		Yes		Yes
Month of sample-Weekend-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.949	0.954	0.957	0.960

Notes: This table presents estimates of Equation (2). Controls include hourly temperature and the interaction for whether the temperature is above or below 20ºC Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

temperature, the minimum and maximum temperature of the day, and all possible interactions between these variables. In practice, we estimate a separate regression for each firm-hour unit, so effectively, each variable is interacted with firm and hour fixed-effects. Importantly, we only use pre-treatment observations to estimate Equation (3). We treat Portugal as another firm unit (and the only non-treated unit).

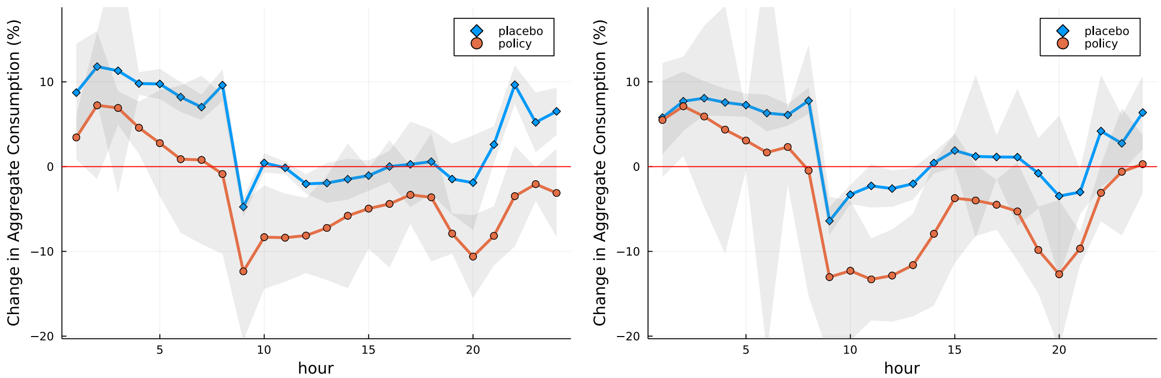
We estimate regression 3 using the LASSO regularization technique to choose among all possible predictors with nonzero significant values. Note that these predictors may vary between hours of the day or firms.[22] We then use these models to generate in-sample and out-of-sample predictions of electricity consumption per capita in the post-treatment period. We compute prediction errors by comparing predictions with the actual outcome.

Figure 11 in Appendix B shows the fraction of models for which a variable is selected. We find that most models include the intercept. National holidays and weekends also explain electricity consumption in both, Spain and Portugal. The maximum daily temperature appears to be even more important than the hourly temperature. As for months, Figure 11 suggests that Portugal has greater seasonal effects, with monthly dummies being an important factor for most hours of the day.

In the second step, we use the prediction errors as the dependent variable in the DiD and TD model presented

[22] We choose the level of regularization through cross-validation, where we partition the data into ten subsamples. We follow the commonly used "one standard error" rule to select the parameters with the best performance.

Figure 5: Panel fixed effects: Hourly Policy effects by type of day



(a) Weekends and Holidays

(b) Weekdays

Notes: These figures present hourly estimates of Equation (2). Observations are weighted by the number of consumers in each distribution area. Controls include temperature, the interaction for whether the temperature is above or below 20ºC, and firm-month-hour, firm-year-hour, and month-of-sample fixed effects. Standard errors clustered at the firm-month level and confidence intervals.

in Section 4.1. Figure 6 plots the prediction errors for Spain and Portugal, with a break for the year 2020. The red line marks the date of the reform. Before its implementation, residuals are centered around zero, meaning that in-sample predictions generally approximate the observed data.[23] On the other hand, there is a decoupling in prediction errors between Spain and Portugal after June 2021, with prediction errors for Spanish firms turning negative, meaning that the predicted consumption overstates the realized one.[24] Intuitively, this should be related to the effect of the policy that we want to identify. Note that the standard identification assumptions of the DiD model have to hold for prediction errors rather than in electricity consumption directly.[25]

Table 4 presents the results for the DiD model of Equation (1) using prediction errors as the dependent variable. We only show the policy coefficients. Placebo coefficients can be found in Table 10 in the Appendix. Focusing on our preferred specification in Column (4), mid-peak and peak coefficients are -7.8% and -12.6%, respectively, slightly larger than the ones found using the standard panel fixed-effect regression. However, we observe a significant reduction in off-peak consumption. Recall that this coefficient is an average of changes in consumption during weekdays off-peak hours, and weekends, and thus this coefficient may be explained by household behavior during weekends.

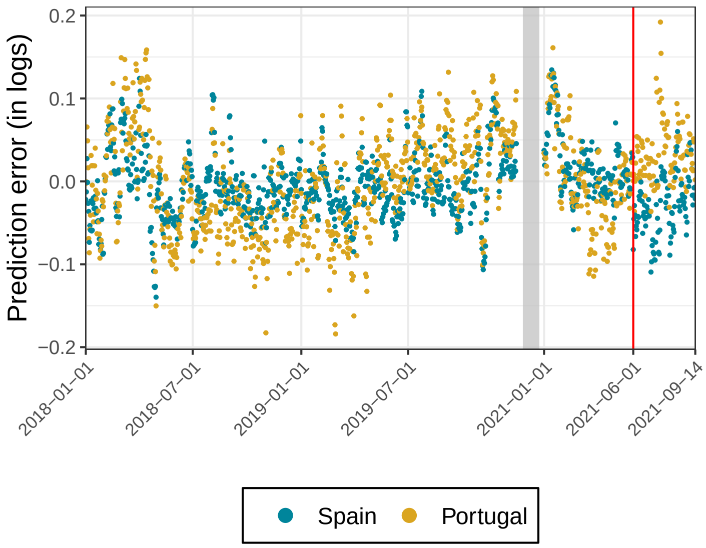
The results do not appear to be more stable across different specifications. A possible explanation is that having only six aggregated models (five distribution companies and Portugal) with the same treatment date,

[23]For Spain, the plotted residuals in Figure 6 show the average among distribution areas, which lowers its variance with respect to Portugal.

[24]We also considered a random forest approach instead of the LASSO regression for the first step. Figure 12 plots the difference in prediction errors between the two methods for Spain and Portugal, and Figure 11c shows the most relevant variables. Even though the random forest approach seems to provide better in-sample predictions, especially in capturing seasonal patterns, regression results in step two are more stable under LASSO, suggesting that these patterns were then captured in the panel fixed effect regression. Table 12 shows these results. Note that the random forest approach is unable to capture pre-trends, evidenced by the significance of multiple placebo coefficients.

[25]Figure 10 shows the individual prediction for each firm and Portugal. The bigger differences between out-of-sample predictions and actual consumption are found for the EDP and Naturgy regions. In contrast, for Endesa and Repsol, predictions reasonably approximate the true evolution of electricity consumption. We can match these differences with the policy effects shown in Figures 9a and 9b derived from estimating the Panel FE and Machine learning models for each distribution area. In all cases, the biggest effects are found in the EDP, Naturgy, and Iberdrola regions, while for Endesa and Repsol, the results are less conclusive.

