Utility Pricing in the Prosumer Era: An Analysis of Residential Electricity Pricing in California

(Job Market Paper Draft)

Felipe A. Castro * Duncan S. Callaway†

January 1, 2017

Abstract

Advanced metering infrastructure (AMI) as well as home energy management systems (HEMS) are enabling residential consumers, or "prosumers", to actively interact with electricity systems. In regulated retail sectors, residential rate structures greatly influence this interaction. Despite its important role, no consensus exists with respect to the ideal design of rates. Among the many factors that keep the debate open, the limited empirical literature turns the focus of the discussion around theoretical prescriptions. This study contributes with an applied analysis of residential rate structures, in the context of California's electricity sector. We focus on a scenario in which AMI and HEMS are widely adopted. To measure welfare changes under different pricing regimes, we embed a detailed model of household behavior into a framework that combines elements from peak-load pricing and capacity expansion. This approach allows us to capture a wide variety of temporal and spatial demand substitution patterns without the need of estimating a large number of parameters. We calibrate the model using data of appliance ownership, census household counts, weather patterns, and a model of California's electricity network. Our analysis compares four tariffs, including real-time pricing and three variants of a time-of-use program, and uses a flat rate structure as a reference case. We find that the increase in consumer surplus for the average household is not greater than 2 dollars per month. However, depending on the appliance stock, weather patterns and the specific rate structure some households may benefit up to ten times more than others. Overall our results suggests that targeting different rates to household with different characteristics is a superior strategy than defaulting all customers into any variant of a time-of-program.

JEL Classification: C44, C61, D41, D61, L51, L94
Keywords: Real-time pricing, Tariff design, Public utility pricing

* Correspondence: Department of Energy and Resources Group, University of California, Berkeley, CA 94709. Tel.: (510) 610-5262. Email: fcastro@berkeley.edu.

† Department of Energy and Resources Group, University of California, Berkeley, CA 94709. Email: dcal@berkeley.edu.

1 Introduction

The increasing penetration of *distributed energy resources* (DER) including advanced metering infrastructure (AMI), energy management systems (HEMS), solar photovoltaic and battery storage systems are enabling residential consumers, or "prosumers", to interact bidirectionally with electricity systems. Customers can not only purchase electricity from the grid but also provide the system with energy and reliability services (Schleicher-Tappeser, 2012). In regulated retail sectors, residential rate structures greatly influence this interaction. Rates can impact adoption decisions, by changing the economic value of the technology, and can influence the usage of this resource. Rates may also increase distributional disparities, creating cross subsidies between those who can and cannot adopt the distributed technology (Eid et al., 2014). Despite this crucial role, no consensus exists with respect to the ideal rate structure. Factors explaining this reality include limitations of the theory and the focus of the empirical work. While the theory asserts that real-time pricing (RTP) is optimal from an economic efficiency perspective (Joskow and Tirole, 2006), regulators must balance efficiency with other public goals, such as rate simplicity, equity or meeting environmental directives (Stanton, 2015). Absent a comprehensive quantification of the impacts associated with implementing RTP, it is difficult for regulators to judge the value of this alternative, especially when it may compromise other regulatory goals. The empirical literature has tried to quantify impacts, however, the scope has been somewhat limited. Researchers have focused on estimating price responsiveness to quantify changes in efficiency in the short-run.[1] Other relevant metrics such as long-run welfare effects, equity or environmental implications have been explored either in isolation or with stylized analyses, but never with an applied approach, within a unified framework.[2]

This work contributes to rate regulation policy with an applied study in the context of California's residential electricity sector. Our focus is the comparison of the long-run welfare effects as well as the equity and environmental implications of a set rate structures. The analysis considers the hypothetical scenario in which HEMS is widely adopted. Under this circumstance the household responses to price signals are likely to be fully rational, driven by an algorithm optimizing the consumption of the appliances (Beaudin and Zareipour, 2015).

[1]Examples include the work of Allcott (2011), Caves et al. (1984), Faruqui and Sergici (2010), Faruqui and Sergici (2011) and Herter (2007).

[2]See, for instance, Borenstein (2006), Borenstein (2013), Crew et al. (1995) and Joskow and Tirole (2006).

A second contribution is the development of a framework that permits comparing rate structures along each of the metrics we consider in our analysis. We develop a model of optimal pricing that accommodates a wide variety of tariffs. Our framework builds upon *peak-load pricing*[3], and borrows some elements from the literature of *generating capacity expansion*.[4] We embed a detailed model of household behavior in this setting, expanding the basic model of peak-load pricing to include heterogeneous households and the adoption of DERs. Our approach allows us to capture a wide variety of temporal and spatial demand substitution patterns, without needing to use a large number of estimates.

Our analysis is particularly relevant to the current regulatory situation in California, in which the default increasing block pricing program is gradually being retired and time-of-use becomes the default rate. We compare this prospective structure with two variants of this program, a TOU combined with demand charges (TOU&DC) and a TOU combined with a critical peak pricing program (TOU&CPP). We also include in the comparison the cases in which households enroll in a flat rate (FR) and in a real-time pricing program.

Considering all households under the FR rate as a reference case, our analysis shows that implementing any other pricing alternative produces gains for the average household that are mild at best. When implementing time-of-use pricing, or any of its variants, the average gain is not greater than 1.2 dollars per month. Even though real-time pricing performs better than the other rate structures, the improvement over the FR tariff seems mild. The average household increases its net surplus by 2 dollars per month under this rate scenario.

In addition, our analysis shows that the net surplus gain varies considerably across households. Factors such as the presence of an air conditioning system or the temperature outside the dwelling are major drivers of this variation. For all rates, households with air conditioners experience higher average gains than household without these appliances. However, the gains in net surplus vary more across the former group of households. The exterior temperature profile is a key factor. Its relationship with the net surplus gains, however, is not simple, and depends on the specific rate. For instance, under the RTP case customers experience greater gains in areas with higher average temperature. But this statistic has no correlation with gains under the TOU&DC program.

[3]For a comprehensive survey of the literature we refer the reader to Crew et al. (1995).

[4]Sauma and Oren (2006) provide a good example of these models.

These two results combined suggest that defaulting all residential customers into a time-of-use rate structure, which is the current path California is following for the residential sector, may not be the best strategy. Targeting different rates to households with different appliance stocks and in different locations will likely be a superior policy.

2 California Electricity Sector and the Emergence of Prosumers

2.1 An overview of the sector

The California electricity sector serves approximately 30 million people across the state. With 59 GW of power plant capacity, the sector delivers near 309 TWh of electricity annually. Its market size is close to $8 billion per year and its transmission system, spanning 25,627 circuit-miles, is part of the Western Interconnection. In terms of market regulation and oversight, there are three institutions involved, each with different roles. The first, the Federal Energy Regulatory Commission (FERC), has jurisdiction over the interstate transmission of electricity. Its responsibilities include the oversight of important merger and acquisitions, reviewing applications for transmission projects, as well as licensing and inspecting private, municipal and state hydroelectric projects. The commission also sets mandatory reliability standards and monitors energy markets across the US. The other two regulatory agencies have jurisdiction in the sate of California. One is the California Energy Commission which is the primary energy policy and planning agency of the state. The other is the California Public Utilities Commission (CPUC). Its main role is regulating the three investor-owned electric utilities of California, including Pacific Gas and Electric Company (PG&E), Southern California Edison (SCE) and San Diego Gas and Electric Company (SDG&E), which collectively serve two thirds of the electricity demand throughout California. Among other functions, the CPUC sets and approves the retail rates, is responsible for ensuring that utilities meet state environmental policies and ensures electricity safety at the distribution level.

In terms of system and market operations, the California's Independent System Operator (CAISO) is responsible for maintaining a reliable transmission of power as well as the comprehensive long-term planning of grid infrastructure. This entity also coordinates forward and spot markets for energy and ancillary services. In addition, CAISO complies with the reliability standards set by the North American Electric Reliability Corporation (NERC) and the Western

Electricity Coordinating Council (WECC). While the former is a non-profit organization developing and enforcing reliability standards for the continental United States, Canada and Baja California, Mexico, the latter is a regional entity promoting bulk electric system reliability in the Western Interconnection.

2.2 Residential rates in California

The distinctive characteristic of residential electricity rates in California is its increasing block structure. They have had this form since 1976, when the Miller-Warren Energy Lifeline Act was enacted. This legislation sought to provide California's residential customers with a minimum necessary quantity of gas and electricity at a fair price, and also to encourage conservation. The legislation set a precedent, providing a conceptual justification for implementing increasing block rates. Since 1976 rates did not change meaningfully until California's electricity crisis.

