MIT CEEPR
Center for Energy and Environmental Policy Research

Working Paper Series

The Efficiency of Dynamic Energy Prices

Andrew J. Hinchberger, Mark R. Jacobsen, Christopher R. Knittel, James M. Sallee, and Arthur A. van Benthem

OCTOBER 2024

CEEPR WP 2024-15

Working Paper Series.

Since 1977, the Center for Energy and Environmental Policy Research (CEEPR) has been a focal point for research on energy and environmental policy at MIT. CEEPR promotes rigorous, objective research for improved decision making in government and the private sector, and secures the relevance of its work through close cooperation with industry partners from around the globe. Drawing on the unparalleled resources available at MIT, affiliated faculty and research staff as well as international research associates contribute to the empirical study of a wide range of policy issues related to energy supply, energy demand, and the environment.

An important dissemination channel for these research efforts is the MIT CEEPR Working Paper series. CEEPR releases Working Papers written by researchers from MIT and other academic institutions in order to enable timely consideration and reaction to energy and environmental policy research, but does not conduct a selection process or peer review prior to posting. CEEPR's posting of a Working Paper, therefore, does not constitute an endorsement of the accuracy or merit of the Working Paper. If you have questions about a particular Working Paper, please contact the authors or their home institutions.

The Efficiency of Dynamic Electricity Prices

Andrew J. Hinchberger,[1] Mark R. Jacobsen,[2,6] Christopher R. Knittel,[3,6]
James M. Sallee,[4,6] Arthur A. van Benthem[5,6*]

September 12, 2024

Abstract

The marginal cost of electricity fluctuates hour-by-hour, yet retail customers typically face flat prices. Using data from all seven US wholesale markets and a new method to evaluate alternative rates set in advance that accounts for equilibrium price effects, we estimate efficiency gains from time-varying price schedules that better align price with cost. We have three main results. First, time-of-use rates and critical-peak pricing, the two most common time-varying rate plans, each correct about 10% of mispricing. Second, complex rate structures based on historical prices often backfire. Third, real-time pricing with price ceilings can capture most potential efficiency gains while limiting customer risk.

Keywords: electricity, time-of-use pricing, critical-peak pricing, real-time pricing, efficiency

JEL: L94, L97, Q41, Q48

*We thank Susanna Berkouwer, Severin Borenstein, Vivek Farias, Jim Hines, Stephen Jarvis, Akshaya Jha, Ryan Kellogg, Angela Pachon, and seminar participants at the MIT EPS LCEC Workshop, the 2024 EAERE conference, Energy Camp at UC Berkeley, the Haas POWER conference, Carnegie Mellon University, the University of Chicago, the University of Pennsylvania, and the University of Southern California for helpful comments and suggestions. We thank Kamen Velichkov and Jay Theriault for their excellent assistance with the research. Van Benthem thanks the Kleinman Center for Energy Policy, Analytics at Wharton, the Mack Institute, the National Science Foundation (SES-2117158), and the Wharton Dean's Research Fund for generous support.

[1]Department of Economics, Northwestern University. [2]Department of Economics, UC San Diego. [3]Sloan School of Management, Center for Energy and Environmental Policy Research, and MIT Climate Policy Center, Massachusetts Institute of Technology. [4]Haas School of Business, UC Berkeley. [5]The Wharton School, University of Pennsylvania. [6]National Bureau of Economic Research.

1 Introduction

This paper is concerned with the efficiency consequences of the mispricing of a key commodity, retail electricity. Inefficient pricing is one of the cornerstone topics in economics, and there are active literatures concerned with its persistence due to factors including market power (e.g., De Loecker, Eeckhout, and Unger 2020), environmental externalities (e.g., Fowlie, Reguant, and Ryan 2016), speculative bubbles (e.g., Scheinkman and Xiong 2003), coarse pricing (e.g., Stevens 2020), and rate of return regulation (e.g., Cicala 2015). In the space of public utilities, the efficiency consequences of inefficient pricing have been studied for water (e.g., Timmins 2002), natural gas (e.g., Davis and Kilian 2011), and electricity (e.g., Borenstein and Bushnell 2022a).

We study inefficient retail electricity pricing that results when customers face time-invariant prices, whereas hourly costs vary substantially. We quantify the inefficiencies of the mispricing of electricity and estimate the potential efficiency gains from a wide range of possible pricing reforms using detailed measures of marginal costs taken from all seven wholesale power markets in the US over two decades.

We estimate that inefficiencies from mispricing are on the order of $2 billion annually. Relatively simple, feasible time-of-use rates and critical-peak pricing—which we refer to as dynamic or time-varying pricing—can each reduce that deadweight loss by about 10%. On the other hand, more complicated pricing schemes based on historical data often backfire because they lead to poor out-of-sample performance. In a novel machine learning exercise, we demonstrate that the vast majority of the efficiency gains come from a time-of-use plan having just two rates. Our results provide guidance to rate setters regarding both the magnitude of potential gains from time-varying prices and the types of reforms that have the most potential. Some of our conclusions may be surprising and run counter to industry trends: we find little gain to increasing time-of-use complexity, and we find robust evidence that basing critical-peak rates on more timely cost information is critical for efficiency. Finally, we show that exposing consumers to real-time wholesale prices while protecting them from extreme price risks using price caps reduces mispricing much more dramatically than either time-of-use or critical-peak pricing.

The efficiency gains of time-varying electricity rates have long been of interest to economists, but the topic is more important now than ever before for two reasons. First, more complex prices are now feasible due to the large-scale adoption of advanced metering infrastructure, or smart meters, that enable utilities to receive real-time information on consumer demand and to send real-time price signals to consumers. Second, gains from better pricing are likely rising due to market conditions. At the heart of virtually every clean energy transition plan is the goal to rapidly build clean power generation and electrify end uses. Electrification of transportation and buildings, as well as explosive growth in data centers, has led industry to expect a step change in load growth (North American Electric Reliability Corporation 2023; EFI Foundation 2024). Expansion of renewable energy increases price volatility and raises the value of accurate pricing (Imelda, Fripp, and Roberts Forthcoming). Moreover, electric vehicles, heat pumps, and data centers are especially likely to respond to time-varying prices (Bailey, Brown, Myers, Shaffer, and Wolak 2024). Failing to set efficient prices for these growing sources of demand would be a critical missed opportunity.

Economic theory advocates for "real-time pricing," under which final electricity consumers would face prices that vary at a high frequency according to the marginal cost of providing electricity in a given moment. This allows market prices to signal the scarcity value of capacity, as well as variation in the marginal cost of operations within an existing set of heterogeneous generators (Boiteux 1949; Joskow 1976; Borenstein and Holland 2005).

Real-time pricing represents one endpoint on a spectrum, with the other endpoint being entirely flat pricing schemes, where electricity rates are the same every hour. Time-of-use (TOU) and critical-peak pricing (CPP) schemes lie between these extremes. TOU prices specify rates for different hours of the day, days of the week, or seasons. CPP rates add a cost premium to specific blocks of time (referred to as 'events') when a system is expected to be near capacity based on the current or expected price in the market. Many hold to the view that real-time pricing is too complex, confusing, or risky to impose on residential customers, and so a central question of the literature has been whether feasible TOU and CPP prices can substantially improve the alignment between cost and price. Utilities have been moving in this direction, and many are now creating more and more complex rates.

We evaluate the potential gains from these commonly-applied rate structures, and also study an alternative in which consumers face real-time pricing, but the prices are subject to aggressive price ceilings to protect them from extreme pricing events. The basis of our analysis is equilibrium prices from wholesale electricity markets, which provide a granular measure of the marginal cost of power in each hour. The basic idea of the paper is to ask how well alternative retail pricing schemes can align the prices customers face with these marginal costs. Our data span almost a million hourly prices. Where we extend our analysis to cover spatial variation within each market, we have on the order of a billion prices.

We start our analysis by fitting alternative rate schedules onto historical wholesale prices using standard regression tools. In every market and year, we show the goodness of fit of alternative rate schemes with varying levels of complexity using the in-sample R^2. This provides a descriptive analysis of mispricing, and it also provides an approximation of efficiency costs following Jacobsen, Knittel, Sallee, and van Benthem (2020). Intuitively, the R^2 measures the efficiency gains achieved by a given rate schedule as compared to a flat-rate baseline, as a percentage of the efficiency gains achievable by real-time pricing over a flat-rate baseline.

We adapt our methodology to address two issues: out-of-sample fit and equilibrium effects of alternative pricing schemes. In practice, rates are set based on backward-looking data. We develop an out-of-sample fit statistic that allows us to quantify mispricing resulting from setting rates based on historical data benchmarked against a flat price based on the same historical data. Our method also accounts for changes in equilibrium prices and quantities that would result from a pricing reform. We estimate cost curves from industry data in each market and demand curves based on load-response estimates from the literature. This provides an improved estimate of the efficiency effects of informationally-achievable policies compared to the simple in-sample R^2. Our revised metric, the "equilibrium basis renormalized R^2", allows us to approximate efficiency gains of many possible alternatives in a simple and transparent fashion while accounting both for out-of-sample prediction and for the equilibrium price effects of alternative pricing structures.

We show that realistic TOU and CPP plans capture only a modest fraction, each on the order of 10% when averaged across markets and years, of the mispricing created by flat rates. Increasingly complex TOU rates—for example, rates that vary flexibly by month of the year and hour of the day—can do much more to align prices with marginal cost in the backward-looking data, but when we recognize that rates must be set ahead of time, the value of complexity reverses. More complex rates trained on historical data frequently fit future prices worse than simpler rates. In doing so, we empirically characterize a textbook lesson from the statistical learning literature: complicated functional forms have the potential to better explain variation but are subject to overfitting and thus can perform worse out-of-sample (James, Witten, Hastie, and Tibshirani 2021). This leads to a concave, non-monotonic relationship between out-of-sample performance and model complexity with an interior optimum. In the case of TOU pricing, our data suggest this optimum typically falls below 10 pricing tiers per year.

We evaluate CPP plans under a range of assumptions and calculate the marginal return to adding events across years and markets. What matters most for CPP efficiency is the ability of the utility to set CPP rates close to the events. In practice, CPP rates are typically set in advance during rate hearings and are often set for multiple years; thus, they are not tailored to specific peak events. We find that CPP rates set in this manner fail to meaningfully improve efficiency, whereas pricing based on day-ahead prices for 20 yearly events would capture 10% of the efficiency gain that real-time pricing could achieve.

A pricing-policy counterfactual in which customers face real-time pricing paired with price caps achieves considerably larger efficiency gains than either TOU or CPP. Even in scenarios where the price cap binds in 25% of hours, real-time pricing with caps achieves approximately two-thirds of the efficiency gains of full real-time pricing. If the price cap is set higher, such that it only binds in 5% of hours, the efficiency gain is over 90%. These price caps are not extreme—between \$28 and \$72 per megawatt hour depending on the power market—yet deliver sizable efficiency gains. These efficiency gains are premised on the notion that customers can respond to rapidly changing prices, which might be most realistic for automated load, as we discuss below. As such, fully realizing the gains from real-time pricing may only be possible in the future.