Figure 6: Machine Learning: Prediction errors



Notes: This figure plots daily prediction errors, defined as the difference between the log of the demand per capita and the log of the prediction of our LASSO model. For Spain, the prediction error is the average over distribution groups.

Table 4: Machine learning: Differences-in-Differences

	Prediction error			
	(1)	(2)	(3)	(4)
Off-Peak	0.006	0.022**	-0.029***	-0.016**
	(0.008)	(0.011)	(0.007)	(0.006)
Mid-Peak	-0.023*	-0.047	-0.064***	-0.078***
	(0.012)	(0.030)	(0.007)	(0.017)
Peak	-0.042***	-0.094***	-0.095***	-0.126***
	(0.005)	(0.008)	(0.007)	(0.018)
Firm-Month-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-TOU-Hour		Yes		Yes
Month of sample-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.044	0.129	0.199	0.242

Notes: This table presents estimates of Equation (1). Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

there is limited heterogeneity between treated and control units.[26]

In Table 5, we split the results between weekdays and weekends. Again, results for mid-peak and peak mirror those of the DiD model, but now the results stay statistically significant, not as in the standard panel fixed-effect model. Also, once we decompose off-peak hours, we find a significant intra-day load shifting on weekdays, with an average increase in consumption of 1.7% during off-peak hours and a 12.4% decrease during peak hours. We still find significant effects during the weekends. All in all, using a machine learning algorithm in the first step allows the researcher to reduce noise in the dependent variable, explaining the smaller standard errors in our second step.

For the standard errors to be comparable, one should account for the noise generated in the first step of the estimation. To assess the relevance of these concerns, we implement a two-step bootstrap procedure. In the first step, we sample with replacement months of the sample during the pre-treatment period for each firm. We then use these data to estimate the firm-specific LASSO model and obtain predictions for all months in the original data (this is, all pre- and post-treatment periods). We repeat this step twenty times to have twenty alternative predictions for each firm. In the second step, we draw one prediction for each firm, compute the prediction error and estimate the panel fixed-effect model. We repeat this process a thousand times.[27] Figure 13 in the Appendix shows the results of this exercise for each group of coefficients. It is reassuring that the mean of the distribution of the bootstrapped coefficients converges to the triple-differences coefficients estimated in the main analysis, suggesting that our results are not driven by a particular sample during the estimation of the machine learning method.

Figure 7 plots the hourly coefficients of the policy and placebo effects. The overall hourly change in consumption due to the policy follows the same pattern that we found using the standard fixed effect panel regression, but the standard errors are more precise. Indeed, the fact that hourly placebo tests are now significant, especially during off-peak hours, can lead to the conclusion that the parallel trend assumption does not hold. Therefore, one should consider these unobserved confounders by estimating the additional policy effects during weekdays with respect to weekends with a formal triple-differences model.

However, this approach may be too conservative for two reasons. First, placebo tests differ only from zero during off-peak hours, so we can still attribute mid-peak and peak coefficients to the policy. Indeed, Table 11 shows placebo coefficients grouped by TOU tariffs, and almost all of them fail to pass the significance test. Second, in identifying only additional effects during weekdays with respect to weekends, we are implicitly shutting down effects other than the price effects of the policy, such as habit formation. Taking this scenario as a lower bound, we still find a decrease in consumption during peak hours of 4.5%, although the coefficient is not statistically significant anymore.

[26] This is, it is plausible that this setting meets all the requirements and identifying assumptions for the two-way fixed effects estimator to produce consistent estimates. Thus, the potential gains for improving the estimator may not be large enough (Roth et al., 2022).

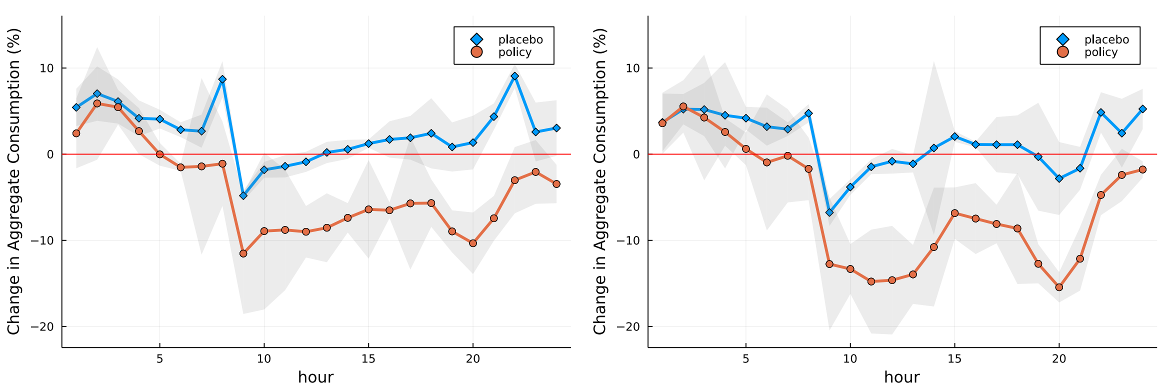
[27] Usually, at this stage, one would sample at the level of the events, estimate Equation (2), and get a distribution of policy coefficients. The standard deviation of this distribution would be the analogue to the clustered standard errors computed in the main analysis. This is the approach followed in Patnaik et al. (2013) based on Davison et al. (1986). Unfortunately, our treatment consists of only one event, which makes this event study bootstrapping strategy not applicable.

Table 5: Machine Learning: triple-differences

	Prediction error			
	(1)	(2)	(3)	(4)
Policy Weekend				
Off-Peak	0.010	0.065***	-0.007	0.015*
	(0.010)	(0.005)	(0.009)	(0.008)
Mid-Peak	-0.011	-0.035	-0.073***	-0.063**
	(0.007)	(0.034)	(0.007)	(0.023)
Peak	-0.010	-0.050*	-0.084***	-0.079***
	(0.017)	(0.028)	(0.019)	(0.021)
Policy Week				
Off-Peak	0.017*	0.059**	0.003	0.017***
	(0.009)	(0.023)	(0.007)	(0.004)
Mid-Peak	-0.023**	-0.047*	-0.064***	-0.077***
	(0.011)	(0.025)	(0.004)	(0.015)
Peak	-0.042***	-0.094**	-0.096***	-0.124***
	(0.009)	(0.036)	(0.009)	(0.023)
Firm-Month-Weekend-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-Weekend-TOU-Hour		Yes		Yes
Month of sample-Weekend-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.049	0.142	0.202	0.243

Notes: This table presents estimates of Equation (2). Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

Figure 7: Machine Learning: Hourly Policy effects by type of day



(a) Weekends and Holidays

(b) Weekdays

Notes: These figures present hourly estimates of Equation (2). Observations are weighted by the number of consumers in each distribution area. We include firm-month-hour, firm-year-hour, and month-of-sample fixed effects. Standard errors clustered at the firm-month level and confidence intervals.