Beginning in the summer of 2000, tight supply margins, weak federal oversight, lack of an elastic demand and flaws in the market design yielded a period of highly volatile electricity prices, known as California's Electricity Crisis. As a result of the crisis the sector underwent a period of drastic reforms. At the retail level, a first response to the high wholesale prices was to lift the retail price cap. This triggered notorious increases in electricity bills, which were then mitigated by freezing the charges of the lower two tiers. The result of this legislation was the replacement of a two tier system by a five tier structure, with prices of Tiers 3 to 5 considerably higher than those of the remaining lower tiers. From 2000 to 2009 differences among tiers increased. However, the enactment of SB695 in 2009 began to allow limited annual increases for Tiers 1 and 2.

At the time of this writing, decision $D.15-07-001$ is the main piece of regulation laying the path for the future of residential electricity rates in California. Key elements of this regulation include the promotion of the consolidation of the tiers and the development of rates that reflect better cost causation. In particular, the decision approves transitioning all residential customers to a default time-of-use tariff by 2019.

2.3 The emergence of prosumers

There are two main forces pushing the emergence of prosumers in California: Environmental policy directives and distributed technologies reaching maturity. The relevant environmental policy is the renewable portfolio standard which established, among others, targets for distributed generation. The policy has triggered the development of an array of incentives for generation at the customer's premises which has caused the massive deployment of distributed energy resources, such as solar photovoltaic panels. As for technological evolution, California has taken major steps towards modernizing its distribution grid, having the largest installation of AMI in the US. This technology constitutes a vital element for implementing time-varying-rates. By enabling two way communication between customer and utility on time intervals of an hour or less, AMI allows utilities measuring consumption on an hourly basis as well as sending price signals on a consistent time scale.

3 A Modeling Framework to Compare Rate Structures

3.1 Utility pricing: An overview of the theory

An important function of regulators is determining the rates that a regulated utility can charge for the provision of its services. This process, also known as *rate regulation*, includes the determination of the rate *rate level* and the *design of rate structures* (Phillips Jr, 1993, pp. 176 - 180). Establishing the rate level entails specifying the total compensation that the utility, or load serving entity (LSE), receives for its services. The design of rate structures, which is the focus of this work, defines how the LSE collects its compensation. Designing rate structures is far from trivial. So it is not surprising that a myriad of methodologies have been suggested and adopted in different jurisdictions. Brown and Sibley (1986, pp. 44 - 60) divide the approaches according to how they assign *common* or *non-attributable* costs across different services and consumer classes. There are two broad categories: the *cost-based pricing* and pricing based on the concept of *marginal cost*. The first group of approaches allocate costs based on criteria other than efficiency. One example is the *fully distributed costs method*, in which common costs are assigned according to the relative shares of magnitudes that can be attributed to a service or group of customers, such as peak-demand, output or revenue.[5] On the other hand, in pricing based on the concept of marginal cost efficiency has a prominent position. How common costs are attributed to the different groups of customers is a byproduct of a welfare maximization

[5]Other approaches seek to minimize cross subsidies across services and consumer classes and others build a set of axioms and derive rate structures consistent with them. For more details on the subject of cost-based pricing, see (Brown and Sibley, 1986, pp. 44 - 60).

process. Our framework falls into the second category of approaches. The model that this work introduces produces a set of rates and allocations of costs that emerge from the welfare maximization of the system under study. This is the approach to rate design that the literature of peak-load pricing studies.

Peak-load pricing develops a normative theory of efficient or welfare maximizing pricing for industries with limited storage capability and time-varying demand. The modern version of the theory originates with the contributions of Boiteux (1960) and Steiner (1957), and intended to provide guidelines in the context of price regulation of natural monopolies, such as vertically integrated electric utilities (Crew et al., 1995). The basic model considers the problem of a social planner choosing prices that maximize welfare, i.e., the surplus of customers and the public utility's profits. Prices coordinate production and consumption decisions over a time horizon. The monopolist invests in production capacity at the beginning of the horizon and prices are such that the utilization and the level of the installed capacity are optimal (Drèze, 1964). Studies including Carlton (1977), Chao (1983), Crew and Kleindorfer (1976) and Panzar (1976) further refine the model to include a stochastic demand, supply-side uncertainties and multiple technologies.

Borenstein and Holland (2005) and Joskow and Tirole (2006), in examining the merits of retail competition in the electricity industry, show that the theory also applies to restructured electricity sectors. In these models, a competitive wholesale market replaces the production side of the vertically integrated utility. At the retail level, both papers distinguish the cases of a regulated distribution company and competitive retailers. While Borenstein and Holland (2005) consider a setting with linear or uniform prices, Joskow and Tirole (2006) contemplate the case of a two-part, non-linear price. Further, Borenstein and Holland (2005) explore the long-run effects of different pricing policies, analyzing the equilibria that emerge at the wholesale and retail levels. More recently, Zöttl (2010) investigates a setting in which there is imperfect competition at the wholesale level, and Chao (2011) updates earlier work exploring the interaction of different pricing policies and renewable technologies.

Our model departs from previous work first by generalizing the type of rate structures present in peak-load pricing. In addition, we introduce a mechanism linking pricing and technology adoption decisions. Finally, our model accommodates household heterogeneity beyond a scale factor.

3.2 The regulator's problem

The regulator's problem combines elements of the peak-load pricing and capacity expansion literature. The key element from capacity expansion not present in peak-load pricing is a transmission network. For simplicity, we do not detail this element of the model in this section. We refer the interested reader to the appendix A.4. As in peak-load pricing, our model falls into the broad category of two-stage stochastic optimization models. Agents in these models make long-run decisions at the beginning of the horizon before uncertainty is realized and define state contingent strategies for the short-run stage. These are static models that can describe systems in steady state. Our framework, therefore, is not suitable for studying system dynamics. In terms of the institutional setting, at the retail level we consider a distribution utility as the load serving entity. In general, however, one can consider settings within two polar cases. While the utility could be fully integrated with the supply side in one case, in the opposite, it could be just a distribution company. Under the assumption of perfect competition at the wholesale level, both cases are equivalent (Joskow and Tirole, 2007), however.

Let ω index a finite and countable set Ω of states of nature, π the corresponding probability vector, $E[\cdot]$ the expectation operator, and a time horizon of $t \in T$ time steps. Given a random vector of consumption d, the household pays $l+\eta(d,p)$ to the utility, where l is a fixed charge, $p \in \mathcal{P}$ a vector of rate parameters and $\eta(\cdot)$ a fee contingent on consumption and the rate parameters. We call the triple $(l, \eta(\cdot), p)$ a rate structure. Our setting is similar to the one in (Joskow and Tirole, 2006) insofar we focus on two-part structures with a state and time contingent demands and prices. However, we generalize this model to accommodate more complex rate structures. In our setting the vectors d do not only have one component for every time and state of nature but also may include other relevant metrics associated to the demand profile, such as peak or total consumption across the time horizon. Similarly, price parameters may include charges for peak or total demand. Specifically, we focus on the case in which $\eta(\cdot)$ is bilinear on the demand vector and price parameters. The following assumption formalizes this specification.

Assumption 3.1 *The demand contingent charge* $\eta(d,p) = d^\top Mp + Ind\left\{\tilde{b} - \tilde{A}d\right\}$*, where* $Ind\{x\}$ *is 0 if* $x \geq 0$*, and* ∞ *otherwise.*

As subsection 3.5 shows, this specification is fairly general, allowing researchers to model a wide range of the rate structures used in practice.