As long as a move towards (partial) real-time pricing is unlikely in the near future, our results suggest that a combination of relatively simple TOU rates and a limited number of yearly CPP days with rates set based on day-ahead prices strikes a balance between simplicity, price uncertainty, and efficiency. An additional observation is that the gains from TOU and CPP policies are mostly additive: when implemented in combination, the two policies deliver 17-20% of the efficiency gain under real-time pricing. This is because CPP targets idiosyncratic price extremes, whereas TOU targets predictable cyclical variation in prices over days, weeks, or seasons. While our analysis suggests that efficiency gains are achievable, much of the pricing reform being rolled out by utilities today—often focusing on more granular TOU schedules and CPP schemes that use uniform, pre-determined prices—may be relatively ineffective.

These conclusions complement findings in other areas of economics where simpler prices are a constrained optimum. DellaVigna and Gentzkow (2019) find that US retailers use uniform pricing across stores and that this has substantial efficiency costs. They suggest that managerial and customer costs of price differentiation may explain the reliance on simpler pricing schedules. Chu, Leslie, and Sorensen (2011) find that simple subsets of complex bundled pricing are nearly optimal among firms selling multiple retail products. In our setting, this complexity might be eschewed because of the difficulty in setting rates in advance.

Our findings also speak to a distinction suggested by Borenstein (2005b) between rate granularity (how many different pricing tiers are there?) versus timeliness (how far in advance must rates be set?). A big-picture question for rate designers is which dimension is most valuable. Our results point firmly towards a greater importance of timeliness. Additional granularity in TOU schemes quickly loses value—and can easily backfire—if granular prices must be set ahead of time. Setting TOU rates ahead of time creates two sources of mispricing: not only will there be errors in matching the *relative* prices between tariff levels (e.g., peak vs. off-peak), but also the *average* price will be wrong ex-post. We show that this second inefficiency dominates the former, and therefore, what matters most in CPP and TOU schemes is the ability to set rates close to the actual events. It also implies that the efficiency advantage from real-time pricing not only stems from matching relative price differentials but also from charging the right price on average.

Time-varying electricity pricing schemes have been extensively piloted and fully deployed in some locations (Faruqui and Sergici 2010; Faruqui and Tang 2021), with several papers evaluating consumer response and efficiency gains from critical-peak pricing (Wolak 2007 2011a; Ito, Ida, and Tanaka 2018; Blonz 2022; Ito, Ida, and Tanaka 2023), time-of-use rates (Aigner 1984; Train and Mehrez 1994; Enrich, Li, Mizrahi, and Reguant 2024), and real-time pricing pilots (Allcott 2011; Andersen, Hansen, Jensen, and Wolak 2017; Fabra, Rapson, Reguant, and Wang 2021). Closest to our work is a prior literature that studies the benefits of alternative schemes using market simulations. Our conclusion that TOU and CPP rates can capture only a modest fraction of the potential gains of real-time pricing is similar to Borenstein (2005a), Borenstein (2005b), Borenstein and Holland (2005), and Holland and Mansur (2006), but those prior papers either used simulations based on hypothetical generator costs or deployed wholesale market data from only one market. In contrast, we examine all major US power markets over twenty years. We also differ by using a new, renormalized R^2 metric that accounts for out-of-sample prediction, the equilibrium price effects of policies, and is directly tied to economic welfare.[1] Our paper sheds new light on why complex, high-dimensional TOU policies have limited potential and on the importance of errors in average prices when prices are set far in advance.[2] We additionally extend our equilibrium model to consider inter-temporal load-shifting where demand can move across hours.

Before proceeding, we mention here three key caveats associated with the starting presumption of this paper, which is that we want retail rates to reflect the real-time variation in wholesale electricity prices as a theoretical ideal. First, we abstract from fixed cost recovery, or other cost components, that are included in customer bills, as well as from environmental externalities. As pointed out by Borenstein and Bushnell (2022ab), these issues can also create an important wedge between marginal costs and benefits, so passing through wholesale prices to final customers can have more complex welfare interpretations when these other price wedges exist. Our analysis demonstrates the potential to align prices with energy market costs, which is ideal when other inefficiencies are resolved through separate rate reforms, like emissions pricing.

Second, in reality, electricity customers may have limited rationality and may find real-time price variation overwhelming. In particular, prior research has called into question the ability of customers to understand marginal prices for electricity (Ito 2014), though evidence suggests that improved understanding can be taught (Kahn and Wolak 2013). Jessoe and Rapson (2014) show experimentally that it is essential to provide readily accessible information about price surges in order to generate a response. These types of findings call into question whether real-time variation in prices would really be as efficient as standard theory suggests, and we interpret this as yet another reason why simple dynamic pricing schemes may be preferred to

[1] Also related is Schittekatte, Mallapragada, Joskow, and Schmalensee (2024), which studies recent data from three wholesale power markets and reaches a more optimistic conclusion about TOU rates. The difference stems from their focus on getting rank correlation of prices right within the day, consistent with a world of fully automated demand that effortlessly reshuffles. In contrast to their approach, we link our measures of fit directly to welfare measures and account for temporal load-shifting using recent empirical estimates.

[2] This aligns with the finding in Holland and Mansur (2006), which is based on two years of data from one market, that monthly flat rates can outperform TOU rates, which they describe as surprising. Our conclusion that getting average prices right in a period is more important than matching patterns, which is based on additional markets, years, and alternative rates, as well as a different method to assess out-of-sample efficiency, provides an explanation for their finding.

complex ones in most settings. Existing evidence shows that automation increases demand response, as one would expect (Gillan 2017; Bailey, Brown, Shaffer, and Wolak 2023; Blonz, Palmer, Wichman, and Wietelman Forthcoming). As home automation expands, it may be feasible to introduce additional complexity, but our results suggest this must leverage more timely generation cost data (i.e., from wholesale prices), not just more rate granularity.

Third, we do not consider the distributional impacts of alternative rate designs, focusing only on the aggregate efficiency implications. Shifts to alternative rate structures are likely to create winners and losers among different customer groups. Burger, Knittel, Perez-Arriaga, Schneider, and vom Scheidt (2020) and Cahana, Fabra, Reguant, and Wang (2024) consider the distributional impacts of real-time pricing across income groups but do not consider the implications of TOU and CPP programs.

This paper proceeds as follows. Section 2 gives background about electricity pricing in the United States. Section 3 describes our data. Section 4 discusses our methodology. Section 5 presents the results for a wide range of potential pricing schemes. Section 6 presents a set of refinements and extensions. Section 7 concludes.

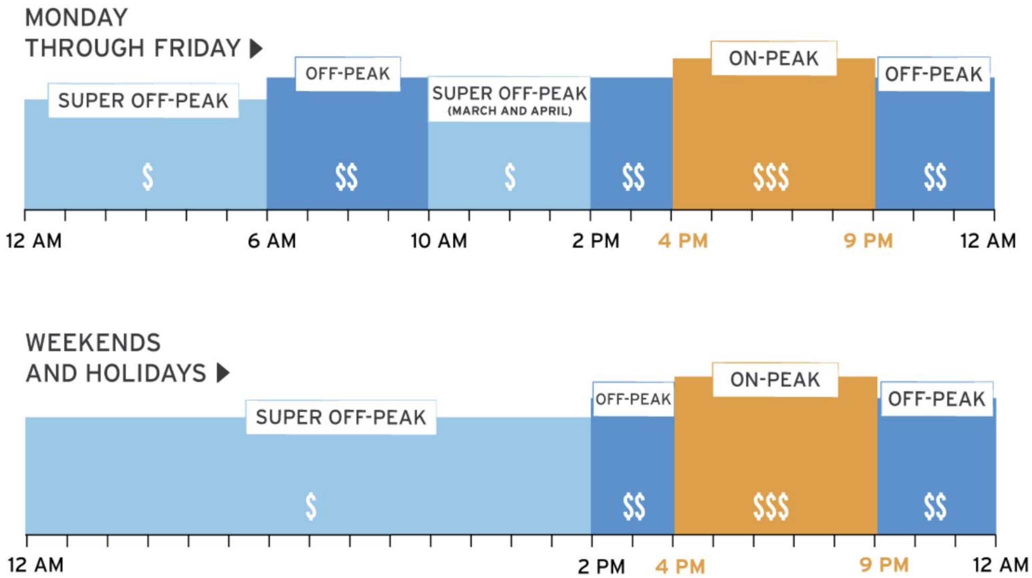
2 Background and Institutions

In many cases, electricity consumers pay the same, or nearly the same, price to their electric utility regardless of when they consume electricity. Data from the EIA suggest that, nationally, 90-95% of end-users face a flat price schedule (see Appendix Figure A.1 for a sector breakdown and trends over time). However, the marginal cost of electricity generation varies substantially across time. The generation required for low levels of demand is typically met by solar, wind, and inexpensive baseload power plants, resulting in a low marginal cost of electricity. In contrast, at high levels of demand, higher-cost peaker plants produce the marginal megawatt hour of electricity. As a result, the marginal cost of electricity can frequently vary by multiple orders of magnitude within a given day.

A wave of deregulation created wholesale electricity markets that yield prices which achieve a real-time balance between supply and demand in ways that follow the textbook economic recipe (Borenstein and Bushnell 2015). These prices, however, are rarely passed through to final consumers who instead face prices that represent costs averaged over an extended time period, often a year or more. This creates an inefficiency because end users purchase power at prices that are well below the cost of provision in some hours and well above that cost in others.

Historically, it was infeasible to measure electricity consumption in real time for each end user, so this mispricing was a necessary compromise. As discussed above, with the advent and roll-out of computerized "smart" electricity meters, high-frequency measurement at the

Figure 1: A Time-of-Use Pricing Schedule from San Diego Gas & Electric



Note: San Diego Gas & Electric "TOU-DR1" price schedule (the default residential electricity pricing plan) as of May 2024. Summer prices (three of six prices) are displayed. New versions of the schedule (changes to at least one of the price levels) were released 21 times between January 2018 and May 2024. Source: https://www.sdge.com/total-electric-rates.

customer level is already a reality in most parts of the United States. Even so, real-time electricity pricing has been met with considerable resistance from utilities and regulators, who fear that consumers will complain about price surges and unpredictable bills. As one example, the skyrocketing wholesale-market prices in Texas in the winter of 2021 made these concerns especially salient, as a small group of customers who had opted into real-time pricing were faced with electricity bills in the tens of thousands of dollars.[3]

As an alternative to real-time pricing, utilities have begun to set increasingly intricate time-varying pricing schemes (Badtke-Berkow, Centore, Mohlin, and Spiller 2015; Faruqui and Tang 2021). As an illustration, San Diego Gas and Electric uses a time-of-use pricing schedule with four different peak periods that change depending on the month (see Figure 1). While time-varying, these tariffs are set in advance; pricing is not adjusted based on realized market conditions.