5 Extensions

5.1 Price Elasticities

This section uses the exogeneity of the policy to estimate household price elasticities. Albeit far from a formal demand estimation analysis, we believe that it can improve the comparability of the results to previous studies, and, therefore, may be helpful for policy evaluation and guidance. For that, we consider a modified version of the DiD model presented in Equation (1) substituting the policy indicator with actual electricity prices:

$$y_{ith} = \eta p_{ith} + \gamma X_{ith} + \alpha_{ith} + \epsilon_{ith}, \tag{4}$$

where y_{ith} is the log of the demand per capita in the distribution area of firm i at time t and hour h, and p_{iht} is the log of the electricity price. Thus, η can be interpreted as the price elasticity.

Estimating Equation (4) for only the treatment group would be problematic because the price coefficient is biased in a context of simultaneous equations where prices and quantities are jointly determined in equilibrium. However, by including Portugal data as our control group and using the policy as an instrument for electricity prices, we can break the simultaneity problem. We instrument the logarithm of the price with the logarithm of the charges component affected by the new TOU regime to get a direct interpretation of the effect and performance of our instrument. Moreover, given that the instrument is used to construct the total price, it is relevant by definition.[28]

The Spanish electricity prices come from the archives of the Spanish System Operator, *Red Eléctrica de España* (REE). See Table 1 and Figure 2a for summary statistics. Traditionally, regulated prices in Portugal were decided for the year ahead. However, given the recent increases in wholesale prices, regulated tariffs kept increasing during 2020 and 2021.[29] Nonetheless, these prices were only updated monthly. Thus, they add limited variation given that our regressions include month-of-sample fixed effects. To estimate Equation (4), we control for hourly temperature as well as an interaction for whether the temperature is above or below 20ºC. Also, we include firm-month-hour, firm-year-hour, and month-of-sample fixed effects, which was our preferred specification in analyzing the effect of the policy.

Table 6 presents the demand parameters estimated using Equation (4) for the two dependent variables studied in the main analysis: demand per capita (in logs) and the prediction error resulting from subtracting the log of the prediction of our LASSO model to the log of the demand per capita. In both cases, we compare OLS estimates with IV estimates using the policy as an instrument. For both outcome variables, OLS estimates do not show any significant demand response, a common result of the simultaneity issue in this framework. However, once the price is instrumented, we observe significant price elasticities ranging between -0.03 and -0.04. These estimates align with the literature on electricity demand estimation, where consumers are usually found to be relatively insensitive to price variations.

We can use these estimates to compute back-of-the-envelope demand responses given price changes and, thus, connect these results with our policy estimates. For that, note that the policy effects of a given TOU

[28] Another problem in estimating Equation (4) is that these estimates do not take into account intra-day substitution patterns (cross-elasticities), where consumption in hour h could be affected by a price of a different hour of the day. Given the limited price variation, we only estimate the own elasticity.

[29] See the Portuguese regulator website.

Table 6: Elasticity estimates

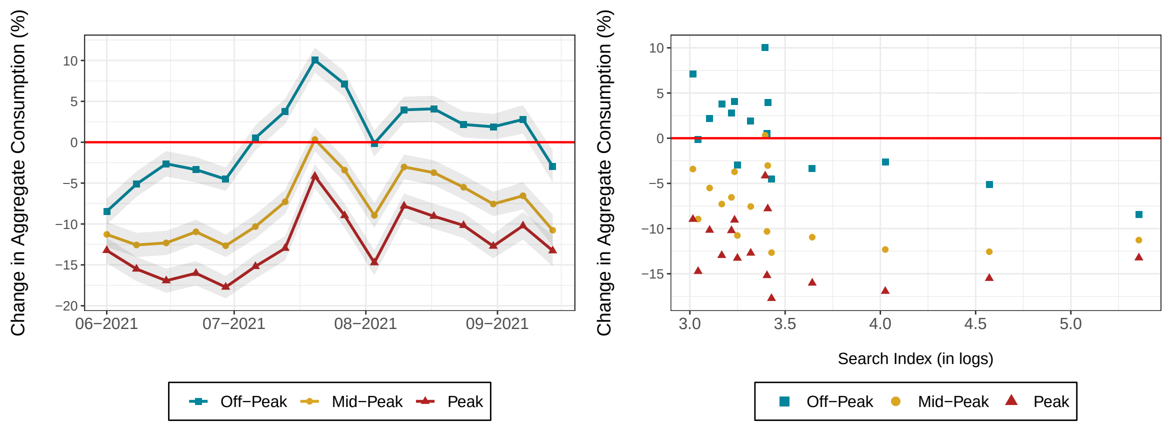
	Demand per capita		Prediction error	
	OLS	IV	OLS	IV
Electricity price	0.033	-0.043***	0.022	-0.031***
	(0.123)	(0.011)	(0.016)	(0.011)
First stage				
TOU tariff		0.238***		0.238***
		(0.020)		(0.020)
Temp Control	Yes	Yes	Yes	Yes
Firm-Month-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-TOU-Hour	Yes	Yes	Yes	Yes
Month of sample-TOU-Hour	Yes	Yes	Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.944	0.943	0.275	0.261

Notes: This table presents estimates of Equation (4). All variables are in log-form. Prediction errors are the difference between the observed demand per capita (in logs) and the predicted one (in logs). Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

tariff also captured increases in the energy cost component of the price during the months the policy was in place. However, given that those increases were constant between hours, we can still use the difference in policy effects to identify changes in the price component affected by the new TOU rates. For that, recall from Figure 2a shows that TOU rates changed from a constant price of 44€/MWh to 6, 42, and 133€/MWh for off-, mid-, and peak hours, respectively. In what follows, we take the elasticity using prediction errors, and the policy effects during weekdays are also estimated using prediction errors (see column (4) in Table 5). According to the observed price changes and the estimated elasticity, electricity consumption should have decreased by 2.5% and 8.9% between off-peak hours and mid- and peak hours, respectively. These differences are significantly lower than the change in consumption implied by our policy estimates, with a difference in consumption of 9.4% and 14.1% derived from subtracting the effect of the policy during off-peak hours from the effect during mid-and peak hours.

One possible explanation for these differences is that the response to variations in TOU tariffs is stronger than the response to changes in energy costs. In contrast to changes in real-time prices, pre-determined tariffs can improve consumer awareness and, therefore, trigger higher responses. One limitation of this analysis is that we cannot separately estimate an elasticity for weekends, since we do not have enough variation in our instrument. *A priori*, however, it is difficult to guess the magnitude of this elasticity. On the one hand, consumers could appear more price elastic and better adjust their daily consumption (e.g., allocate activities such as clothes washing or cooking to low-price hours). On the other hand, our main results point to some form of habit formation, with consumers following the same weekday behavior during weekends. In that case, price changes are a poor predictor of consumer demand response.