As in peak-load pricing, we consider homogeneous households, with a mapping $D_\omega(p) : \mathcal{P} \to \mathbb{R}^{|T|}$ and a real valued function $U_\omega(d_\omega)$ representing their demand and gross surplus metric, respectively. Given a set of wholesale prices $\{\lambda_\omega\}$, the planner problem optimizes the household net surplus, $E\left[U_\omega\left(D_\omega(p)\right) - \eta\left(D_\omega(p), p\right)\right] - l$, and guarantees that the utility meets its revenue requirement, $E\left[\eta\left(D_\omega(p), p\right) - D_\omega(p)^\top \lambda_\omega\right] + l - \Pi$, with Π an exogenous fixed cost. As (Joskow and Tirole, 2006) shows, this amounts to find (l^*, p^*) such that

$$(d^*, p^*) \in \underset{(d,p)}{\arg\max} \left\{ E\left[U_\omega(d_\omega) - d_\omega^\top \lambda_\omega\right] : \, d_\omega = D_\omega(p) \ \forall \omega \in \Omega \right\}, \tag{1}$$

$$l^* = E\left[\lambda_\omega^\top d_\omega^*\right] - \eta(d^*, p^*) + \Pi. \tag{2}$$

3.3 A competitive wholesale electricity market

The wholesale market representation in this model is a variant of the supply-side model studied in the peak-load pricing and capacity expansion literature. More specifically, we follow closely the representation in (Chao, 2011). In this model, infinitesimal competitive firms interact in a spot market for electricity. Each decides on their long-run installed capacity and short-run generation profiles. We denote the total installed capacity of technology $k \in K$ as x_k and its cost of carrying capacity as $\tilde{r}_k$. The aggregated production profile of this technology in state of nature ω is $y_{\omega k} \in \mathbb{R}_+^T$, and variable costs per unit of power production is $c_{\omega k} \in \mathbb{R}_+^T$. We capture variability in a technology's availability – e.g. due to outages – with an availability factor per technology contingent on the states of nature, $\rho_{\omega k} \in R^T$. In a perfectly competitive market firms are price takers, thus, production and capacity for technology k are the solution of the problem

$$\max_{(y_{\omega k}, x_k)} \left\{ E\left[(\lambda_\omega - c_{\omega k})^\top y_{\omega k}\right] - x_k \tilde{r}_k \colon 0 \leq y_{\omega k} \leq x_k \rho_{\omega k} \right\}.$$

The market equilibrium is a tuple $(d^*, p^*, y^*, x^*, \lambda^*)$ such that (d^*, p^*) solves the regulator's problem at λ^*, (y^*, x^*) solves the problem of the producer at that price, and supply equals demand. It is easy to verify that the market equilibrium is the solution of

$$\max_{(d,p,x,y)} E\left[U_\omega(d_\omega) - \sum_{k\in K} y_{k\omega}^\top c_{k\omega}\right] - x^\top \tilde{r} \tag{3}$$

subject to

$$d_\omega = \sum_{k\in K} y_{\omega k} : \lambda_\omega, \tag{4}$$

$$0 \le y_{k\omega t} \le x_k \rho_{\omega k}, \tag{5}$$

$$p \in \mathcal{P}, \tag{6}$$

$$d_\omega = D_\omega(p). \tag{7}$$

3.4 The household behavior

Except for our specification of the demand contingent fee, $\eta(\cdot)$, (3) – (6) is the classic peak-load pricing problem. Researchers can use the model to analyze theoretically and numerically implications of different constraint sets for the vector of retail prices. A key assumption that facilitates the study of these models is a demand system with analytic expression. Our framework drops this assumption because our specification of $\eta(\cdot)$ implies, in general, demands with no analytic definition. Consistently, our model updates the peak-load pricing problem replacing (7) with the following condition,

$$d \in \arg\max_d \left\{ E\left[U_\omega\left(d_\omega\right) - d_\omega^\top M_\omega p_\omega\right] : \bar{b}_\omega - \bar{A}_\omega d_\omega \ge 0,\ \forall \omega \in \Omega \right\}, \tag{8}$$

where $(\bar{b}, \bar{A})$ contain the parameters of the rate structure, $(\tilde{b}, \tilde{A})$, and possibly others. Henceforth we refer to (8) as the *household problem*, and to (3) – (6), (8) as the *pricing problem.*

3.5 Illustrative examples

Our specification of the household demand allows to model the influence on demand of several rate structures and, also, represent demand-side technologies of interest. Here we show how to implement the models that we use in our analysis. The appendix A.2 provides additional examples. Some notation will prove useful. The matrix I_m corresponds to the identity of m by m. The vectors e_m and z_m are, correspondingly, vectors of ones and zeros of m dimension.

Modeling rate structures. Our analysis compares *time-varying pricing* (TVP) and a

TVP combined with a *demand charge* (DC). A time varying pricing is the simplest type of rate to model. Set $M_\omega = I_{|T|}$, let the vectors $p_\omega \in \mathbb{R}^T_+$ and $d_\omega \in \mathbb{R}^T$, and define define $\mathcal{P}$ as follows,

$$\mathcal{P} := \begin{cases} \left\{p \in \mathbb{R}_+^{|T|\times|\Omega|} : \, p_{\omega t} = p_{\omega' t'} \, \forall (\omega, t), (\omega', t')\right\} & \text{for FR,} \\ \left\{p \in \mathbb{R}_+^{|T|\times|\Omega|} : \, p_{\omega t} = p_{\omega' t'} \, \forall (\omega, t), (\omega', t') \in TW(\omega, t)\right\} & \text{for TOU,} \\ \left\{p \in \mathbb{R}_+^{|T|\times|\Omega|}\right\} & \text{for RTP,} \end{cases}$$

where $TW(\omega, t)$ is the set of time windows (ω', t') in the same time window as (ω, t).

Adding a demand charge to any of these structures requires redefining $d := [\bar{d}, \hat{d}]$ and $p := [\bar{p}, \hat{p}]$, where $\bar{d}, \bar{p} \in \mathbb{R}_+^{|T|\times|\Omega|}$, and $\hat{d}_\omega, \hat{p}_\omega$ correspond to the maximum consumption and demand charges under ω, respectively. The matrix M_ω is now equal to $I_{|T|+1}$, and the analyst may add additional conditions to the set $\mathcal{P}$ to model demand charges constant across some scenarios. A final element of this structure is the constraint linking the hourly consumption profile $\bar{d}$ and the maximum consumption $\hat{d}$, which we model via the following definitions

$$\tilde{b} := z_{|\Omega|\cdot|T|}, \;\; \tilde{A} := \begin{bmatrix} I_{|\Omega|\times|T|} & -I_{|\Omega|} \otimes e_{|T|} \end{bmatrix}.$$

Household as composite of devices. Following the approach of Reiss and White (2005), we consider that households are composite of devices and assume their utility functions are additively separable. For reasons we explain later, we consider in our analysis two devices, a central air conditioning unit and a rooftop solar panel, and the household baseline. Central air conditioning falls in the more general category of *thermostatically controlled loads* (TCL's), whose behavior follows the laws of thermodynamics. Mathieu et al. (2015) presents a model describing the behavior of these appliances, which links the household's inside temperature, θ, with the outdoor temperature $\tilde{\theta}$, the thermal characteristics of the dwelling, ξ, and the electricity consumption of this appliance, d. It is possible to show that the inside temperature profile has the following form

$$\theta(d; \xi, \tilde{\theta}) = \Theta_1(\xi) d + \theta_2(\xi, \tilde{\theta}), \tag{9}$$

where Θ_1 and θ_2 are a matrix and a vector, functions of the thermal parameters and temperature outside the dwelling. Appendix A.3 shows the full derivation of this relationship. Here we close the model of the TCL behavior introducing a mechanism capturing household's preferences for thermal comfort. The simplest approach involves a penalty for deviating from an ideal inside temperature, $\hat{\theta}$. Equation (10) shows a utility function consistent with this approach, which we use in our analysis.

$$U(d) = -\beta \left\| \theta(d; \xi, \tilde{\theta}) - \hat{\theta} \right\|^2 \tag{10}$$

For modeling the household baseline, we assume a linear demand system and compute the associated utility function using a standard procedure. For the rooftop solar panel we add to the household problem a constraint limiting its hourly production given the hourly availability of the solar resource. A final element of the household model links the demands of each device with the net demand of the customer,

$$\bar{d} = d_{baline} + d_{ac} - d_{solar}. \tag{11}$$

3.6 Household heterogeneity and DER adoption

The pricing problem contemplates one representative customer and there is no mechanism modeling customer adoption. In our analysis, however, we consider heterogeneous customers and analyze impacts of pricing on adoption. We incorporate these two elements using the framework that Castro and Callaway (2016) develop. The paper distinguishes different customer types $i \in I$,[6] each of which decides a set of technologies to adopt $j \in J$. [7] Calling the combination $h := (i, j)$ a *segment* and defining α_h and r_h, respectively, as the number of households and cost associated to a segment,[8] the paper shows how modifying the pricing problem permits modeling adoption decisions. Specifically, equation (3) becomes

$$\max_{(\alpha,d,p,x,y)} E\left[\sum_h \alpha_h \left[U_{h\omega}(d_{h\omega}) - r_h\right] - \sum_{k\in K} y_{k\omega}^\top c_{k\omega}\right] - x^\top \tilde{r} \tag{12}$$

and (4) updates to

$$\sum_h \alpha_h d_{h\omega} = \sum_{k\in K} y_{\omega k} : \lambda_\omega. \tag{13}$$

[6] In order to fix ideas consider the following two examples: $I = \{\text{with central AC}, \text{without central AC}\}$ or $I = \{\text{live in hot weather}, \text{live in cold weather}\}$.