In addition, utilities are implementing policies such as critical-peak pricing, in which they inform customers, typically a day in advance, about significantly higher rates the following day. This usually happens during the hottest days of the summer when the grid is expected to operate near capacity constraints. Most often, CPP price levels and hours of the day are

[3] https://www.nytimes.com/2021/02/20/us/texas-storm-electric-bills.html

set in advance (for example, using data from the previous year), but events are called flexibly. Southern California Edison, for example, calls 12 to 15 annual CPP events with each event lasting between 4pm and 9pm on a given day, and it announces these critical peak events a day in advance.[4] The number of CPP events called each year as well as the rate charged during CPP events varies across utilities.

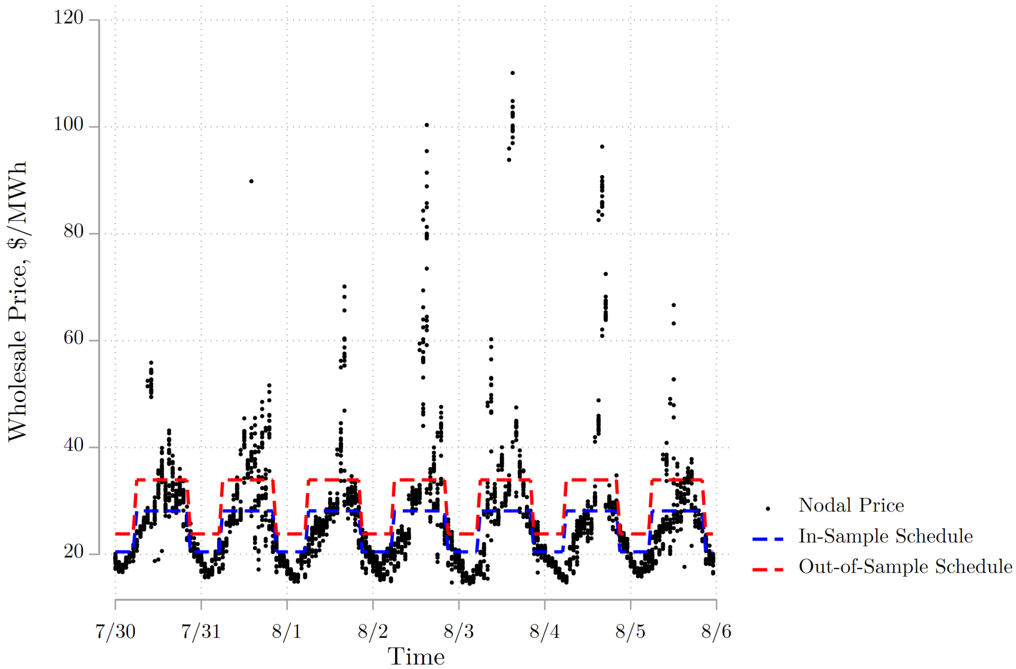
The basic idea of TOU and CPP schemes is to better align wholesale and retail prices. However, TOU and CPP schemes, by construction, create a limited number of unique prices and/or affect a limited number of hours. If wholesale market price variation is large and unpredictable, then simple TOU and CPP schemes may be too blunt of an instrument to deliver the expected benefits.

Figure 2 provides an example to illustrate the limitations of a simple daily on-peak vs. off-peak pricing scheme. In the graph, each dot is a wholesale market price from a node in the PJM market for one summer week in 2019. The dotted blue line shows the best-fit two rate schedule based on in-sample data for the entire year, with the peak hours defined as 6 am to 8 pm. This represents an improvement over a flat rate, but it falls far short of a true real-time rate. By using the in-sample average for the year, we will get the average price right by construction, but the peak-pricing scheme does not capture variation across hours within each day's peak period, variation across days, or variation across space within an hour. This pricing schedule thus fails to provide efficient incentives to conserve electricity, though it does represent an improvement over a single flat rate.

In reality, however, rates must be set ahead of time. This creates two potential sources of mispricing: errors in matching relative price differentials between different tariff levels and, importantly also, errors from getting the average price wrong. As an example, the peak pricing schedule in the dotted red line in Figure 2 is the best fit schedule based on ex-ante predictions for 2019 using pricing data from the previous three years. Relative peak vs. off-peak prices remain similar to the in-sample prediction in the blue dotted line, but the average retail price no longer matches the average price in the wholesale market overall in 2019. These features, the substantial variation not accounted for by simple TOU rates and the importance of out-of-sample prediction errors in the level of prices, turn out to play a persistent role when analyzing our full data set.

[4] https://www.sce.com/business/rates/cpp

Figure 2: Wholesale Prices and On-Peak vs. Off-Peak Pricing Schedule



Note: Data points in black are drawn from a random sample of 20 PJM network nodes in the 31st week of 2019. The dotted blue line represents the yearly in-sample optimal peak pricing schedule estimated using 2019 pricing data. The dotted red line represents the predicted schedule based on the preceding three years. Peak hours are defined as 6am-8pm. Price outliers below $0 and above $150 are omitted for visual clarity.

3 Data

The pricing data for this paper consist of locational marginal prices for all wholesale electricity markets' hourly auctions in the US: PJM (Mid-Atlantic and some Chicagoland), ISO-New England, New York ISO, ERCOT (Texas), MISO (Midcontinent), SPP (South Central), and CAISO (California).[5] Together, these markets cover approximately two-thirds of the US population. We source these price data for the first six markets from SNL Financial, a subsidiary of S&P Global Market Intelligence that collects the data directly from the Independent System Operators (ISOs), the organizations that coordinate the operations of the electrical grid. The data include both real-time (spot) and day-ahead (forward) prices, with each market's data beginning somewhere between 2000 and 2011 and ending in 2020. We obtain CAISO node-level pricing data from 2009 through 2015 from an archive of downloads directly from the ISO.[6] We

[5]Restructured wholesale electricity markets in the United States do not clear with a single price. Rather, many "nodes" in the market are given their own locational marginal price, which represents the marginal cost of providing electricity at that location in the network, accounting for transmission losses and congestion constraints.

[6]We are grateful to Akshaya Jha for sharing these data.

present some results using the full geographic data, but we often collapse the data across nodes and use an hourly dataset at the ISO level due to a lack of nodal electrical load (quantity) data. Results are shown at the ISO level unless labeled otherwise. We also source load data at the ISO-hour level from SNL Financial, who aggregated these data from each ISO's FERC 714 filing. These data span from 2009 through 2020 for all markets.

Table 1 shows summary statistics. Mean electricity prices are usually in the $20-40 per megawatt hour range, but prices are volatile and extreme pricing happens—either negative or in the thousands of dollars per megawatt hour. This volatility is not completely eliminated by collapsing to the hourly ISO mean.

To translate goodness of fit statistics into dollar deadweight loss (DWL) estimates, we need measures of the slope of supply and demand. For the demand side, we take a linear functional form and calibrate the slope using observed prices and quantities and an elasticity of -0.2 (see Appendix A.1). We perform sensitivity analysis with alternative values of -0.1 and -0.3 in Appendix B. We discuss load-shifting in Section 6.3 and extend the model by applying a range of cross-price elasticities drawn from the literature.

For the supply side, we have collected merit orders—data on the engineering capacities and marginal fuel and other operating costs of power plants in each ISO. We again source these data from SNL Financial. We assign marginal costs for each facility by taking the sum of fuel costs, emissions allowance costs, and other variable operation and maintenance costs. For renewable sources, we first generate baseline capacity for each step by taking the product of the facility's maximum capacity and its capacity factor[7] over the relevant timeframe, and then calibrate hourly generation using observed price and quantity pairs (see Section 4). For other sources, we take the product of the facility's capacity and 1 minus the equivalent forced outage rate demand (EFORd).[8]

[7]The capacity factor is the ratio of a power generator's production over the theoretical maximum possible production under ideal circumstances.

[8]We source this parameter from various technical reports from the ISOs, assigning each facility the appropriate average EFORd by generation type and ISO when possible. When these detailed data are unavailable, we instead assign each facility the ISO-level average EFORd (this is the case for SPP and ISO-NE).

Table 1: Descriptive Statistics for Electricity Prices, All Markets

	PJM	ISO-NE	NYISO	ERCOT	MISO	SPP	CAISO
Node-Level Real-Time Prices ($/MWh)							
Median	29.04	38.60	35.71	21.67	24.16	21.04	30.92
Mean	37.41	47.09	47.09	29.67	29.53	26.40	35.52
Std Dev	36.00	38.93	51.92	104.52	27.51	37.28	43.83
Minimum	-2003.43	-3713.87	-5150.37	-4104.96	-1981.98	-1475.10	-4446.70
Maximum	3941.31	6095.74	5492.07	9139.77	3500	2552.31	3583.17
99^{th} Percentile	169.34	193.97	230.17	164.38	125.83	126.63	199.51
Number of Obs.	209,670,520	142,935,036	84,901,403	48,922,457	156,927,172	642,431,694	304,076,362
Number of Nodes	2124	1133	558	660	1684	7570	8574
Node-Level Day-Ahead Prices ($/MWh)							
Median	31.15	40.05	39.06	24.11	26.01	21.40	34.93
Mean	37.58	47.67	47.25	31.33	30.24	23.24	36.29
Std Dev	27.57	32.19	33.93	76.94	18.02	13.17	15.46
Minimum	-1032.06	-941.23	-349.94	-283.82	-977.11	-1202.67	-1240.47
Maximum	1605.91	3983.00	1629.02	8214.04	2088.92	2885.29	4699.57
99^{th} Percentile	135.59	177.36	172.82	122.95	99.48	60.77	75.60
Number of Obs.	209,670,520	142,935,036	84,901,403	48,922,457	156,923,692	405,844,841	298,087,253
Number of Nodes	2124	1133	558	660	1684	7570	8574
ISO-Level Real-Time Prices ($/MWh)							
Median	31.57	40.37	38.95	22.18	25.26	22.50	31.46
Mean	39.72	48.46	48.20	29.88	31.74	28.59	35.81
Std Dev	30.61	37.89	42.45	95.74	22.36	29.63	36.50
Minimum	-202.03	-155.77	-880.73	-36.51	-198.19	-160.29	-196.76
Maximum	1787.87	2422.89	1623.82	9000.00	1370.52	1423.23	1036.18
99^{th} Percentile	148.06	190.19	191.07	148.65	110.09	110.28	186.52
Number of Obs.	154,946	156,409	184,128	88,440	138,123	122,053	58,720
ISO-Level Day-Ahead Prices ($/MWh)							
Median	33.80	41.99	41.75	24.67	27.22	21.66	35.37
Mean	39.82	49.07	48.38	31.57	32.56	23.62	36.34
Std Dev	25.23	31.39	30.43	73.79	18.37	10.77	12.98
Minimum	0.04	-5.01	0.01	0.87	-30.9	-13.09	-14.14
Maximum	929.89	674.57	976.20	5004.53	316.51	173.41	490.76
99^{th} Percentile	127.30	172.78	152.34	108.33	104.07	56.52	71.35
Number of Obs.	154,946	156,409	184,128	88,440	138,123	59,983	57,208
First Year Observed	2003	2003	2000	2011	2005	2007*	2009
Last Year Observed	2020	2020	2020	2020	2020	2020	2015

Note: The variable of interest is the hourly price of electricity observed in the PJM, ISO-NE, NYISO, ERCOT, MISO, SPP, and CAISO wholesale markets. Data from the first six markets are downloaded from SNL Financial; data from CAISO are downloaded from OASIS. For SPP, real-time price data begin in 2007, but day-ahead data begin in 2014 with the launch of SPP's integrated marketplace. ISO-level data represent mean prices across all network nodes in the market in each hour.