Figure 8: Google Trends and Policy Effects



(a) Policy effects by week

(b) Google Searches and Policy effects

Notes: Figure 8a presents estimates of Equation (2) aggregating the treatment dummies at the TOU tariff level and interacting them with week-of-sample fixed effects. Observations are weighted by the number of consumers in each distribution area. We include firm-month-hour, firm-year-hour, and month-of-sample fixed effects. Standard errors clustered at the firm-month level and confidence intervals.

5.2 Google Trends

In this section, we present an exploratory analysis of the relationship between Google searches and the effect of the policy. As shown in Figure 4, there was only a modest increase in searching during the weeks previous to the introduction of the policy, while most searches were concentrated during the first week, returning to initial levels afterward. It is this one-time shock nature and the subsequent lack of variation that did not allow us to proceed with a more formal analysis to establish some sort of causality between searching behavior and consumer response.

To study the relationship between search and consumption behavior, Figure 8a shows the weekly effect of the policy since its introduction by TOU tariff, estimated using Equation (1) on weekdays consumption data with prediction errors as the outcome variable. Figure 8b combines both graphs by plotting the relation between searches and policy effects. We find the following results. First, during the first week, there is a generalized decrease in consumption throughout all hours of the day that, combined with the significant amount in search, could imply that households were not sure yet about the functioning of the new system and preferred to hedge against possible bill increases. Then, during the following weeks, there is a negative relation between search and consumption patterns, especially reflected in an increase in consumption during off-peak hours by the end of June with respect to periods previous to the policy. Even though it is not possible to draw causal conclusions from this relationship, it could be that during the first month, households were adapting their consumption patterns. Finally, from mid-July, both the amount of searching and the effect of the policy remained constant after that. These estimates are consistent with the end of the adaptation process and a permanent change in household behavior.

6 Conclusions

We study the effects of a TOU pricing program for residential electricity demand introduced in Spain in June 2021. Under the new regulation, system and network charges (accounting for 50% of the overall electricity bill) had three tiers during weekdays, while weekends were fully off-peak. To identify the causal impact of the policy, we estimate a Differences-in-Differences model using Portugal to control for cross-sectional and temporal confounders. We compare two empirical models, the first being the standard fixed effects panel model. Using this method, we find a significant reduction in consumption of 10.3% for peak hours on average, while we do not find any load-shifting to hours with lower prices. We then split off-peak hours between weekdays and weekends. We observe a significant demand response during weekends for all three periods, pointing to some sort of habit formation.

We then turn to machine learning techniques for variable selection, given the richness of fixed effects in this setting. Surprisingly, we do not find evidence that using machine learning techniques helps to reduce the sensitivity of the estimates across different specifications. A possible explanation for the sensitivity of this approach could come from the lack of heterogeneity of treated units and treatment dates. Nonetheless, the algorithm helps to reduce the noise in the dependent variable, leading to more precise estimates. In particular, during weekdays, consumption increased by 1.7% during off-peak hours, while it decreased by 7.7%, and 12.4% during mid- and peak hours, respectively.

This paper contributes to the discussion on whether dynamic pricing schemes are an effective tool to change consumer behavior and help to reduce overall energy consumption. We find that salience can be a crucial factor in driving consumer response and that, while a system with frequent price changes may need to be accompanied by a process of automation, we find that a TOU pricing scheme can have significant effects given its foreseeable nature and thus, forming new consumption habits. All in all, economic incentives can be an effective tool, especially when the long-run effects such as habit formation are taken into account.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2022). When Should You Adjust Standard Errors for Clustering?*. *The Quarterly Journal of Economics 138*(1), 1–35. 12

Allcott, H. (2011). Rethinking real-time electricity pricing. *Resource and Energy Economics 33*(4), 820–842. Special section: Sustainable Resource Use and Economic Dynamics. 5

Boiteux, M. (1960). Peak-load pricing. *The Journal of Business 33*(2), 157–179. 4

Bollinger, B. K. and W. R. Hartmann (2020). Information vs. automation and implications for dynamic pricing. *Management Science 66*(1), 290–314. 2, 5

Borenstein, S. (2002). The trouble with electricity markets: Understanding california's restructuring disaster. *Journal of Economic Perspectives 16*(1), 191–211. 4

Borenstein, S. (2005). The long-run efficiency of real-time electricity pricing. *The Energy Journal 26*, 93–116. 4, 5

Borenstein, S. (2013). Effective and equitable adoption of opt-in residential dynamic electricity pricing. *Review of Industrial Organization 42*, 127–160. 5

Borenstein, S. and S. Holland (2005). On the efficiency of competitive electricity markets with time-invariant retail prices. *RAND Journal of Economics 36*, 469–493. 4

Burger, S., C. Knittel, I. Perez-Arriaga, I. Schneider, and F. Vom Scheidt (2019). The efficiency and distributional effects of alternative residential electricity rate designs. *The Energy Journal 41*. 6

Burlig, F., C. Knittel, D. Rapson, M. Reguant, and C. Wolfram (2020). Machine learning from schools about energy efficiency. *Journal of the Association of Environmental and Resource Economists 7*, 1181–1217. 3, 6, 11

Byrne, D. P., L. A. Martin, and J. S. Nah (2022, 04). Price Discrimination by Negotiation: a Field Experiment in Retail Electricity. *The Quarterly Journal of Economics 137*(4), 2499–2537. 6

Cahana, M., N. Fabra, M. Reguant, and J. Wang (2022). The Distributional Impacts of Real-Time Pricing. CEPR Discussion Papers 17200, C.E.P.R. Discussion Papers. 6

Davison, A. C., D. V. Hinkley, and E. Schechtman (1986). Efficient bootstrap simulation. *Biometrika 73*(3), 555–566. 17

Elmaghraby, W. and P. Keskinocak (2003). Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science 49*, 1287–1309. 4

Enrich, J., R. Li, A. Mizrahi, and M. Reguant (2022). Smart Meters and Retail Competition: Trends and Challenges. *AEA Papers and Proceedings 112*, 461–465. 6

Fabra, N., M. Reguant, D. Rapson, and J. Wang (2021). Estimating the Elasticity to Real Time Pricing: Evidence from the Spanish Electricity Market. *AEA Papers and Proceedings, American Economic Association 111*, 425–29. 2, 5, 6

Faruqui, A., S. Sergici, and L. Lam (2020). Bridging the chasm between pilots and full-scale deployment of time-of-use rates. *The Electricity Journal 33*(10). 5

Fowlie, M., C. Wolfram, P. Baylis, C. Spurlock, A. Todd-Blick, and P. Cappers (2021). Default effects and follow-on behaviour: Evidence from an electricity pricing program. *Review of Economic Studies 88*(6), 2886–2934. 5, 13