[7] An instance of this set is $J = \{\{\text{solar PV, battery storage}\}, \{\text{solar PV}\}, \{\text{battery storage}\}\}$.

[8] The cost associated to a segment (i, j) is the annualized capital cost of the set of technologies j.

A final element to include for modeling adoption is the feasible region for α, which ensures that the number of households per segment is consistent with the number of households per customer type.

3.7 Comparing rate structures

Subsections 3.2 to 3.6 develop an analytic tool to compare rate structures. Researchers can explore the effects of different structures on welfare and other metrics by changing the specification of the consumption contingent fee, $\eta(\cdot)$, solving the pricing problem and comparing the metrics of interest. While the method is not intended to predict what would happen were the tariff under analysis in place, it provides a consistent assessment of the potential differences between them.

Solving the pricing problem is not straightforward. The problem falls into the broad category of *bilevel problems*, in which a leader – in our setting, the regulator – indirectly controls the actions of the follower – the household – changing one or more parameters of her problem. Bilevel problems are hard to solve in general, and state of the art solvers can only handle problems of moderate size. For large instances, researchers have to devise specialized algorithms. Given the size of the instance we explore in this study, we had to develop a specialized algorithm as well. However, the development of the algorithm is beyond the scope of this work. It constitutes a completely separate research effort, which Castro et al. (2016) describes in detail. The basic idea is to decompose the pricing problem into one problem per household and state of nature, and one problem that coordinates the demands of the households. The algorithm iterates solving all problems at each repetition and stops when consecutive solutions do not change. The key aspect of the algorithm is its distributed nature which, by enabling its implementation in cluster computing facilities, makes our modeling framework practical.

4 Modeling California's Electricity Sector

We construct our model of the California electricity sector supplementing the network model that Price and Goodin (2011) developed for market analysis. The model consists of a network with 240 nodes, or buses, which corresponds to a topological reduction of the transmission system encompassing the Western Interconnection.[9] This reduced system also provides generation technologies and demands at each node, and the physical characteristics of the network, including transmission constraints. The generating power plants correspond to aggregations per type of fuel. Non-dispatchable generating technologies[10] such as solar or wind generation, and reservoirs such a geothermal or hydro power plants come with a year of hourly energy production. Fossil fuel technologies, on the other hand, only include physical and short-run economic parameters, such as heat rates and fuel costs. As for the demands, the network model includes a year of hourly energy consumption at nodes of the network corresponding to demand centers.

Before describing how we complete this data set, we make two clarifications. At first sight, the network model has more information than this analysis requires, because the Western Interconnection includes more states than just California. Using the full interconnection model, however, allows us to produce realistic import and export flows and provides the opportunity to study the impacts of residential pricing policy outside California. Because the main computational difficulties emerge from our detailed modeling of the demand, and because we do not model in detail demand at nodes outside California, the full network model did not increase significantly the computational complexity of our analysis.

A second clarification relates to our treatment of the Western Interconnection as an integrated market. That is, in our model the physics of the transmission lines is the unique factor limiting the flow of electricity through the interconnection. In practice, the administration of this system is divided among 38 *balancing authorities*, each controlling one portion of the network, and whose central role is to guarantee the reliable operation of their respective sub regions. This adds additional limitations to the flows of electricity which our model does not capture. At the time of this writing, however, California is leading the efforts to assess the impacts of a multistate regional market for the Western Interconnection.[11] Thus, an integrated market is plausible for the future of the interconnection.

4.1 Generating technologies

Because we are interested in studying long-run impacts, we replace the cost functions in the network model with the functions we described in subsection 3.3. A fixed and variable cost

[9]We refer the reader to the Western Electric Coordinating Council website for detailed information on the interconnection.

[10]Plants with outputs that are determined to great extent by exogenous factors such as weather conditions.

[11]See Brattle et al. (2016) for further detail.

implement these functions. The fixed cost includes the annuity associated with developing and installing the generating technology and the fixed O&M costs. The variable cost, on the other hand, encompasses fuel and variable O&M costs.

In addition to these economic parameters, technologies have associated emissions and availability factors. While the former captures the fact that different fuels have different GHG emissions, the latter reflects the fact that power plants experience unplanned outages. Emissions factors as well as the economic parameters of the generating technologies come from EIA (2016), and table 1 summarizes the specific values we use in this study. As for availability factors, we use the magnitudes that NERC makes publicly available through its Generating Availability Data System (GADS).

Table 1: Generating technologies costs and GHG emissions

technology	variable cost	fixed cost	emissions factor
	[2015 $/MWh]	[2015 k$/MW-year]	[tCO2eq/MWh]
nuclear	11	448	-
biomass	35	424	1.93
coal	32	361	0.65
geothermal	-	341	-
solar	-	192	-
wind	-	184	-
gas adv CC	39	84	0.33
gas conv CC	42	77	0.35
gas adv CT	63	52	0.52

We include in the model the existing plants, by technology, for each node.[12] As mentioned, investment decisions to expand this fleet are endogenous to the model. We treat hydro power generation as exogenous because a correct treatment of this technology, which involves the stochastic dynamic optimization of reservoirs, is beyond the scope of our model. Even though solar and wind are non-dispatchable generators, we use the time series in the data set only to compute hourly availability factors.[13] The actual hourly production for these technologies is ultimately determined by their installed capacity, an outcome of our model.

[12]Appendix A.1 shows the geographic distribution.

[13]Hourly availability factors are the ratio between the hourly production and the nameplate capacity of the technology in the data set.

4.2 Developing a model of California's residential demand

We construct a model of the residential sector of California calibrating our model of household behavior at each node of the network. Households can consume and produce electricity on an hourly basis, and impact the system via their net demands. In terms of consumption, we consider two major categories of household end uses: cooling and non-cooling. We take this approach for two reasons. In California central air conditioning is a major source of electricity consumption, and approximately one out of every two households has this type of appliance (Palmgren et al., 2010). In addition, studies at the appliance level report air conditioning to be a major source of demand responsiveness (Mathieu et al., 2015; Reiss and White, 2005).

The second reason relates to the hourly demand data available for this study. In our framework the baseline of a household is the fraction of its demand not modeled as any particular device. In order to calibrate this function one needs an intercept, i.e., a time series of electricity consumption and the prices in effect when this happened. For the demand part of the intercept we use the load shapes developed by Itron Inc., described in Wei et al. (2013). This data set disaggregates residential consumption into space conditioning and other loads.

Utility functions summarize household preferences for each end use. We calibrate them using the appliance level elasticities and marginal effects estimates that Reiss and White (2005) report. We model the baseline consumption as a linear demand system and use an elasticity of -0.08, corresponding to the estimate for households with no space conditioning. For the price intercept, we follow the procedure that Borenstein and Holland (2005) describe, assuming the rate structure of the intercept to be a flat rate.

The model for cooling corresponds to the TCL model developed in subsection 3.5. This has three groups of parameters which can be categorized as technical, behavioral and weather related. Technical parameters include the thermal resistance and capacitance of the household, and the efficiency of the air conditioner. Mathieu et al. (2015) provide ranges for these parameters in California, and we use the midpoint of those ranges in their study. The behavioral parameters are the ideal interior temperature, which we set to 22°C (or 72°F), and the discomfort penalty, β. Using our model of the TCL, we link this latter parameter with the estimate of the marginal effect for central air conditioning of Reiss and White (2005). The expression linking the two magnitudes is

$$\beta = -\frac{e_{|T|}^{\top}\Theta_1(\xi)^{\top}\Theta_1(\xi)e_{|T|}}{\frac{d}{dp}E[e_{|T|}^{\top}d_{\omega}]}. \tag{14}$$

The weather related parameter corresponds to the outside temperature. The National Renewable Energy Laboratory (NREL) develops the Typical Meteorological Year (TMY) data set for modeling energy conversion systems.[14] This data set contains 12 months of hourly data at selected locations across the US. The data for each month typifies conditions for the location over a longer period of time, such as 30 years. There are 73 locations corresponding to California. Based on distance we assign each of these locations to one of the buses with demand within the California portion of the network model. The TMY data set also comes with hourly values for solar radiation. We use them for our model of rooftop solar panels.