4 Methodology

The primary empirical question in this paper is how well TOU and CPP rates can approximate wholesale energy price fluctuations. We rely on a standard "goodness of fit" metric, the R^2, as well as variants of that metric appropriate to our situation. We define and explain these metrics here.

Our analysis begins by partitioning the year into groups of hours and running regressions that describe how much of the variation across all hours is explained by dummy variables representing those partitions. These regressions take the following form:

$$y_{is} = \sum_k \beta_k D_{is}^k + \varepsilon_{is}, \tag{1}$$

where y_{is} is the wholesale price of electricity in hour i in sample s (where sample refers to a specific year of data in a specific wholesale market), k indexes the set of unique price tiers (there is a unique k for every price charged), D_{is}^k are indicator variables that are coded as one if a given hour falls within the specific tier k, β_k are pricing-policy parameters to be estimated, and ε_{is} is an error term.

After estimating such an equation for a given sample s, we first assume that the utility charges the price $\hat{\beta}_k$ in each pricing tier k. This assigns to each block of hours the average wholesale price in that block. Our focus is not on those prices per se, but rather on how well those prices are able to match the pattern of prices in the wholesale market. Thus, we focus not on the $\hat{\beta}$ coefficients themselves but rather on goodness of fit statistics.

In-sample R^2: We begin with the traditional in-sample R^2, which is defined as one minus the ratio of the sum of squared residuals from an ordinary least squares (OLS) regression divided by the total sum of squares of the dependent variable. That is, $R^2 = 1 - \sum_i (y_i - \hat{y}_i)^2 / \sum_i (y_i - \bar{y})^2$, where y_i is the dependent variable, $\bar{y}$ is that variable's mean, $\hat{y}_i$ is the predicted value, and i indexes the observations in the sample. The R^2 is bounded between zero and 1 when it comes from a regression that includes an intercept term.

In our case, y corresponds to wholesale electricity prices and $\hat{y}$ to retail prices charged to customers. Implicitly, we imagine a utility would set the best possible rates given the allowed degree of flexibility. Jacobsen et al. (2020) show that, under certain conditions, the efficiency-maximizing rate can be found by simply running OLS regressions with indicator variables coded as one for the hours when a given rate tier applies, as in our Equation (1). Intuitively, in any given hour, the deadweight loss (inefficiency) from retail prices that differ from wholesale prices is (approximately) proportional to the square of the difference between the two prices. Thus, inefficiency is minimized by minimizing the sum of squared errors (differences between wholesale and retail prices) across all hours. This is exactly what an OLS regression does. In turn, this means that the R^2 from this regression directly measures the fraction of the efficiency gain that a given pricing scheme can achieve as the proportion of the efficiency gain that real-time pricing could achieve over a baseline flat price.[9]

This approximation is most accurate when demand and supply curves are close to linear, there is little correlation between the slopes of demand curves and pricing errors across hours, when there is little correlation between cross-hour elasticities and products of pricing errors,[10] and when equilibrium effects of policy are small (i.e., the observed price closely approximates the price under real-time pricing). Correlation between cross-hour elasticities and products of pricing errors is zero or small by definition when demand response is entirely or mostly within an hour; i.e., if "load shifting" is small relative to the main demand response. When load-shifting between hours is strong then correlations with products of errors could appear, though Jacobsen et al. (2020) show that the R^2 remains a good approximation in many cases. We investigate cross-hour substitution in detail here using a range of possibilities for load-shifting from the literature in Section 6.3.[11] We find only small biases away from the R^2 measure, and running in either direction (positive or negative correlations) depending on the setting. Therefore for our main analysis below we simplify the model to abstract from cross-effects.

Out-of-sample R^2: The above in-sample analysis assumes that the optimal tariff schedules prescribed by OLS are feasible in practice. In reality, utilities can only set tariffs in advance based on past pricing data and cost projections. It is, therefore, the out-of-sample (OoS) performance that matters. To incorporate an electric utility's ability to set efficient prices into our analysis, we derive several novel out-of-sample welfare metrics.

As an intermediate step, consider an out-of-sample R^2 statistic, which we label OoS R^2, that simply applies the standard R^2 formula to our out-of-sample price schedule. That is, it uses the equation $R^2 = 1 - \sum_i (y_i - \hat{y}_{i,OoS})^2 / \sum_i (y_i - \bar{y})^2$, where $\hat{y}_{i,OoS}$ are fitted values using coefficients estimated from a different sample. For example suppose a utility uses 3 years of data to choose a TOU pricing system for the year 2020: The coefficient estimates would be derived from 2017-2019 data, but OoS R^2 is calculated using 2020 data with $\bar{y}$ still representing the mean of the (realized) y_i. In this case, it is possible that the R^2 becomes negative, which means that the proposed prices created larger squared errors than would result from applying an unbiased flat price (e.g., the in-sample mean for 2020).

The advantage of the OoS R^2 is that it is on the same scale as the in-sample R^2—both

[9] Note that a flat pricing policy has $R^2 = 0$ (regression on a constant), whereas $R^2 = 1$ for real-time pricing (regression on a fully-saturated set of indicator variables for each hour i in each sample s).

[10] Jacobsen et al. (2020) provide conditions on cross-price elasticities where the R^2 approximation holds exactly.

[11] Schittekatte et al. (2024) take a different approach. They focus on the rank correlation of prices with marginal costs within a day, which is consistent with an assumption that all electricity load can freely shift across hours. In that case, a simple R^2 may be a poor proxy. This scenario may someday be true if enough load is fully automated. Our approach is instead to use recent empirical estimates of load-shifting to divide demand responses into load-shifting and total reductions.

statistics measure performance relative to $\bar{y}$. As in the case of the in-sample R^2, the numerator is a sum of squared residuals that is proportional to deadweight loss. Scaling this as a percentage of the gains achieved by an unbiased flat tariff is, however, not obviously meaningful because the unbiased flat tariff is not a feasible policy—it too requires information about the *ex-post* realized mean price. This information is unavailable to the electric utility when prices are set based on historical data or inaccurate forecasts. Hence, out-of-sample, a flat pricing policy is also biased. We therefore propose an alternative measure described next.

Renormalized R^2: Our preferred measure of out-of-sample goodness of fit is what we label renormalized R^2. It is the out-of-sample R^2 of a given policy of interest P minus the out-of-sample R^2 of a flat tariff B given the same information (that is, using the mean price from the estimation sample of past prices), divided by one minus the out-of-sample R^2 of the flat tariff given the same information:

$$\text{Renormalized } R^2 = \frac{R_P^2 - R_B^2}{1 - R_B^2}, \tag{2}$$

where R_P^2 is the OoS R^2 of the policy we want to evaluate and R_B^2 is the OoS R^2 of the equally-informed flat tariff. This renormalized R^2 is equal to the percentage of deadweight loss recovered by applying this policy relative to an equally-informed flat tariff.[12] The benefit of using such a metric is its ease of interpretation—a renormalized R^2 of 1 is equivalent to true real-time pricing (as is the case for the in-sample R^2), and a renormalized R^2 of 0 is welfare-equivalent to the best flat tariff given the information the policymaker has ex-ante.[13]

Equilibrium Effects: The final extension we make to the R^2 method concerns the left-hand side of the regression. In previous analyses, the regressand was the observed wholesale price since it represents the per-unit cost of electricity of the marginal generator. This has a direct welfare implication for a TOU program that affects a small portion of the market, but we are also interested here in the welfare implications of broad adoption. In that case, because the short-run wholesale supply curve is sloped, the marginal cost of electricity varies with equilibrium in the market.

The R^2 method described above would suggest that the first-best counterfactual real-time pricing schedule would be a replication of the wholesale prices we observe in the data (which we call the "observed basis"). However, this policy would change the marginal cost of electricity

[12] To see this, let P, B, U be the OoS policy to be evaluated, the equally-informed flat tariff, and the unbiased flat tariff, respectively. Then, substituting using the result from Jacobsen et al. (2020): $\frac{R_P^2 - R_B^2}{1-R_B^2} = \frac{\left(1-\frac{DWL(P)}{DWL(U)}\right)-\left(1-\frac{DWL(B)}{DWL(U)}\right)}{1-\left(1-\frac{DWL(B)}{DWL(U)}\right)} = \frac{\frac{DWL(B)-DWL(P)}{DWL(U)}}{\frac{DWL(B)}{DWL(U)}} = 1 - \frac{DWL(P)}{DWL(B)}$.

[13] Note that this is on a different scale than the in-sample R^2. Both are linear in policy deadweight loss with a value of 1 representing real-time pricing, but a value of 0 represents the deadweight loss from either unbiased or biased flat pricing for the in-sample and out-of-sample, renormalized R^2, respectively.

in equilibrium; the first-best real-time pricing schedule is therefore a function of an equilibrium outcome. For small shifts in retail prices, like those usually prescribed in peak/off-peak rules, these equilibrium effects will be relatively minor, leaving the observed basis R^2 a good approximation of the true deadweight loss ratio. However, for greater absolute changes in price, like those implied by CPP or complex TOU schemes, the equilibrium effects may be significant.

We address this possibility by simulating equilibrium prices under real-time pricing and under the out-of-sample price schedule in question. In all of our out-of-sample analysis (with the exception of the analysis of locational pricing in Section 6.2), we then evaluate policies using simulated equilibrium prices and refer to this approach as the "equilibrium basis." We generate policies using regressions of simulated equilibrium prices on policy indicators, and then construct the renormalized R^2 metric using the deadweight losses implied by hourly supply and demand curves. This approach also allows us to account for convex generation costs when computing efficiency losses. We discuss the construction of supply and demand curves in detail in Appendix A.1. Although our qualitative conclusions are similar between the observed base and equilibrium approaches, this methodology allows greater precision especially during peak periods.

Operationalization of Policies: Our regression-based methodology allows for the evaluation of any arbitrarily shaped price schedule. Here we discuss the method used for producing TOU and CPP price schedules generally; the specific schedules explored appear in Section 5.