Fritsch, F. N. and J. Butland (1984). A method for constructing local monotone piecewise cubic interpolants. *SIAM Journal on Scientific and Statistical Computing 5*(2), 300–304. 9

Gibbs, C., D. Guttentag, U. Gretzel, L. Yao, and J. Morton (2017). Use of dynamic pricing strategies by airbnb hosts. *International Journal of Contemporary Hospitality Management 30*, 2–60. 4

Holland, S. and E. Mansur (2008). Is real-time pricing green? the environmental impacts of electricity demand variance. *The Review of Economics and Statistics 90*(3), 550–561. 4

Holland, S. P. and E. T. Mansur (2006). The short-run effects of time-varying prices in competitive electricity markets. *The Energy Journal 27*(4), 127–155. 4

Jessoe, K. and D. Rapson (2014). Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review 104*(4), 1417–38. 5

Joskow, P. L. and C. D. Wolfram (2012). Dynamic pricing of electricity. *American Economic Review 102*(3), 381–85. 2, 4

Leslie, G., A. Pourkhanali, and G. Roger (2021). Can real-time pricing be progressive? identifying cross-subsidies under fixed-rate electricity tariffs. 6

McAfee, R. P. and V. Te Velde (2006). Dynamic pricing in the airline industry. *Handbook on economics and information systems 1*, 527–67. 4

Mullainathan, S. and J. Spiess (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106. 11

Patnaik, I., A. Shah, and N. Singh (2013). Foreign Investors under Stress: Evidence From India. *International Finance 16*(2), 213–244. 17

Poletti, S. and J. Wright (2020). Real-time pricing and imperfect competition in electricity markets*. *The Journal of Industrial Economics 68*(1), 93–135. 4

Roth, J., P. H. C. Sant'Anna, A. Bilinski, and J. Poe (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Conditionally Accepted, Journal of Econometrics*. 12, 17

Schittekatte, T., D. S. Mallapragada, P. L. Joskow, and R. Schmalensee (2022). Electricity retail rate design in a decarbonized economy: An analysis of time-of-use and critical peak pricing. Working Paper 30560, National Bureau of Economic Research. 5

Steiner, P. O. (1957). Peak Loads and Efficient Pricing*. *The Quarterly Journal of Economics 71*(4), 585–610. 4

Varian, H. (2016). Causal inference in economics and marketing. *Proceedings of the National Academy of Sciences 113*, 7310–7315. 6

Williamson, O. E. (1966). Peak-load pricing and optimal capacity under indivisibility constraints. *The American Economic Review 56*(4), 810–827. 4

Wolak, F. A. (2019). The Role of Efficient Pricing in Enabling a Low-Carbon Electricity Sector. *Economics of Energy & Environmental Policy 8*(2). 2

Woo, C., R. Li, A. Shiu, and I. Horowitz (2013). Residential winter kwh responsiveness under optional time-varying pricing in british columbia. *Applied Energy 108*, 288–297. 5

7 Appendix

A Tables

Table 7: Summary statistics

			Spain			
variable	*units*	*mean*	*st. dev.*	*minimum*	*median*	*maximum*
demand	MWh	3064.91	909.76	1466.20	2983.40	7029.90
consumer	Million	11.07	0.28	10.43	11.10	11.51
demand per capita	Wh/cons.	276.65	81.13	137.13	270.40	648.56
temperature	Celsius	15.84	7.18	-1.30	15.16	35.21
high temperature	Binary	0.30	0.46	0.00	0.00	1.00
			Portugal			
variable	*units*	*mean*	*st. dev.*	*minimum*	*median*	*maximum*
demand	MWh	344.50	97.94	160.60	316.10	816.30
consumer	Million	1.07	0.09	0.92	1.09	1.21
demand per capita	Wh/cons.	320.95	85.62	164.02	297.45	820.30
temperature	Celsius	16.95	8.48	-3.72	15.89	44.63
high temperature	Binary	0.33	0.47	0.00	0.00	1.00

Notes: Sample between January 2018 and September 14, 2021, excluding 2020. The unit of observation is an hour-distribution area. There are five distribution areas in Spain and one distribution area in Portugal. $N = 142,000$.

Table 8: Panel fixed effects: Differences-in-Differences

	ln(demand per capita)			
	(1)	(2)	(3)	(4)
Policy				
Off-Peak	0.012	0.034	-0.025	0.000
	(0.063)	(0.035)	(0.042)	(0.031)
Mid-Peak	-0.016	-0.032	-0.055	-0.056
	(0.026)	(0.167)	(0.066)	(0.155)
Peak	-0.037	-0.083**	-0.083	-0.103**
	(0.150)	(0.029)	(0.132)	(0.032)
Placebo				
Off-Peak	0.013	0.034*	0.022	0.047**
	(0.045)	(0.020)	(0.036)	(0.021)
Mid-Peak	0.034***	0.018	0.009	0.008
	(0.003)	(0.033)	(0.007)	(0.024)
Peak	0.052***	0.006	0.014	-0.006
	(0.005)	(0.022)	(0.010)	(0.011)
Firm-Month-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-TOU-Hour		Yes		Yes
Month of sample-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.943	0.948	0.952	0.954

Notes: This table presents estimates of Equation (1). Controls include hourly temperature and the interaction for whether the temperature is above or below 20ºC. Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

Table 9: Panel fixed effects: triple-differences

	ln(demand per capita)			
	(1)	(2)	(3)	(4)
Policy Weekend				
Off-Peak	0.016	0.082	-0.002	0.032
	(0.011)	(0.069)	(2.305)	(0.032)
Mid-Peak	-0.005	-0.025	-0.067	-0.053**
	(0.031)	(0.032)	(1.335)	(0.025)
Peak	-0.007	-0.045	-0.083	-0.075**
	(0.046)	(0.072)	(2.491)	(0.024)
∆ Policy Week				
Off-Peak	0.009	-0.006	0.007	0.005
	(0.024)	(0.081)	(1.122)	(0.022)
Mid-Peak	-0.010	-0.008	0.013	-0.001
	(0.038)	(0.093)	(1.884)	(0.056)
Peak	-0.029	-0.041	0.001	-0.027
	(0.105)	(0.027)	(5.604)	(0.121)
Placebo Weekend				
Off-Peak	0.008	0.074	0.061	0.095***
	(0.021)	(0.051)	(1.197)	(0.020)
Mid-Peak	0.045	0.025	-0.005	0.009
	(0.066)	(0.081)	(1.779)	(0.067)
Peak	0.061	0.022	-0.004	0.004
	(0.078)	(0.032)	(3.661)	(0.008)
∆ Placebo Week				
Off-Peak	-0.021	-0.035***	-0.022	-0.024
	(0.046)	(0.009)	(3.082)	(0.078)
Mid-Peak	-0.013**	-0.011	0.011	-0.003
	(0.004)	(0.013)	(2.144)	(0.022)
Peak	-0.014***	-0.025***	0.012***	-0.016***
	(0.004)	(0.006)	(0.003)	(0.004)
Firm-Month-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-TOU-Hour		Yes		Yes
Month of sample-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.949	0.954	0.957	0.960