Another input for the analysis is the count of households types per bus. We distinguish four types of households in our analysis, corresponding to the combinations of tenure status and the presence of central air conditioning. While section 5 discusses this categorization further, here we focus on the calibration of the household counts. The sources of data for this task are the 2010 census and the Residential Appliance Saturation Survey 2010 (RASS), described by Palmgren et al. (2010). The General Housing Characteristic data set of the 2010 census contains counts of occupied households by tenure. Based on distance we assign census counts at the tract level to each bus in order to calibrate the total number of renter and owner occupied households. Similarly, we use the distance between the zip code centroid and the bus to assign the RASS responses to each bus. The survey, besides recording appliance ownership per survey participant, includes their tenure status. We assume that the fraction of household under each of the household types is that of the RASS survey and we multiply this fraction by the census counts per bus to estimate the final number of households under each of the categories that we analyze.

A final piece of the demand side is that corresponding to commercial and industrial customers. The test system comes with total load profiles per bus. The commercial and industrial load at each node of the network model is the difference between the total and aggregated residential demand at each node. We assume the latter quantity to be equal to the baseline profiles multiplied by the households counts.

[14]For a detailed description of the data we refer the reader to Wilcox and Marion (2008).

5 An Analysis of Residential Rate Structures in California

This analysis explores efficiency, as well as the distributional and environmental impacts of residential electricity rate structures. We use our model to quantify these metrics for five different tariffs that constitute plausible future residential rates in the Californian electricity sector.

Because our analysis focuses on long-run impacts, ideally one would have to compare net surplus distributions with respect to wealth levels under the different pricing regimes in our study. This approach is impractical, however, for at least two reasons. First, to the best of our knowledge information on households wealth for California is not publicly available. Thus, having this input would require an indirect calculation, which is an effort beyond the scope of this research. A second reason is that our model does not directly account for household wealth. There is no explicit mechanism linking this metric with either short- or long-run decisions. We assume, alternatively, that the level of wealth of a household translates into differences in its technology options. Wealthier customers have access to a wider variety of technologies.

Consistently, we split the population according to whether they can or cannot adopt distributed energy resources. The splitting criteria is household tenure status. That is, we assume homeowners have enough resources to purchase DERs while renters do not. Even though this criteria reflects the reality in California in the past decade (Borenstein, 2015), with better financing alternatives and new business models such as community solar[15] our assumption may not be adequate for future analyses.

An additional clarification relates to the specific rate structures we use in our analysis. An important element determining the final definition of the time-of-use schedules are the time windows associated with the different charges. Commonly, these rate structures distinguish valley and peak periods and also seasons in order to set the volumetric charges. The traditional approach is to consider existing time windows as inputs. In this analysis we take a different path which avoids two difficulties involved with the traditional approach. One is that existing time windows are likely inadequate for future system conditions. In California only a small fraction of households have been enrolled in TOU's programs. As TOU becomes the default rate for residential customers and the *net load*[16] shape changes due to the increasing penetration of

[15] We refer the reader to Huijben and Verbong (2013) for further discussion.

[16] Net load is the net of the aggregated demand, or system load, and the non-dispatchable generating technologies.

renewable generation, the existing time windows will likely be obsolete. Furthermore, a TOU with a demand charge rate has not yet been implemented at the residential level in California. A second difficulty is the sub-optimality of setting time windows exogenously. This makes the comparison among rates inconsistent because our framework computes optimal retail prices for the RTP and FR programs.

In order to avoid these shortcomings, we consider the most flexible type of TOU possible. That is, tariffs in which the energy charge can vary hourly and across seasons but not for days occurring in the same season. Similarly, we assume a demand charge that changes across seasons. Our preliminary analysis indicates that four to five time-windows, depending on the season, can approximate with no meaningful efficiency loss the hourly windows. The results we report in this analysis, however, correspond to the hourly energy charges.

A final clarification relates to the DERs we include in our exercise. Besides considering all residential customers having AMI and HEMS, originally, homeowners were able to adopt either rooftop solar PV systems or battery storage units. Preliminary results indicated that the latter two DERs were not cost-effective alternatives. No customer under any of the rates we study, nor under current or projected economic parameters for these technologies, adopted these DERs. In the case of the solar PV systems, this result indicates that the factors driving the current adoption levels of this DER are policies specially designed to promote this technology. These include the California Solar Initiative (CSI), federal subsidies and the increasing block rate structure for residential customers. Because the CSI is not effective anymore, and the future of federal subsidies is uncertain, we do not include these policies in our analysis. On the other hand, even though increasing block structures are being phased out, a surcharge for high monthly consumption will remain. This makes the study of this structure relevant. Future versions of this analysis will include this rate.

In the case of battery storage systems, our preliminary results do not make a case against this technology. It simply reflects the fact that our model only accounts for the energy arbitrage value that a battery storage unit can create. In addition, however, this technology can provide ancillary services for the distribution grid and serve as a mean of transportation when being part of an electric vehicle. Because our model does not capture any of these value streams, we do not include this technology in the final analysis.

5.1 Aggregated efficiency gains

Table 2 shows efficiency gains of four tariffs with respect to the base case scenario: the flat rate structure. At the level of the Western Interconnection, the RTP rate achieves greater efficiency gains followed by the time-of-use combined with a critical peak pricing program. The TOU combined with a demand charge produces similar gains but the time-of-use program alone only increases the net benefit by one half of the value when combined with another program. All programs reduce the aggregated benefit - or gross surplus - of the residential sector. However, the reductions in costs more than compensate the reductions in gross surplus.

In terms of efficiency increases, the same ranking does not hold when focusing on the residential sector in California. The main difference is that the TOU&DC rate structure increases the net benefit the least. This is the result of differences in bill reductions for customers inside and outside California. The presence of a transmission network explains this outcome. In our framework, the bill of a customer is equal to the multiplication of the *locational marginal prices* (LMPs)[17] by her consumption profile. The topology of the network significantly influences the magnitude of the LMPs at different nodes and, thus, the household bill at different locations. Differences in LMPs then explain differences in the distribution of bill reductions, in and outside California. In the case of the TOU&CPP rate, customers outside California capture an important fraction of the cost reductions with respect to the FR case. In all the other cases, on the other hand, the Californian residential sector captures most of the reductions in costs.

The average efficiency gains per household are mild at best, being not greater than 2 dollars per month. Importantly, in all cases, with the exception of the RTP program, the gains appear insufficient to justify the implementation of time-varying rate structures. Implementing any time-varying rate requires the deployment of AMI. Estimates of the cost of this infrastructure vary. However, one can construct a reasonable range using the documentation of pilot projects conducted under the American Recovery and Reinvestment Act of 2009 in DOE (2012a,b). Considering the cost of AMI, the net of the average expenditure per household on advanced metering infrastructure and the operational savings, plus the cost of a standard meter, a reasonable approximation of this cost lays between 1 and 2.5 dollars per month. The lower bound at least doubles the gains of TOU and TOU&CPP, warning against the deployment of AMI if these

[17] In many jurisdiction, in particular in California, the wholesale electricity prices differ at different nodes of the transmission network, reflecting network congestion and transmission losses. These nodal prices are called locational marginal prices.

Table 2: Benefits and costs: Changes with respect to flat rate structure

Level	Tariff	Net benefit	Benefit	Cost	Net benefit as a percentage of the cost
		[millions $/year]			[%]
Western Interconnection	RTP	340	-100	-440	1.50
	TOU & CPP	155	-38	-193	0.68
	TOU & DC	137	-22	-159	0.61
	TOU	77	-22	-99	0.34
California's residential sector	RTP	274	-100	-374	5.18
	TOU & CPP	46	-38	-84	0.87
	TOU & DC	176	-22	-198	3.28
	TOU	82	-22	-104	1.53
		[$/year]			
Average per household in California	RTP	22	-8	-30	5.18
	TOU & CPP	4	-3	-7	0.87
	TOU & DC	14	-2	-16	3.28
	TOU	6	-2	-8	1.53

tariffs are the pricing alternatives. Even tough in California AMI is already deployed, unless the cost of AMI decreases, in the long run the state might do as well with simpler rates and a simpler infrastructure.

5.2 Implications for different households

Figure 1 shows the distribution of households across net surplus gains with respect to the flat rate tariff. The figure has four panels, one per each type of household we distinguish in this analysis. In terms of the average gain, the ranking that we observe at the household level in Table 2 also holds when disaggregating per type of household. While RTP remains the most beneficial rate structure, the combination of a TOU and critical peak pricing program is the least favorable. Even some household types would be better off with a simple flat rate tariff than with the TOU&DC structure.