For all out-of-sample results shown in the paper, we assume that price data from the three previous years are used to create the listed year's price schedule. Other papers in this literature have used a similar prediction time window (Schittekatte et al. 2024). The numbers of years chosen reflects a tradeoff between including information about recent price patterns and reducing the influence of noise. In Appendix B, we vary the number of years we use as the training sample and find no significant trend in the resulting R^2. This suggests that there are few predictive gains to increasing the number of years used. In specifications using the observed basis, the regressand is the observed real-time price. In specifications using the equilibrium basis, the regressand is instead the simulated real-time equilibrium price.

When applying TOU pricing schemes we consider varying degrees of complexity using a set of indicators and their interactions. To capture typical within-day variation, for example, we define "peak" as an indicator equal to 1 during hours of high expected prices and 0 otherwise. We define the set of peak hours using an in-sample variant of the best subsets selection algorithm (described in more detail in Appendix A.2). In most markets, this amounts to a day/night indicator, with the exception of ERCOT, where the algorithm chooses peak hours as noon through 6pm.[14] Energy-use patterns also vary between weekends and weekdays, so we introduce

[14]For other markets, we use 6am-8pm (PJM/ISO-NE/MISO/SPP), 7am-8pm (NYISO), and 10am-10pm (CAISO).

a "weekend" indicator variable to represent this. Finally, energy prices vary seasonally as well, so we define "season" as an indicator equal to 1 from April through September and 0 otherwise. In defining some more complex TOU schemes we also use hour of day, day of week, and month of year indicators in a similar fashion. To define CPP policies, we order days by their peak-period price, assign the top-n days a critical-peak price based on either forward or spot prices, and assign all other hours an off-peak price. For further details on CPP schedules, see Appendix A.3.

5 Results

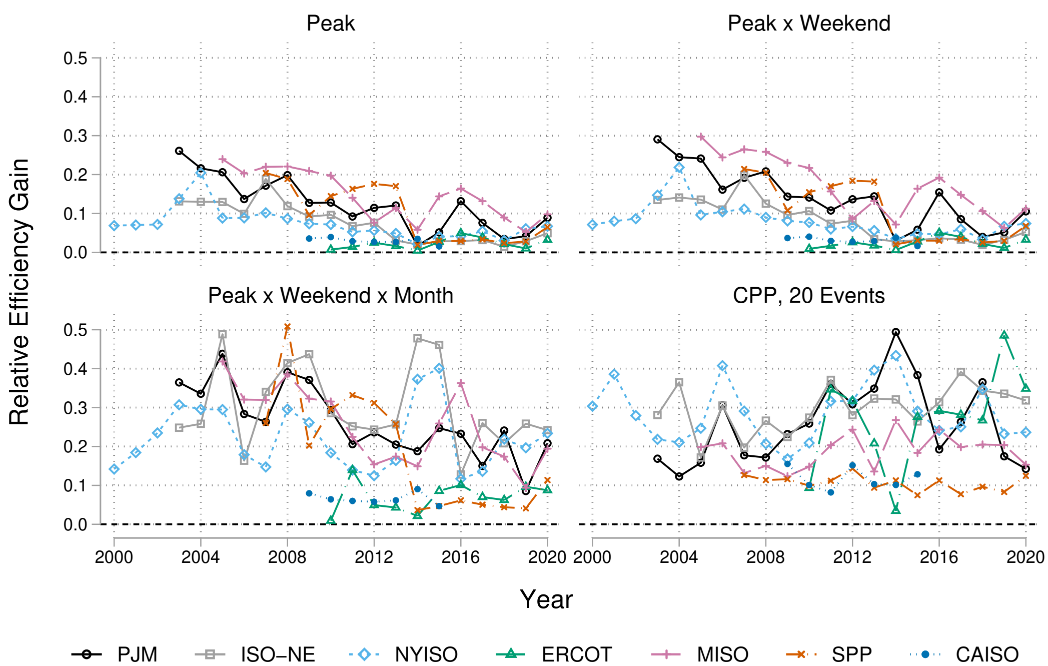
5.1 In-sample Results

We begin with the observed-basis in-sample analysis as a benchmark, showing results across years and markets at the ISO level.[15] That is, for a given pricing policy, we use OLS to define the optimal in-sample price schedule and interpret the R^2 of this regression to be the proportion of deadweight loss recovered relative to the deadweight loss with a flat price. As discussed, these efficiency gains are not actually achievable in practice (since prices are set using historical rather than contemporaneous price data) and so this exercise provides an upper bound on the possible efficiency gains of dynamic pricing policies.

Figure 3 shows the in-sample R^2 for three candidate TOU policies and one candidate CPP policy across all market-years in our sample. The R^2 values vary by year and market. The top-left panel shows that the simplest peak pricing policy explains, on average, 8.9% of the total variation in wholesale prices (median: 7.2%). While small as a percentage, we show below that these gains are still economically significant. Additional complexity increases potential efficiency gains in this in-sample context: gains are weakly increasing when adding variables by construction. These gains turn out to be economically meaningful in our context. Means (medians) of the R^2 in the remaining three panels are 0.099, 0.216, and 0.231 (0.080, 0.208, and 0.224), respectively. Table 2 reports an expanded range of policies for one market, PJM. This shows clearly that increasing complexity, using observed-basis prices and analyzing in-sample fit, strongly improves efficiency. The most detailed pricing policy recovers over 50% of the efficiency benefits of real-time pricing, but such a scheme would require 24 * 7 * 12 = 2,016 different prices (set for each day-of-week by hour-of-day by month combination). This is not a tariff scheme that any utility would realistically consider, but we think it serves as a useful upper bound.

[15] In general, ISO-level R^2s are somewhat lower than node-level R^2s, as we are mechanically removing any variation in price across locations. See Appendix Tables C.1 and C.2 for a comparison of in-sample regressions.

Figure 3: In-Sample Efficiency Gain of Different TOU and CPP Designs



Note: Each point shows the observed-basis in-sample R^2 for a regression of wholesale prices on indicator variables representing the given policy in a market-year. "CPP" constitutes a set of 20 indicators equal to 1 during each of the 20 highest price peak periods throughout the year and 0 otherwise. Note that temporal coverage varies across markets due to data availability.

Table 2: R^2 from In-Sample Electricity Tariff Regressions - PJM

Pricing Scheme	2003	2004	2005	2006	2007	2008	2009	2010	2011
Flat Tariff	0	0	0	0	0	0	0	0	0
Peak	.261	.216	.206	.136	.171	.199	.127	.128	.092
Peak x Weekend	.291	.245	.240	.160	.192	.208	.143	.141	.108
Peak x Season	.277	.225	.216	.150	.185	.282	.207	.153	.125
Peak x Weekend x Season	.308	.257	.250	.177	.207	.291	.224	.168	.143
Peak x Weekend x Month	.365	.335	.437	.282	.263	.391	.370	.293	.206
Hour x DoW x Month	.601	.601	.665	.481	.490	.611	.558	.493	.411
CPP, 20 Events	.168	.123	.158	.305	.177	.172	.231	.259	.360
Pricing Scheme	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Flat Tariff	0	0	0	0	0	0	0	0	0
Peak	.114	.120	.017	.050	.131	.075	.034	.041	.089
Peak x Weekend	.136	.143	.029	.057	.153	.085	.039	.052	.105
Peak x Season	.125	.137	.050	.070	.150	.089	.054	.048	.100
Peak x Weekend x Season	.149	.160	.067	.078	.172	.101	.061	.059	.115
Peak x Weekend x Month	.236	.204	.187	.246	.232	.150	.240	.086	.208
Hour x DoW x Month	.437	.398	.349	.444	.500	.356	.361	.312	.453
CPP, 20 Events	.308	.349	.493	.382	.193	.263	.365	.174	.143

Note: Cells of this table present R^2 values for regressions of the hourly ISO-average price of electricity observed in the PJM wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use or critical-peak price schedule (given by the row). Peak hours are defined as 6am-8pm. Season is an indicator variable splitting the year into April through September and October through March. Versions for other markets are found in Appendix Table C.1. Versions at the node level are found in Appendix Table C.2.

We also find significant cross-market heterogeneity—prices in MISO and PJM are generally better matched by simple intermediate TOU pricing schemes than those in ERCOT or CAISO. This suggests that the performance of pricing policies is highly market-specific; some markets experience substantially higher efficiency gains from time-varying pricing than others.

The in-sample R^2 for the CPP policy (represented by a regression with indicator variables for each of the 20 highest peak price periods) is similar in both mean and variance to the peak $\times$ weekend $\times$ month policy. The R^2 of the CPP policy has an interquartile range between 0.148 and 0.306, and the R^2 of the complex TOU has an interquartile range between 0.117 and 0.295. This suggests that, in-sample, the potential efficiency gains of CPP policies are similar to those from a rather complex TOU policy.

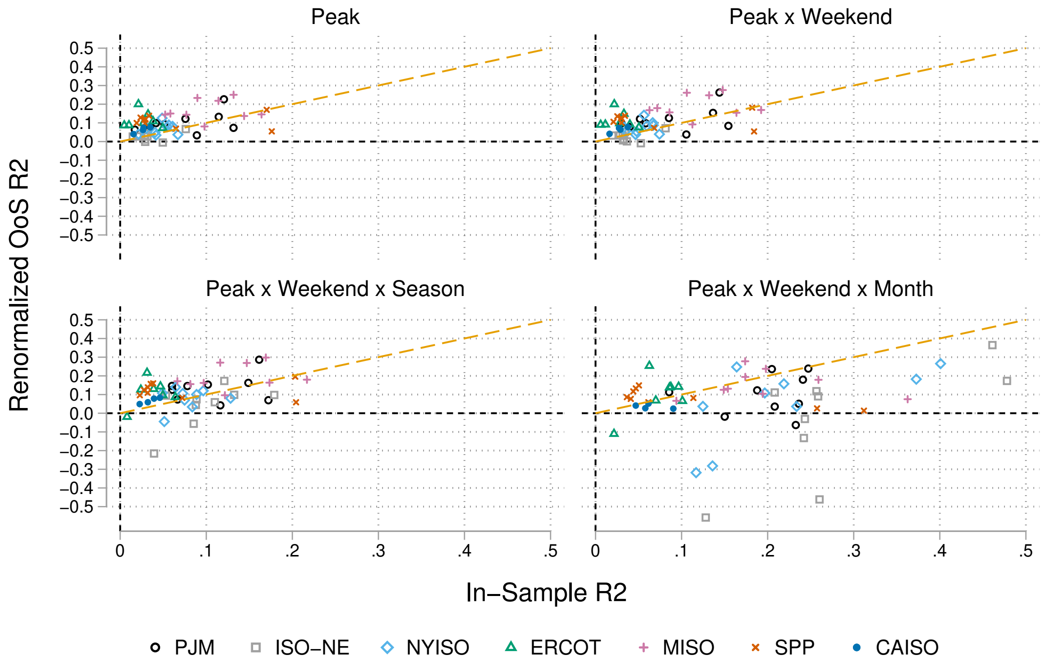
5.2 Out-of-Sample TOU Policies

We now turn to the efficiency gains possible with more realistic, out-of-sample policies. Are historical prices informative enough to produce efficiency-improving price schedules, and do they offer economically meaningful improvements? We use the equilibrium-basis renormalized R^2 metric described in Section 4 to consider policies set using out-of-sample data. The renormalized R^2 represents the proportion of deadweight loss that is recovered by the policy relative to an equally-informed (i.e., set using the same historical dataset) flat tariff. This concisely summarizes the aggregate efficiency gains available. Our simulation method also allows us to explore further decomposition below. The results in this section differ from the prior results both in that they use out-of-sample data and in using simulated real-time prices as the dependent variable.