Notes: This table presents estimates related to Equation (2) but computing the additional effect of the policy during weekdays with respect to weekends. Controls include hourly temperature and the interaction for whether the temperature is above or below 20ºC. Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

Table 10: Machine learning: Differences-in-Differences

	Prediction error			
	(1)	(2)	(3)	(4)
Policy				
Off-Peak	0.006	0.022**	-0.029***	-0.016**
	(0.008)	(0.011)	(0.007)	(0.006)
Mid-Peak	-0.023*	-0.047	-0.064***	-0.078***
	(0.012)	(0.030)	(0.007)	(0.017)
Peak	-0.042***	-0.094***	-0.095***	-0.126***
	(0.005)	(0.008)	(0.007)	(0.018)
Placebo				
Off-Peak	0.017***	0.033	0.019***	0.033**
	(0.004)	(0.028)	(0.004)	(0.013)
Mid-Peak	0.036***	0.012	0.016**	0.001
	(0.005)	(0.030)	(0.007)	(0.024)
Peak	0.053***	0.002	0.026***	-0.005
	(0.007)	(0.017)	(0.007)	(0.005)
Firm-Month-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-TOU-Hour		Yes		Yes
Month of sample-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.044	0.129	0.199	0.242

Notes: This table presents estimates of Equation (1). Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

Table 11: Machine learning: triple-differences

	Prediction error			
	(1)	(2)	(3)	(4)
Policy Weekend				
Off-Peak	0.010*	0.065***	-0.007	0.015
	(0.005)	(0.012)	(0.115)	(0.017)
Mid-Peak	-0.011	-0.035	-0.073	-0.063**
	(0.008)	(0.035)	(0.119)	(0.020)
Peak	-0.010	-0.050***	-0.084	-0.079***
	(0.008)	(0.006)	(0.865)	(0.016)
Δ Policy Week				
Off-Peak	0.007*	-0.006	0.011	0.002
	(0.004)	(0.018)	(10.270)	(0.033)
Mid-Peak	-0.012	-0.012	0.009	-0.014
	(0.009)	(0.013)	(2.769)	(0.019)
Peak	-0.033***	-0.044***	-0.011	-0.045
	(0.004)	(0.011)	(1.547)	(0.047)
Placebo Weekend				
Off-Peak	-0.006	0.049*	0.029	0.051
	(0.016)	(0.026)	(3.658)	(0.047)
Mid-Peak	0.048***	0.024*	-0.002	0.008
	(0.006)	(0.013)	(0.859)	(0.019)
Peak	0.074***	0.034	0.012	0.018
	(0.012)	(0.029)	(4.672)	(0.031)
Δ Placebo Week				
Off-Peak	-0.004***	-0.017**	-0.000	-0.009
	(0.001)	(0.007)	(0.137)	(0.040)
Mid-Peak	-0.012***	-0.013	0.018	-0.005
	(0.003)	(0.012)	(0.245)	(0.019)
Peak	-0.022***	-0.033***	0.013***	-0.021**
	(0.002)	(0.002)	(0.002)	(0.007)
Firm-Month-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-TOU-Hour		Yes		Yes
Month of sample-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.049	0.142	0.202	0.243

Notes: This table presents estimates related to Equation (2) but computing the additional effect of the policy during weekdays with respect to weekends. Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

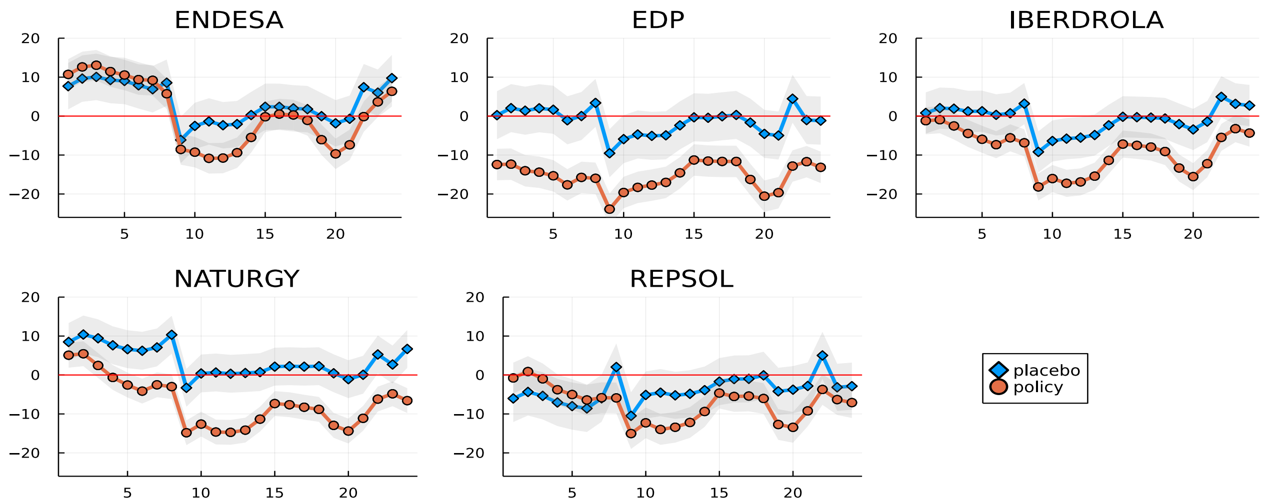
Table 12: Random Forests: Differences-in-Differences

	Prediction error			
	(1)	(2)	(3)	(4)
Policy				
Off-Peak	0.008	0.015**	-0.018**	-0.010***
	(0.007)	(0.006)	(0.007)	(0.003)
Mid-Peak	-0.019***	-0.028**	-0.055***	-0.060***
	(0.004)	(0.009)	(0.002)	(0.004)
Peak	-0.038***	-0.058***	-0.083***	-0.095***
	(0.006)	(0.008)	(0.008)	(0.006)
Placebo				
Off-Peak	0.005	0.013	0.008***	0.016**
	(0.004)	(0.011)	(0.002)	(0.006)
Mid-Peak	0.017**	0.008**	0.010	0.005**
	(0.007)	(0.004)	(0.006)	(0.002)
Peak	0.025***	0.005	0.015**	0.002***
	(0.005)	(0.008)	(0.005)	(0.001)
Firm-Month-TOU-Hour	Yes	Yes	Yes	Yes
Firm-Year-TOU-Hour		Yes		Yes
Month of sample-TOU-Hour			Yes	Yes
N	142,000	142,000	142,000	142,000
Adjusted R^2	0.084	0.155	0.175	0.219