For all rates, the average net surplus gain is different for different households. Those with central air conditioning have a greater average surplus gain when compared to households without this appliance. This difference translates in turn into homeowners having a greater average surplus gain than renters. This happens simply because the proportion of homeowners with central AC is greater than 50%, and the opposite is true for renters. If one consider home-ownership as a proxy for wealth, wealthier customers benefit more from the rate structures we explored in this analysis.

Customers with no AC systems experience small average net surplus gains and this metric has small variance across households. The small increase in net surplus is driven by the elasticity we assume for the baseline consumption. The demands of households with no AC system are inelastic. This small elasticity helps to explain the small variance as well. In addition, other two elements influence the variance. One is the fact that we use one baseline profile for all households. Another is that the LMPs show small variation in the demand nodes of the California portion of the network.

The net surplus gains vary more across households with central AC systems. The variance is driven by differences in the temperature profiles at the different locations. Interestingly, the order in terms of net surplus gains induced by the different temperature profiles is different for each rate we explore. For instance, in the case of real-time pricing locations with higher within-day temperature variance and higher average temperature tend to have greater net surplus gains. One observes a similar pattern when households are enrolled in the TOU&CPP and TOU programs. However, this correlation disappears when the time-of-use rate is combined with demand charges. In this latter case, households located in places where the between-day variance of temperatures is lower tend to benefit the most.

The variance in net surplus for household with AC systems suggest targeting as a strategy for implementing time-varying rates. In particular, the TOU&DC and the RTP program appear to be the most attractive alternatives. However, the non-trivial relationship between surplus gains and temperature profiles suggests that regulators should analyze carefully where to implement these structures.

5.3 On carbon emissions

A final element we explore in this analysis is how the different rate structures impact carbon emissions. A first observation is that not all technologies we consider are economical. Neither coal, nor biomass or nuclear are profitable. Perhaps one could have anticipated this outcome in light of the figures in table 1, which shows that geothermal, solar and wind dominate nuclear,

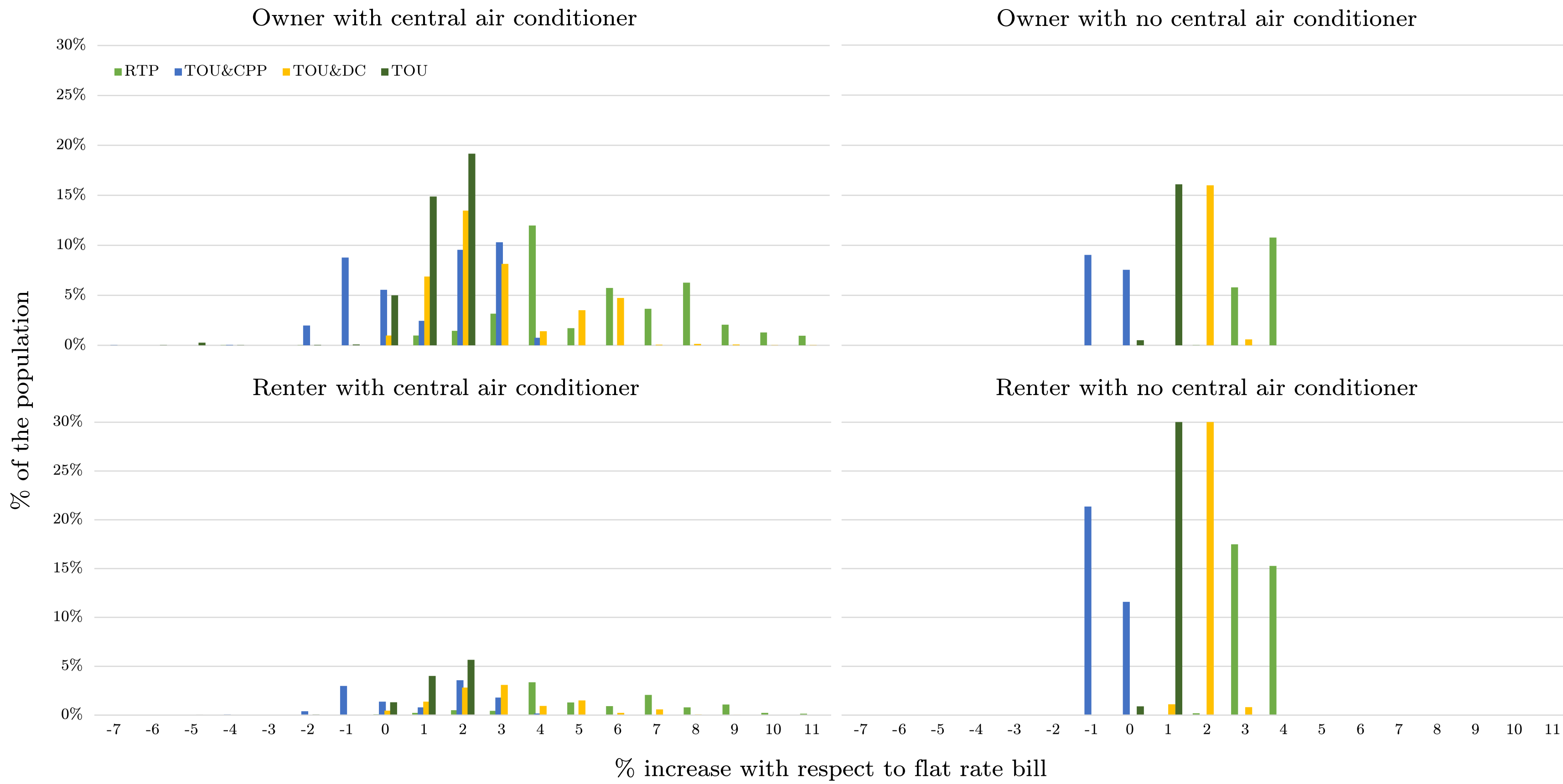

Figure 1: Distribution of households across net surplus gains

biomass and coal. This is not a fair comparison, however, because these resources are of a different nature. While geothermal power plants have important geographic limitations, wind and solar are intermittent resources. Thus, one cannot discard a priori technologies with dominated economic characteristics.

A second observation is that some technologies do not change their total production profile or capacities across the rate scenarios. Consistently, those technologies do not alter their carbon emissions. This technologies include hydro and wind generating power plants. We expected hydro power generation to be invariant because it was exogenous in this analysis. The invariance of wind generation, on the other hand, is a outcome of the model.

Table 3: Capacity, production and emissions changes with respect to FR scenario

Metric	Tariff	Gas adv. CC	Gas adv. CT	Gas conv. CC	Solar	Total	Change relative to FR total
		[MW-year]					%
Capacity	RTP	315	(8,561)	(177)	439	(7,984)	(6.09)
	TOU&CPP	(715)	(3,609)	(754)	2,449	(2,629)	(2.01)
	TOU&DC	(598)	(3,043)	(142)	1,463	(2,320)	(1.77)
	TOU	309	(1,600)	(642)	15	(1,918)	(1.46)
		[GWh/year]					%
Production	RTP	2,197	(2,900)	156	1,620	1,074	0.17
	TOU&CPP	(6,932)	(610)	(1,177)	9,268	549	0.09
	TOU&DC	(4,356)	(793)	(33)	5,536	352	0.06
	TOU	1,693	(216)	(1,135)	55	398	0.06
		[kt of CO2eq/year]					%
Emissions	RTP	725	(1,508)	55	-	(728)	(0.62)
	TOU&CPP	(2,288)	(317)	(412)	-	(3,017)	(2.56)
	TOU&DC	(1,438)	(412)	(12)	-	(1,861)	(1.58)
	TOU	559	(112)	(397)	-	49	0.04

Table 3 shows changes in capacity production and emissions with respect to the reference case. In addition, the table shows total change as a percentage of the Western Interconnection total for the FR scenario. We do not include unprofitable technologies nor technologies that do not vary across rate scenarios.

Agreeing with the basic insight of peak-load pricing, the total installed capacity decreases the most under real-time pricing, followed by TOU&CPP, TOU&DC and TOU. The order in terms of total installed capacity is not the same for total production. Indeed, the figures in Table 3 show that the order is somewhat reversed, with RTP increasing production the most. These changes, however, are a minor fraction of the total production of the Western Interconnection.