In general, R^2 values are low for out-of-sample TOU policies. These low R^2 values persist for complex policies. For peak pricing, the renormalized R^2 has a mean (median) of 0.095 (0.084). The renormalized R^2 from the highly-complex peak $\times$ weekend $\times$ month policy has a mean (median) of 0.071 (0.097). In contrast to the in-sample analysis, our out-of-sample evaluation shows no gain—and even a loss after a certain point—from additional complexity.

Figure 4 compares results from our observed-basis in-sample analysis to results from our equilibrium-basis out-of-sample analysis (see Appendix Figure B.1 for plots of the out-of-sample efficiency gains analogous to Figure 3). While the two methodologies are not directly comparable in welfare terms, this provides an illustration of how seemingly promising in-sample policies can underperform out-of-sample. In each panel, we display the 45-degree line in yellow. Points above this line, under our framework, appear to perform better out-of-sample than they do in-sample. Points below this line appear to perform worse. Comparing panels, we see that more complex policies typically underperform expectations more strongly than simple policies do. Further, within each panel, when a greater proportion of the in-sample variation is explained by the policy, the greater the risk of overfitting. This can be seen using the line of best fit for

Figure 4: In-Sample versus Out-of-Sample Performance



Note: Each point represents the in-sample and out-of-sample performance for a given policy in a given market-year. The horizontal axis represents the simple in-sample R^2 as in Section 5.1. The vertical axis represents the equilibrium-basis renormalized R^2 as described in Section 4 and as used in most of the out-of-sample analysis in the paper. 45-degree line shown in yellow. In-sample results are only shown when there are enough data to produce corresponding out-of-sample values.

each panel—for all four policies shown, this line has a slope below 1.[16] This is consistent with in-sample gains not being realized out-of-sample, even conditional on the tariff structure being used.

This highlights a critical tradeoff between complexity and prediction accuracy. For a peak pricing policy, the utility only needs to correctly gauge the price differential between two groups of hours. For the peak × weekend × month policy, they must accurately set a 48-part tariff. As policies grow more complex, price schedules are determined using fewer observations per price category in the training data, and the category-specific mean prices in the test data are also more volatile for the same reason.

The renormalized R^2 results are robust to our choice of the price elasticity of demand. Appendix Figure B.2 shows that choosing -0.1 or -0.3 as opposed to our base value of -0.2 has very little impact. Likewise, the result that there are diminishing (and even negative) returns to TOU complexity is similar when using observed wholesale prices instead of simulated real-time

[16]The (slope, intercept) of these fit lines in each panel are (0.589, 0.061), (0.639, 0.062), (0.531, 0.065), and (0.285, 0.021), respectively.

prices (see Appendix Figure B.3).[17] Our main conclusions are also robust to variation in the number of years used for training the data to set prices—see Appendix Figure B.5 for details.

We conclude that, for the markets and years in our study, simple peak pricing delivers most of the potential efficiency gains from TOU pricing. Added complexity, where it does render additional benefits, leads to a tradeoff between additional average welfare gains and an associated increase in the variance in those gains, and also to an increasing risk of performing worse than flat pricing.

An advantage of our simulation approach is that we can adjust our policy price schedules to evaluate why promising in-sample policies fail to deliver out-of-sample. We focus on two sources of misprediction: getting the average price wrong and getting the "shape" of the optimal price schedule wrong. The average price can be wrong because the average wholesale electricity price varies from year-to-year—therefore the mean price in-sample need not equal the mean price over the last three years.[18] To decompose the impact of this mispricing, we define a "mean-corrected" price schedule that vertically shifts the policy price schedules to match the mean in the year where the policy is applied.

The error remaining in this mean-corrected schedule comes from the difference in prices between sets of hours differentiated by the policy; i.e., getting the shape wrong. For example, absolute levels notwithstanding, if the optimal policy would charge \$20 per megawatt hour more in peak hours than off-peak hours, an out-of-sample policy charging only \$15 per megawatt hour more in those hours constitutes an error. By comparing the deadweight loss of out-of-sample, mean-corrected price schedules with in-sample price schedules, we are able to quantify the efficiency losses attributed to these "shape" errors versus the efficiency losses from inaccuracy in the mean price level.

Table 3 shows the simulated deadweight loss associated with two pricing policies in each of our markets, averaged over the years of our sample.[19] As reflected in the R^2 values, the difference between the "biased flat" and the two out-of-sample columns shows that there is a modest efficiency gain from implementing these out-of-sample TOU policies. The nationwide

[17] The mean R^2 using the observed wholesale prices is lower, however, as price spikes in observed prices are more extreme than under equilibrium pricing. Consequentially, the observed basis overweights the deadweight loss associated with the highest price hours—see Appendix Figure B.4 for details. This demonstrates why the equilibrium simulations are necessary to fully understand the welfare impacts of intermediate policies.

[18] Note that the "mean prices" in question are the means of the simulated real-time prices, not those of the observed wholesale prices.

[19] Note that the in-sample DWL values in this table are derived from our simulated equilibrium prices, in contrast with the R^2 values used in Section 5.1, which use observed wholesale prices. The relative pattern of results is robust to whether we use the observed or equilibrium-basis prices, but the total dollar numbers are lower when using the equilibrium-basis prices. Thus, we conclude again that the qualitative results about complexity and predictions out of sample are not driven by our method of establishing counterfactual real-time prices, but the total efficiency gain in dollars is more sensitive. As explained in Section 4, simulated equilibrium prices are superior to observed prices as the latter fail to account for equilibrium adjustments caused by the pricing schemes.

Table 3: Out-of-Sample Deadweight Loss by Market

ISO	Biased Flat	Peak: Out-of-Sample	Peak: Mean-Corrected	Peak: In-Sample	Peak x Weekend x Month: Out-of-Sample	Peak x Weekend x Month: Mean-Corrected	Peak x Weekend x Month: In-Sample
CAISO	207	193	118	118	200	122	107
ERCOT	663	594	521	521	581	505	431
ISO-NE	243	237	181	181	239	187	103
MISO	237	199	136	135	201	138	106
NYISO	210	197	145	145	198	146	91
PJM	505	462	306	304	453	299	193
SPP	131	117	91	91	119	94	80
All ISOs	2196	1998	1497	1493	1991	1491	1111

Note: This table presents estimated DWL figures (in millions of dollars per year) for two potential policies across markets (averaged over the period 2014-2020) with an assumed demand elasticity of -0.2. The biased flat column represents DWL from an out-of-sample flat tariff equal to the previous three years' mean. The out-of-sample columns represent the DWL from the tariffs generated with data from the three years preceding the listed year. The mean-corrected columns are the same, but tariffs are shifted to align the mean of the tariff with the true real-time mean price in the listed year. The in-sample columns represent the DWL from the tariffs generated with data from the listed year.

efficiency gains relative to the biased flat policy are ∼\$198 million per year for the peak policy (the difference between the DWL of \$2,196 million for the biased flat policy and the DWL of \$1,998 million for the out-of-sample peak policy) and ∼\$206 million for the peak × weekend × month policy (the difference between the biased flat DWL of \$2,196 million and the peak × weekend × month DWL of \$1,991 million)[20], with significant heterogeneity across markets.

The hypothetical in-sample peak policy reduces DWL by ∼\$505 million relative to the OoS peak policy, nearly all of which is due to *getting the average price right*, not better fitting the relative price movements. For the more granular policy, the hypothetical in-sample policy would reduce DWL by ∼\$880 million relative to OoS, of which 57% is due to getting the average price right and the remaining 43% from fitting the shape of price movements more precisely. These results highlight the inherent weakness of TOU pricing that is often discussed in the literature—tariffs must be set using historical prices, which ex-post will be a poor fit for realized relative price changes. Our analysis highlights a different reason why TOU pricing underperforms relative to real-time or critical-peak pricing: despite the focus on price fluctuations as motivation for CPP and real-time pricing, getting the average price level right appears to be the more important benefit of real-time pricing.

[20] These numbers do not line up arithmetically because of rounding.

In summary, TOU policies are limited in the efficiency gains that they can achieve, but a simple peak vs. off-peak tariff can still recover about 10% of the efficiency gap with real-time pricing—varying from 0-25% across markets and years—and is straightforward to implement. Additional gains to implementing more complex policies are small because predicting the shape of the optimal tariff becomes more difficult as the number of price levels increases.

5.3 Out-of-Sample CPP Policies

We now use the same methodology, the equilibrium-basis renormalized R^2, to evaluate potential CPP policies and their performance out-of-sample. We focus on how the design of CPP policy impacts its efficiency. How are events called? How are prices set? How many events are allowed? While our focus with TOU policies was on their ability to project optimal price schedules a year in advance, a CPP policy that uses day-ahead or real-time price information to call critical-peak events and set critical-peak prices is feasible and may able to match wholesale prices better and generate larger efficiency benefits per hour affected. This leads us to also explore variation in what information is used to set CPP rates that is different from our analysis of TOUs.

In this section, we assume that CPP rates are introduced on top of an otherwise flat rate. This allows us to isolate their efficiency potential. In the next section we demonstrate the interaction of CPP and TOU rates.