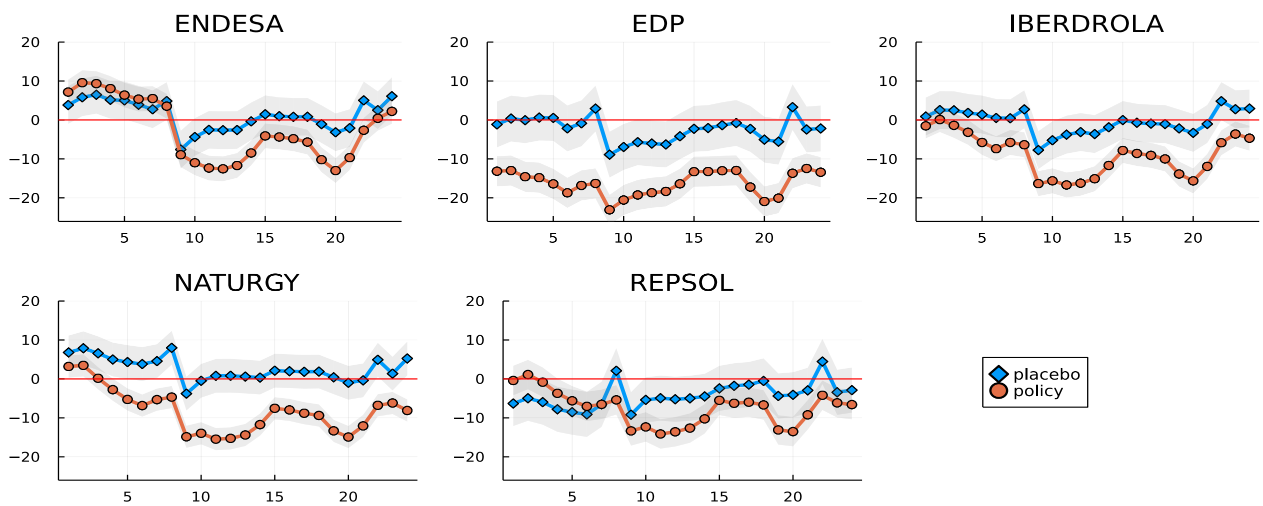
Notes: This table presents estimates of Equation (1). Observations are weighted by the number of consumers in each distribution area. Standard errors clustered at the firm-month level. Significance levels: ***p<0.01, **p<0.05, *p<0.1.

B Figures

Figure 9: Triple-differences coefficients by distribution area and for weekdays

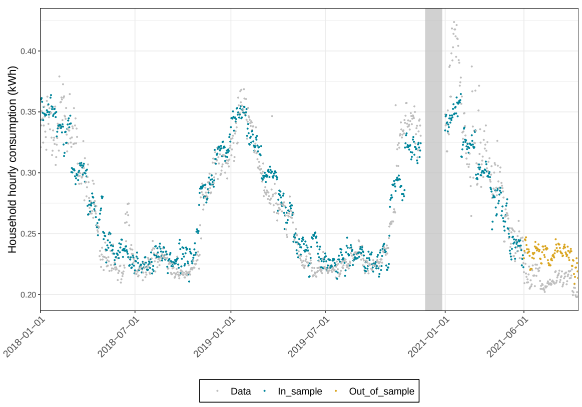


(a) Panel Fixed Effects

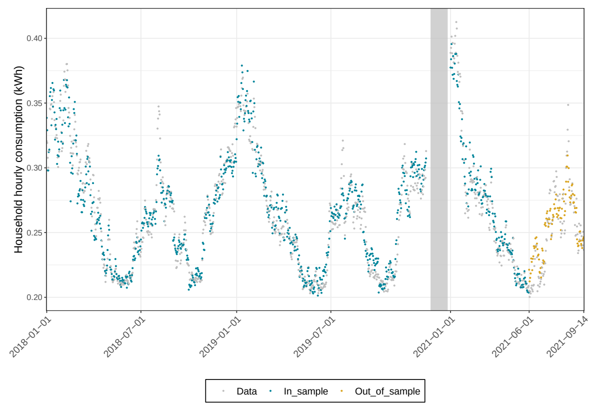


(b) Machine Learning

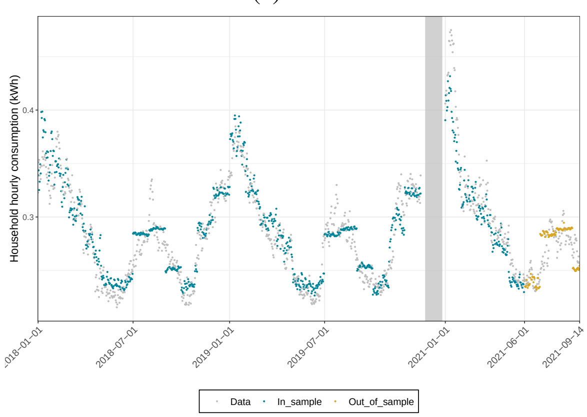
Figure 10: Machine learning: in- and out-of-sample predictions by distribution area