Even though production always changes positively, emissions do not. This is true for all rate scenarios with the exception of the TOU case. What happens is that solar production increases considerably in all but the TOU scenario. This suggests long-run complementarities between the demand responsiveness of the residential sector in California and solar generating plants in the Western Interconnection. Interestingly, RTP is not the rate that increases this complementarity the most. The combination between a TOU and a CPP program is the rate alternative that increases solar production and reduces emissions more notoriously.

6 Conclusions

We conduct an analysis of rate design in California's residential electricity sector. Beyond the applied insights, we contribute with a modeling framework to evaluate rate structures. The framework gives an important step towards bridging top down models of pricing and investment with bottom up models of household behavior. Building upon the theory of peak-load pricing, we illustrate how to modify the basic model to accommodate household heterogeneity, as well as the adoption of distributed energy resources and more general types of rate structures.

Our analysis seeks to quantify efficiency, distributional and environmental implications of rate structures that are plausible alternatives for California's future residential sector. The analysis shows that the average gains of implementing time-varying rates with respect to a simple flat rate program are rather mild, even in the real-time pricing scenario. Our results also show that factors such as the presence of an air conditioning system and the exterior temperature profile can have a meaningful impact on the surplus gains that different rates generate on households. These two results combined suggest that defaulting all residential customers into a time-of-use rate structure, which is the current path California is following for the residential sector, may not be an ideal strategy. Targeting different rates to households with different appliance stocks and in different locations will likely be a superior policy.

A Appendix

A.1 Distribution of technologies

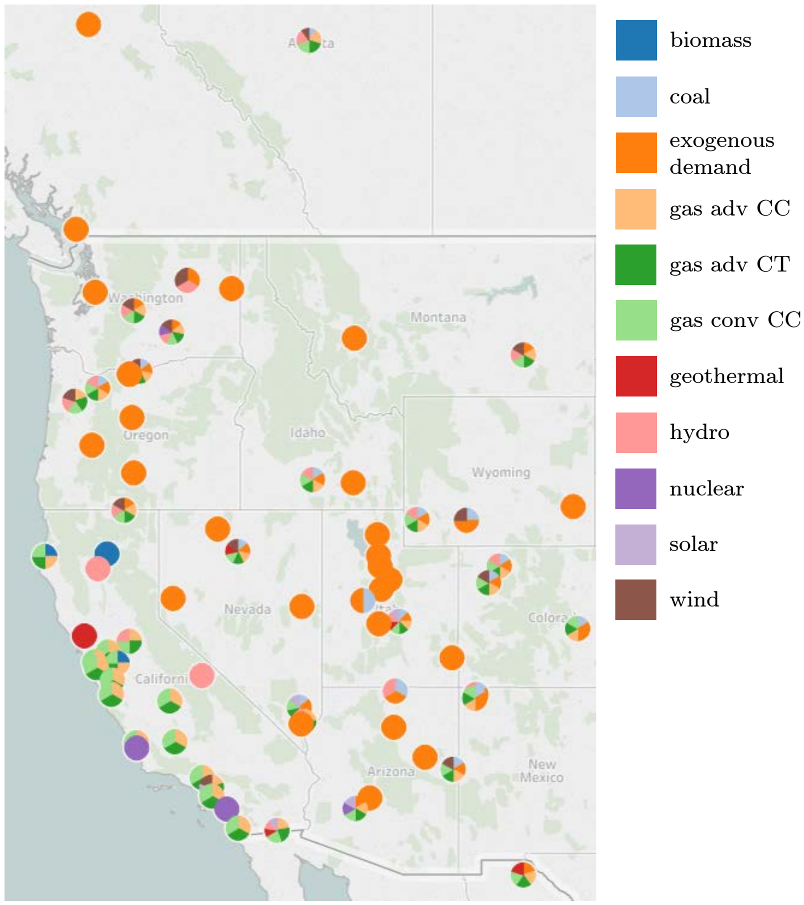

Figure 2: Geographic distribution of generating technologies

A.2 More examples of rates and devices

This appendix provides two additional examples. One describes how to implement an (IBP) structure and the other how to model a battery storage system. We keep the notation in section 3 and define Z_m as a square matrix of zeros of m by m.

A model of an IBP structure. Under this rate a customer pays a volumetric charge p_n if its total consumption falls within the tier $n \in \{1, \ldots, N\}$, that is, if $\bar{d}^\top e_{|T|} \in [q_{n-1}, q_n]$. To model this let the auxiliary variable $\hat{d}_n$ denote the consumption on tier n, and define

$$\tilde{b} := \begin{bmatrix} 0 \\ q \end{bmatrix}, \ \tilde{A} := \begin{bmatrix} e_{|T|}^\top & -e_N^\top \\ z_N z_{|T|}^\top & I_N \end{bmatrix}, \ M := \begin{bmatrix} z_{|T|} z_N^\top \\ I_N \end{bmatrix}.$$

Additionally, define the prices constraint set as follows

$$\mathcal{P} := \left\{ p \in \mathbb{R}_+^N : p_n \geq p_{n-1} \, \forall n > 1 \right\}.$$

Modeling the dynamics of a battery storage system. A battery of *roundtrip efficiency*[18] ζ changes its state of charge in t, s_t, in response to an energy charge, d_t^1, or discharge, d_t^2, according to the following relationship,

$$s_{t+1} = s_t + d_t^1 \sqrt{\zeta} - d_t^2 \frac{1}{\sqrt{\zeta}}.$$

In addition, the battery has a maximum storage capacity s_{max}, and a minimum state of charge s_{min}. The energy charge and discharge also have boundaries: the maximum charging and discharging rates of the battery, d_{max}.

To include this model in our framework define $\bar{d} := [d^1, d^2]$, and

$$b_{storage} := \begin{bmatrix} (s_{max} - s_1) \cdot e_{|T|} \\ -(s_{min} - s_1) \cdot e_{|T|} \\ d_{max} \cdot e_{|T|} \\ z_{|T|} \\ d_{max} \cdot e_{|T|} \\ z_{|T|} \end{bmatrix}, \ A_{storage} := \begin{bmatrix} B\sqrt{\zeta} & -B\frac{1}{\sqrt{\zeta}} \\ -B\sqrt{\zeta} & B\frac{1}{\sqrt{\zeta}} \\ I_{|T|} & Z_{|T|} \\ -I_{|T|} & Z_{|T|} \\ Z_{|T|} & I_{|T|} \\ Z_{|T|} & -I_{||T|} \end{bmatrix},$$

with B a lower triangular matrix of ones, and s_1 the state of charge at the beginning of the first period.

[18]Round trip efficiency is the maximum output per unit of energy input into the battery storage system.

A.3 Deriving a linear model for a TCL

The model of the thermostatically controlled load we use in this paper is similar to the one in (Mathieu et al., 2015). It originates from the thermodynamic identity,

$$C \cdot \dot{\theta}(t) = d(t) \cdot \eta - \frac{\theta(t) - \tilde{\theta}(t)}{R},$$

in which C is the thermal capacity of the interior space, η is the coefficient of performance of central air conditioner and R is thermal resistance of the household. This identity simply sates that the variation in heat in the interior space is equal to the heat extracted by the air conditioner plus the heat that enters the space as a result of a temperature gradient. The solution of this equation is

$$\theta(t) = \theta_0 e^{\frac{-t}{CR}} + e^{\frac{-t}{CR}} \int_0^t e^{\frac{s}{CR}} \kappa(s) ds,$$

where $\kappa(s) := \left[\tilde{\theta}(t) + \eta R d(s))\right] / CR$. Thus, we can write the temperature in $t + \Delta$ as

$$\theta(t + \Delta) = e^{\frac{-\Delta}{CR}} \theta(t) + e^{\frac{-t-\Delta}{CR}} \int_t^{t+\Delta} e^{\frac{s}{CR}} \kappa(s) ds.$$