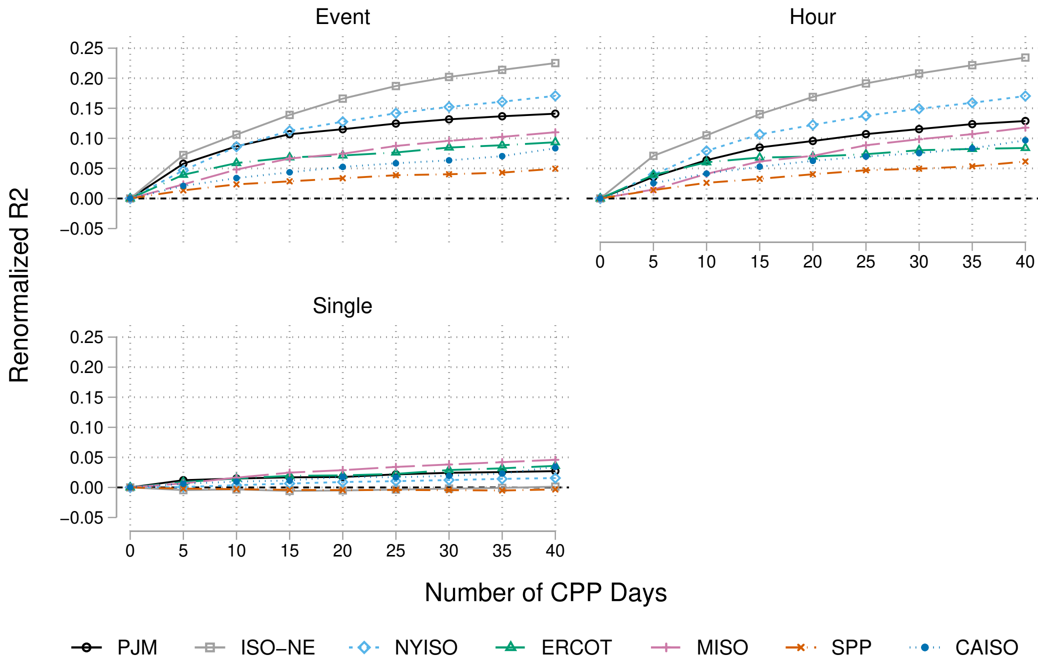
We define a critical-peak event using the same hours of the day as TOU peak/off-peak pricing, but only a limited number of days of the year will be subject to a critical-peak price. By contrast, all off-peak hours and all peak hours on non-event days receive the same flat rate. Critical-peak events can be called a day in advance using price information from the day-ahead market.

Utilities have a number of options regarding the structure of prices charged during these critical peak events. Our default throughout this paper is event-level pricing, which assigns a unique price for all hours during each critical peak event called. In this case, we assume that utilities will charge a price based on the average day-ahead price for each peak period.[21] As an alternative we consider hour-level pricing that assigns a unique price to each hour of each critical peak event—this is between 7 and 16 times as granular as event-level pricing, depending on the number of hours in the peak period. In the opposite extreme, "single" event pricing assigns just one uniform price to every hour of every critical peak event across the year—overwhelmingly, this is the pricing structure that utilities use when implementing CPP policies.[22] In principle,

[21] We use predicted values from a regression of mean simulated equilibrium prices on mean day-ahead prices during peak periods to set event-level prices. See Appendix A.3 for details.

[22] We set this price to the mean of all event-level prices across the year—equivalently, this is the mean of the simulated equilibrium prices as predicted by the observed day-ahead prices among all critical peak hours. Note that this is too optimistic as, in practice, the single price would be set a year in advance, but we show below that even this optimistic version hardly achieves any efficiency gains.

Figure 5: Out-of-Sample Efficiency Gain of Different CPP Designs



Note: Each point represents the mean equilibrium-basis renormalized R^2 for a given CPP policy in a given market, using the simulation methodology described in Section 4. Means are taken across years between 2014-2020, although CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues, and SPP begins in 2015 as their day-ahead market began in March 2014. Prices are based on and called on day-ahead prices.

utilities could call critical peak events in real time and charge a real-time price, but this is far removed from standard practice. Typically, utilities call events around one day in advance, which increases the margin of error for calling events and therefore decreases the efficiency gains.[23]

Figure 5 shows the renormalized R^2 of CPP policies varying across the number of events called, with each panel representing a different critical-peak pricing structure. Most strikingly, our results show that using a single price for all peak events prevents most meaningful efficiency gains—even with 40 events, no market averages even 5% deadweight loss recovery (and SPP is even negative on net). In effect, this policy becomes a TOU peak policy that only affects a small number of days per year. In principle, this could be more effective than the TOU peak policy, as it would allow the policymaker to more precisely target high-price events. However,

[23] An additional question relating to CPP policy design is how accurately utilities can predict the peak periods with the highest prices throughout the year. Most active CPP policies have a fixed maximum allotment of CPP days per year (usually between 10 and 20). We assume that the utility calls event days with perfect hindsight. In reality, the utility must solve an optimal stopping problem to determine when they should call these critical peak events. We abstract from this and instead compare the performance of CPP policies that call and price with either day-ahead or real-time prices. Note that calling and pricing based on day-ahead prices serve as the default in all other figures in this text.

we do not find this to be the case in our data. In contrast, a CPP with 20 event-level prices can consistently recover a significant amount of deadweight loss, with a mean (median) R^2 of 0.092 (0.074). We see little to no change in efficiency from further increasing the granularity to event pricing at the hour-level.

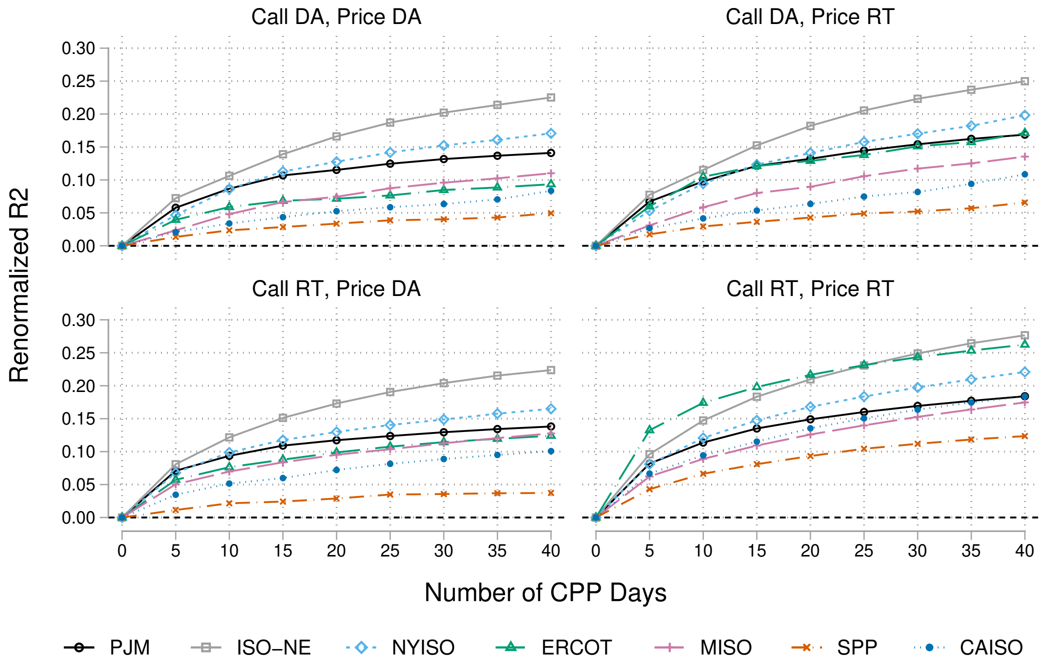
In general, these results suggest that the gains from CPP policies that are priced and called based on day-ahead markets are similar to those from the simple TOU policies in Section 5.2. The mean deadweight loss across all market-years for the biased flat policy is $321 million. With 20 CPP events, it is $273 million, and with 40 events, it is $259 million. The average for the best-performing TOU policy (peak $\times$ weekend $\times$ season) is $285 million. The overall economic magnitude remains somewhat limited, although closing the efficiency gap relative to real-time pricing by approximately 10% is still a meaningful improvement. As with TOU pricing, the efficiency gains from CPP policies vary significantly by market, with 20 events in ISO New England showing large gains close to 17% but SPP achieving just above 3%. This can be explained in part by differences in predictive accuracy in the day-ahead market. Regressions of real-time prices on day-ahead prices yield R^2 values well below 1. This demonstrates that a sizable component of the variation in real-time prices cannot be predicted even a day in advance, though forward markets in some ISOs explain more of the variation in spot prices than in others (see Appendix Figure B.6 for details).

We now consider the efficiency gains from allowing utilities to set appropriate prices shortly before these events occur. We also vary whether the critical peak event prices are defined based on the day-ahead price data or on the real-time price data. We consider the case of event-level pricing unique to each event window.

Figure 6 presents the results of CPP policies that are called and priced based on day-ahead vs. real-time prices. The top-left panel presents the same results as the top-left panel of Figure 5—this is the "realistic" scenario that allows the utility to call and price events based on day-ahead market information. Broadly, we find small benefits to increasing either price accuracy or event-calling accuracy, but larger benefits from increasing both. As before, with 20 CPP events, calling and pricing on the day-ahead market results in a mean (median) renormalized R^2 of 0.092 (0.074). Calling events based on real-time prices leads to slightly higher R^2 at 0.102 (0.099), and pricing events real-time also leads to somewhat higher R^2 at 0.111 (0.0129). When both calling and pricing in real time, however, R^2 increases much more to 0.157 (0.149). This suggests that, while an ex-post real-time CPP policy closes an additional $\sim 7\%$ of the efficiency gap with real-time pricing, a more feasible implementation using day-ahead calling and pricing can still meaningfully improve efficiency.

We are not aware of any prior studies that systematically demonstrate how the efficiency gains from CPP relate to these underlying design features. Of special note is our finding that pre-specifying a single event price for the year ahead of time sharply limits the value of CPP

Figure 6: CPP Price and Call Comparison



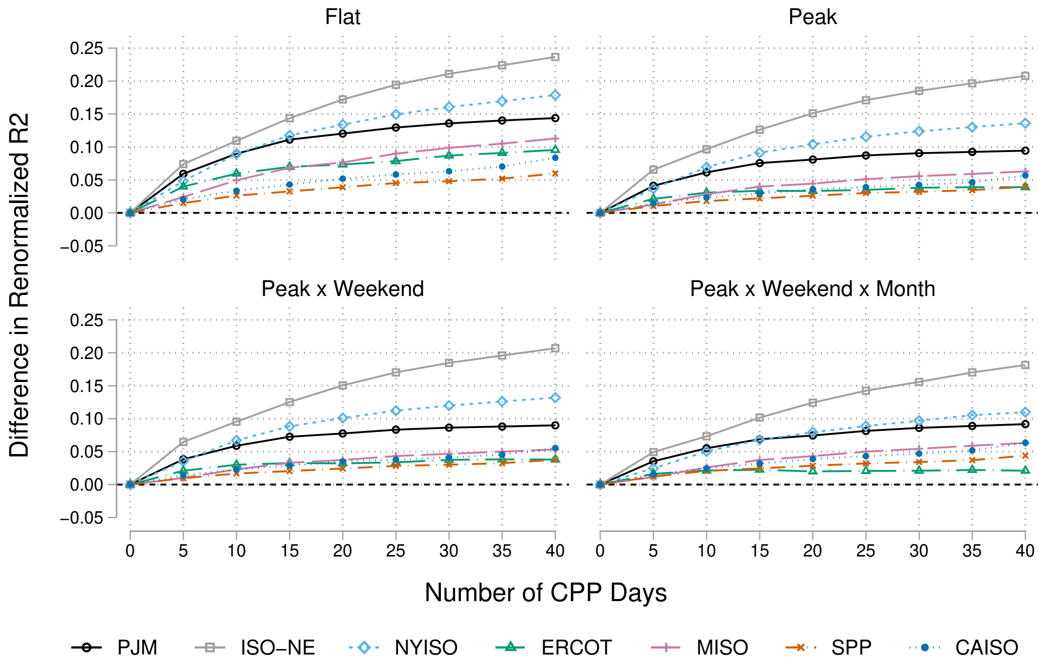
Note: Each point represents the mean equilibrium-basis renormalized R^2 for a given CPP policy in a given market, using the simulation methodology described in Section 4. The two panels on the left price based on day-ahead prices and the two panels on the right price based on real-time prices. The two panels on top call on day-ahead prices, and the two panels on the bottom call on real-time prices. Means are taken across years between 2014-2020, although CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues, and SPP begins in 2015 as their day-ahead market began in March 2014. Each CPP event is assigned a unique critical-peak price.

plans. This stems from the fact that there are large differences in equilibrium prices among the peak days within each market-year in the data and from our assumption that there will be meaningful differences in demand response across "peak" and "super peak" days. This may run counter to the intuition of some because prior studies have suggested that many customers have a binary response to pricing events (Wolak 2011a; Gillan 2017). But, as long as there is still slope in the *aggregate* demand response, getting the price differences right among the very highest price days will still have efficiency gains.