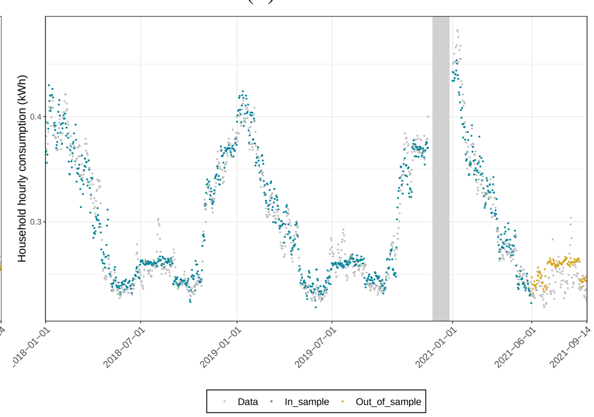
(a) EDP



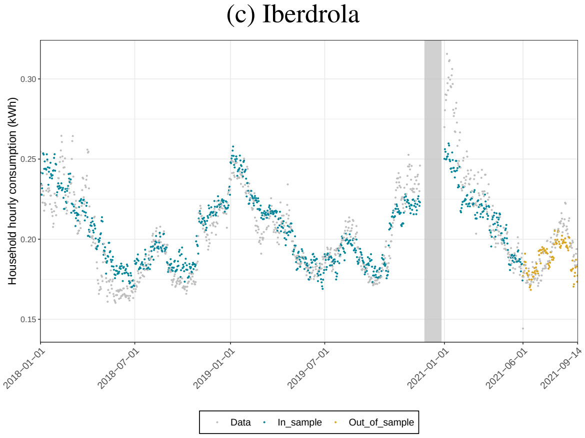
(b) Endesa



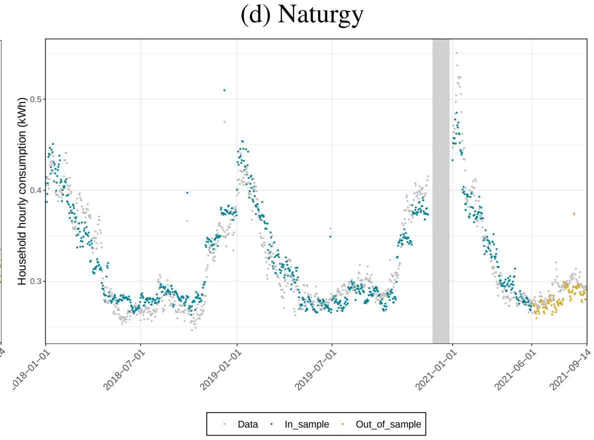
(c) Iberdrola



(d) Naturgy

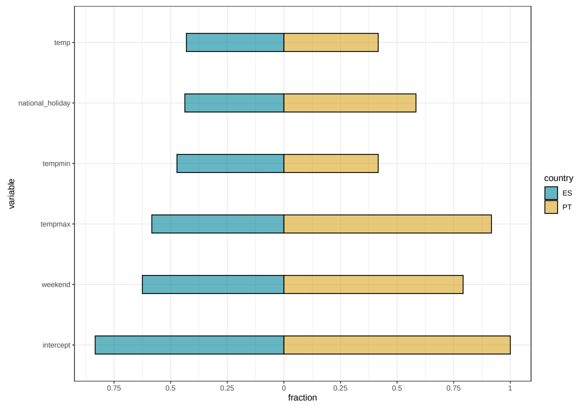


(e) Repsol

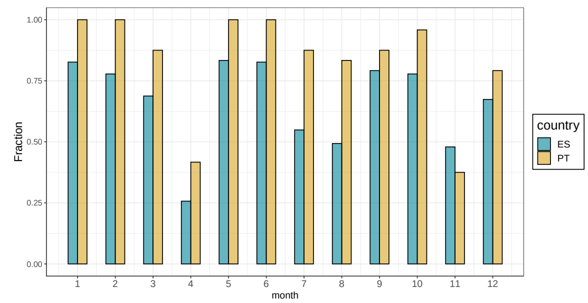


(f) Portugal

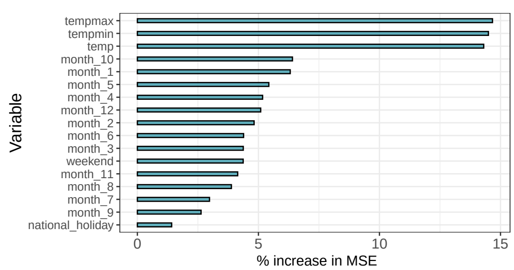
Figure 11: Machine learning: predictive capacity of control variables



(a) LASSO: fraction of models selecting controls

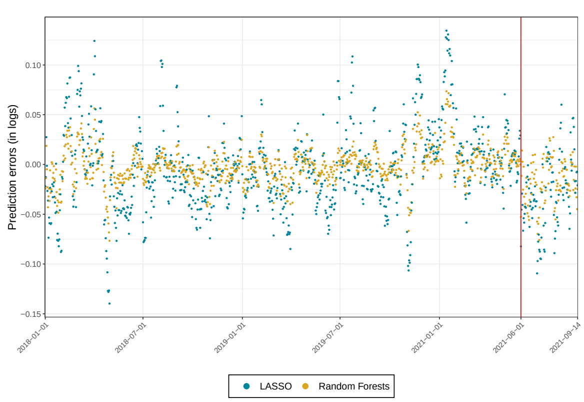


(b) LASSO: fraction of models selecting months

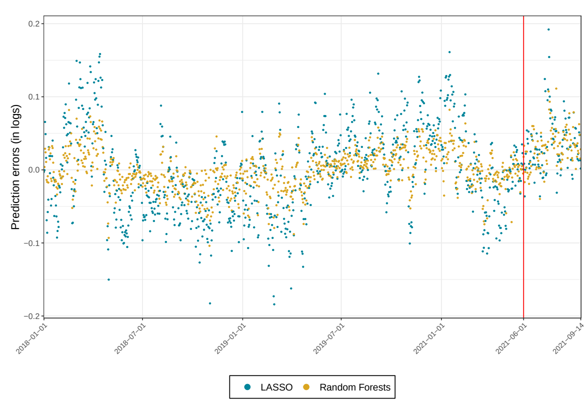


(c) Random Forest: variable importance

Figure 12: Prediction errors by method

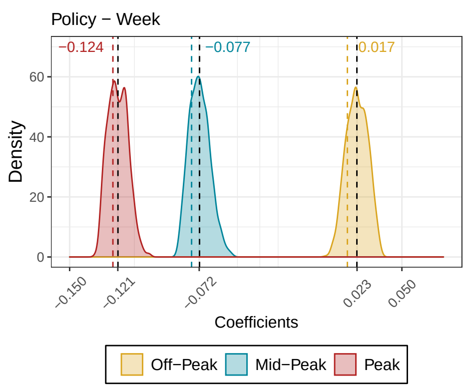


(a) Spain

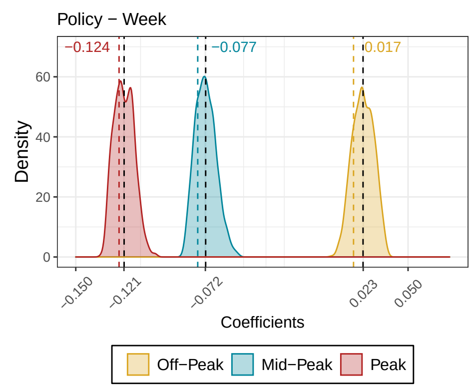


(b) Portugal

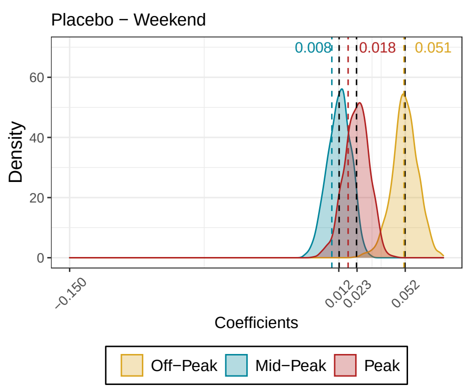
Figure 13: Bootstrapped Distribution of Triple-Difference LASSO coefficients



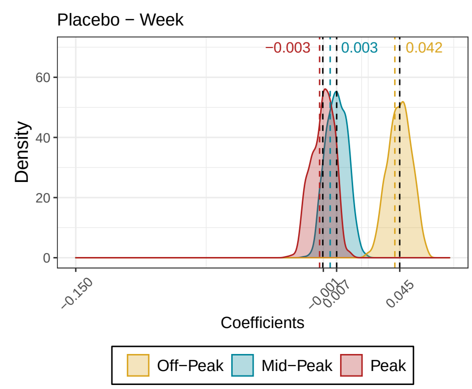
(a) Policy - Weekend



(b) Policy - Week



(c) Placebo - Weekend



(d) Placebo - Week