Assuming that the exterior temperature and the power consumption are constant over the interval $[t, t + \Delta]$ and equal to $\tilde{\theta}(t)$ and $d(t)$, $\kappa(\cdot)$ is also constant. Furthermore, If we normalize Δ to 1, we have that

$$\theta_{t+1} = a\theta_t + (1 - a)\tilde{\theta}_t + (1 - a)\eta R d_t, \tag{15}$$

where $a := e^{\frac{-\Delta}{CR}}$. Equation (15) shows a linear relationship between the consumption of the central air conditioner and the interior temperature. Considering as border condition $\theta_{|T|+1} = \theta_1$, which is consistent with the assumption of similar consecutive days, we can express (15) in matrix form as in (9). The explicit formulae follows

$$\Theta_1(\xi) = \frac{(1-a)\eta R}{1-a^{|T|}} \begin{bmatrix} a^{|T|-1} & a^{|T|-2} & \cdots & \cdots & \cdots & a & 1 \\ 1 & a^{|T|-1} & \cdots & \cdots & \cdots & a^2 & a \\ a & 1 & \cdots & \cdots & \cdots & a^3 & a^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ a^{t-2} & a^{t-3} & \cdots & 1 & \cdots & a^t & a^{t-1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a^{|T|-2} & a^{|T|-3} & \cdots & \cdots & \cdots & 1 & a^{|T|-1} \end{bmatrix} \text{ and } \theta_2(\xi, \tilde{\theta}) = \frac{1}{\eta R}\Theta_1(\xi)\tilde{\theta}.$$

A.4 Including a transmission network

In order to model a transmission network, we take the standard approach in capacity expansion, e.g. (Sauma and Oren, 2006). In principle, a non-liner system, known as the *power flow equations*, describe how electricity circulates in the transmission network. The non-linearity of this system, however, adds significant complexity to the capacity expansion problem. In order to keep this model tractable researchers approximate the power flow equations. In this paper we use the *dc approximation*, which corresponds to a linearization of the original system. The dc approximation considers a vector w of net imports at each node of the network, a matrix $PTDF$ that maps net imports to flows across each of the network's edges, and a bounds for the flows, (f^+, f^-). There are two conditions associated to the approximation. One expresses constraints on flows,

$$f^+ \geq PTDFw \geq f^-,$$

the other establishes that the summation off all net imports is zero. To incorporate this network model in our framework, modify pricing problem adding these conditions. In addition, modify (13) as follows,

$$\sum_h \alpha_h d_{h\omega n} = \sum_{k \in K} y_{\omega k n} + w_{\omega n} \colon \lambda_\omega, \tag{16}$$

where n indexes the nodes of the network.

References

Allcott, H. (2011). Rethinking real-time electricity pricing. *Resource and Energy Economics*, 33(4):820–842.

Beaudin, M. and Zareipour, H. (2015). Home energy management systems: A review of modelling and complexity. *Renewable and Sustainable Energy Reviews*, 45:318 – 335.

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

Borenstein, S. (2006). Customer risk from real-time retail electricity pricing: Bill volatility and hedgability. Technical report, National Bureau of Economic Research.

Borenstein, S. (2013). Effective and Equitable Adoption of Opt-In Residential Dynamic Electricity Pricing. *Review of Industrial Organization*, 42(2):127–160.

Borenstein, S. (2015). The Private Net Benefits of Residential Solar PV: The Role of Electricity Tariffs, Tax Incentives and Rebates. Working Paper 21342, National Bureau of Economic Research.

Borenstein, S. and Holland, S. (2005). On the efficiency of competitive electricity markets with time-invariant retail prices. *The RAND Journal of Economics*, 36(3):pp. 469–493.

Brattle, E3, BEAR, and Aspen (2016). The impacts of a Regional ISO-Operated Power Market in California. Technical report, California Independent System Operator.

Brown, S. J. and Sibley, D. S. (1986). *The theory of public utility pricing*. Cambridge University Press.

Carlton, D. W. (1977). Peak load pricing with stochastic demand. *The American Economic Review*, 67(5):pp. 1006–1010.

Castro, F. A. and Callaway, D. S. (2016). Optimal rate design in modern electricity sectors. Manuscript submitted for publication.

Castro, F. A., Lara, J., and Callaway, D. S. (2016). A mathematical programming approach to utility pricing. Manuscript in preparation.

Caves, D. W., Christensen, L. R., and Herriges, J. A. (1984). Consistency of residential customer response in time-of-use electricity pricing experiments. *Journal of Econometrics*, 26(1):179–203.

Chao, H.-p. (1983). Peak load pricing and capacity planning with demand and supply uncertainty. *The Bell Journal of Economics*, 14(1):pp. 179–190.

Chao, H.-p. (2011). Efficient pricing and investment in electricity markets with intermittent resources. *Energy Policy*, 39(7):3945 – 3953. Special Section: Renewable energy policy and development.

Crew, M. A., Fernando, C. S., and Kleindorfer, P. R. (1995). The theory of peak-load pricing: A survey. *Journal of Regulatory Economics*, 8(3):215–248.

Crew, M. A. and Kleindorfer, P. R. (1976). Peak load pricing with a diverse technology. *The Bell Journal of Economics*, 7(1):pp. 207–231.

DOE (2012a). Demand reductions from the application of advance metering infrastructure, pricing programs and customer-based systems. Technical report, US Department of Energy.

DOE (2012b). Operations and maintenance savings from advance metering infrastructure. Technical report, US Department of Energy.

Drèze, J. H. (1964). Some postwar contributions of French economists to theory and public policy: With special emphasis on problems of resource allocation. *The American Economic Review*, 54(4):pp. 2–64.

EIA (2016). Annual Energy Outlook 2016, with Projections to 2040. Technical report, US Energy Information Administration.

Eid, C., Reneses Guilln, J., Fras Marn, P., and Hakvoort, R. (2014). The economic effect of electricity net-metering with solar PV: Consequences for network cost recovery, cross subsidies and policy objectives. *Energy Policy*, 75:244–254.

Faruqui, A. and Sergici, S. (2010). Household response to dynamic pricing of electricity: A survey of 15 experiments. *Journal of Regulatory Economics*, 38(2):193–225.

Faruqui, A. and Sergici, S. (2011). Dynamic pricing of electricity in the mid-Atlantic region: econometric results from the Baltimore gas and electric company experiment. *Journal of Regulatory Economics*, 40(1):82–109.

Herter, K. (2007). Residential implementation of critical-peak pricing of electricity. *Energy Policy*, 35(4):2121–2130.

Huijben, J. and Verbong, G. (2013). Breakthrough without subsidies? {PV} business model experiments in the netherlands. *Energy Policy*, 56:362 – 370.

Joskow, P. and Tirole, J. (2006). Retail electricity competition. *The RAND Journal of Economics*, 37(4):799–815.

Joskow, P. and Tirole, J. (2007). Reliability and competitive electricity markets. *The Rand Journal of Economics*, 38(1):60–84.

Mathieu, J. L., Dyson, M. E., and Callaway, D. S. (2015). Resource and revenue potential of california residential load participation in ancillary services. *Energy Policy*, 80:76 – 87.

Palmgren, C., Stevens, N., Goldberg, M., Bames, R., and Rothkin, K. (2010). California Residential Appliance Saturation Survey. Technical Report CEC-200-2010-004, California Energy Commission.

Panzar, J. C. (1976). A neoclassical approach to peak load pricing. *The Bell Journal of Economics*, 7(2):pp. 521–530.

Phillips Jr, C. F. (1993). *The regulation of public utilities*. Public Utilities Reports, Incorporated, 1993, third edition.

Price, J. E. and Goodin, J. (2011). Reduced network modeling of WECC as a market design prototype. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1–6. Conference Proceedings.

Reiss, P. C. and White, M. W. (2005). Household Electricity Demand, Revisited. *The Review of Economic Studies*, 72(3):853–883.

Sauma, E. E. and Oren, S. S. (2006). Proactive planning and valuation of transmission investments in restructured electricity markets. *Journal of Regulatory Economics*, 30(3):261–290.

Schleicher-Tappeser, R. (2012). How renewables will change electricity markets in the next five years. *Energy Policy*, 48:64–75.

Stanton, T. (2015). Distributed Energy Resources: Status Report on Evaluating Proposals and Practices for Electric Utility Rate Design. Status Report 15-08, National Regulatory Research Institute, 8611 Second Avenue, Suite 2C, Silver Spring, MD 20910.

Steiner, P. O. (1957). Peak loads and efficient pricing. *The Quarterly Journal of Economics*, 71(4):pp. 585–610.

Wei, M., Nelson, J. H., Greenblatt, J. B., Mileva, A., Johnston, J., Ting, M., Christopher Yang, Jones, C., McMahon, J. E., and Kammen, D. M. (2013). Deep carbon reductions in California require electrification and integration across economic sectors. *Environmental Research Letters*, 8(1):014038.

Wilcox, S. and Marion, W. (2008). User manual for TMY3 data sets. Technical Report NREL/TP-581-43156, National Renewable Energy Laboratory.

Zöttl, G. (2010). A framework of peak load pricing with strategic firms. *Operations Research*, 58(6):1637–1649.