5.4 Are TOU and CPP Policies Complementary?

In the previous two sections, we discussed whether or not TOU or CPP price policies are effective at improving efficiency out-of-sample, and we found mean renormalized OoS R^2's of approximately 10% each. TOU schedules attempt to match the *typical* fluctuations in electricity prices, and CPP events attempt to match *spikes* in electricity prices. As they serve different

Figure 7: CPP with Underlying TOU Policies



Note: Each point represents the mean differenced equilibrium-basis renormalized R^2 for a given CPP policy in a given market, using the simulation methodology described in Section 4. Means are taken across years between 2012-2020, although CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues, and SPP begins in 2015 as their day-ahead market began in March 2014. Prices are based on and called on day-ahead prices. Each CPP event is assigned a unique critical-peak price.

purposes, this raises the possibility that they are largely complementary and could lead to a larger efficiency gain when applied simultaneously.

Concretely, what we mean by complementary policies is that their DWL gains (and thus their R^2s) are additive when layering the two policies on top of one another. To empirically investigate this, we define a set of CPP policies jointly with various TOU policies. To isolate the efficiency gains of the additional CPP events, in addition to renormalizing to a flat baseline, in each panel we subtract the renormalized R^2 values of the respective TOU policy with zero critical-peak price events.[24] Algebraically, this is the difference in DWLs between the TOU-only and the joint policy divided by the DWL of the flat policy—if the policies are perfectly orthogonal, this metric will have the same value regardless of the underlying TOU policy.

Figure 7 presents the results, which align with the hypothesis that TOU and CPP are largely complementary. The top-left panel presents the same results as the top-left panel of Figures 5 and 6. The next three panels present the same, but with increasingly complex underlying TOU

[24]This mechanically sets the zero-event CPP policy to a renormalized R^2 of zero, as in the previous CPP figures.

schedules. R^2 values decrease slightly as underlying policies become more complex. For a CPP policy with 20 events, the mean (median) value for the policies described by the four panels is 0.095, 0.068, 0.065, and 0.058 (0.077, 0.044, 0.037, and 0.043), respectively. This suggests that, on average, less than half of the efficiency gains of CPP would be subsumed by a pre-existing TOU policy. This is more present in some markets than others—ERCOT and MISO incur more significant cannibalization, while efficiency gains in ISO-NE are closer to perfectly orthogonal.

These results suggest that TOU and CPP policies provide separate and largely complementary efficiency gains in retail markets. The approximately 10% DWL recovery from TOU and the approximately 10% DWL recovery from CPP, accounting for a mild overlap, result in a mean 17.2% DWL recovery for a policy with 20 CPP events with a peak $\times$ weekend policy underlying it—with 40 CPP events, this increases to 19.6%. This is a meaningful efficiency gain that seems quite feasible to achieve, so long as utilities are disciplined in the number of TOU rates they choose to set and judiciously use recent price data to create event-level critical-peak prices. Even so, it does not come close to the efficiency gains from real-time pricing.

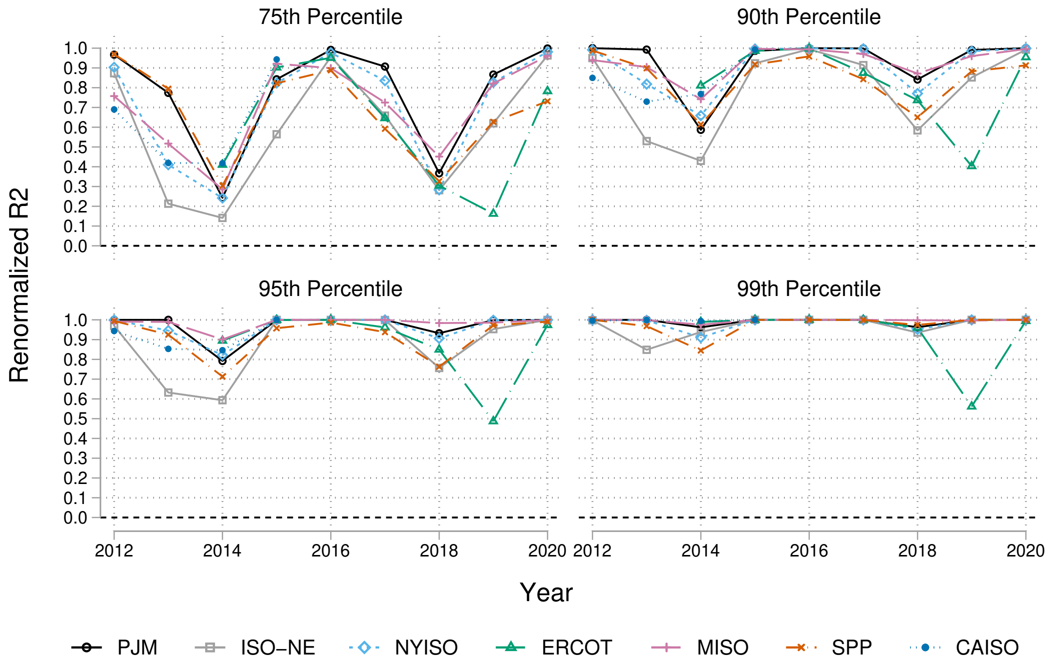
5.5 Real-Time Pricing with Price Caps

So far, we have focused on TOU and CPP because these are the time-varying pricing policies that have been implemented in practice. They are also "manageable" from the perspective of the effort needed from a customer to respond to price signals. In contrast, real-time pricing has drawn skepticism as it might be challenging for customers to effectively respond without automation. Further, it could lead to significant bill uncertainty and leaves people vulnerable to potentially skyrocketing prices during extreme peak hours. Motivated by the potential efficiency benefits of real-time pricing, and the need to address its price-volatility concerns, we now study the efficiency gains of a third set of pricing systems: real-time pricing programs with price caps. The caps prevent retail prices from going to extreme levels. Naturally, this limits the efficiency of real-time pricing during hours when the price ceiling is exceeded, but the system achieves maximum efficiency in the remaining hours.

Our methodology is well-suited to compare real-time pricing programs with price ceilings to those without. We take the schedule of simulated first-best equilibrium prices under real-time pricing, and truncate its values above a certain threshold. We then calculate deadweight loss under both this capped real-time pricing schedule and an out-of-sample flat tariff, and calculate the renormalized R^2 value as usual. This value gives the proportion of deadweight loss that would be abated by the capped real-time pricing program over the biased flat tariff. Equivalently, one minus this metric is proportional to the welfare *losses* associated with having imposed the cap onto a true real-time pricing program, scaled relative to the total deadweight loss associated with a biased flat policy.

Figure 8 presents the results with price caps at the $75^{\text{th}}, 90^{\text{th}}, 95^{\text{th}}$, and 99^{th} percentiles of

Figure 8: Real-Time Pricing with Price Caps



Note: Each point represents the equilibrium-basis renormalized R^2 for real-time pricing with a given price cap in a given market-year, using the simulation methodology described in Section 4. CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues. Panels represent real-time pricing schedules with price caps set at the 75$^{\text{th}}$, 90$^{\text{th}}$, 95$^{\text{th}}$, and 99$^{\text{th}}$ percentiles of simulated equilibrium prices from the three previous years, respectively. A version using in-sample price caps appears in Appendix Figure B.7. A version using a constant price cap over time appears in Appendix Figure B.8.

simulated equilibrium prices over the preceding three years of data.[25] The results reveal large efficiency benefits from real-time pricing with price caps. Across these four panels, the mean (median) renormalized R^2 values are 0.670, 0.874, 0.932, and 0.978 (0.766, 0.932, 0.988, and 1), respectively. Price caps set at the 99$^{\text{th}}$ percentile almost fully restore efficiency, and they do so with price limits that are arguably implementable—the caps average between \$52 in ERCOT and MISO and \$96 in ISO-NE (see Appendix Table B.1 for details). These caps are an order of magnitude smaller than the typical wholesale-market price caps, which are in the thousands of dollars.

On average across markets and years, we conclude that welfare gains from price-capped real-time pricing programs persist, even when price caps bind on a significant fraction of hours. These programs correct for all mispricing in most hours of the year, precisely fitting both the mean and the shape of marginal costs.

Results are highly variable from market to market and (more strongly) from year to year.

[25]We avoid using absolute dollar number caps because the distributions of observed and simulated prices vary significantly between market-years. See Appendix Table B.1 for a sense of scale in dollar terms.

In large part, this is due to the fact that we are setting these caps out of sample. In years where the prices of inputs to marginal electricity generation (oil, natural gas, coal, etc.) increase, the cap (set based on historical data) binds more frequently, leading to a lower capped-policy R^2. In years where the average price of electricity falls nationwide, the cap binds less frequently, leading to higher R^2s. To net out these effects, we include a (not achievable in practice) version in Appendix Figure B.7, which fixes the proportion of hours in which the cap binds by setting it in-sample. Results are less variable across market-years, but are qualitatively similar on average.

These results suggest that the current emphasis on implementing TOU and CPP programs is leaving a pricing policy with potentially superior efficiency performance on the table—even very conservative price caps achieve far higher efficiency gains than either TOU or CPP systems. Importantly, the efficiency gains from real-time pricing with caps that we calculate assume that demand can respond to a high-dimensional, frequently varying price schedule. Many consumers may be overwhelmed by the price variation and might (perhaps rationally) ignore much of it. In contrast, programmable, automated load could handle such a price schedule well. Thus, we interpret the potential gains from real-time pricing with caps as something that could be realized in the future as more load becomes "smart," rather than gains that could be realized immediately.

6 Extensions

6.1 Best Subsets Selection

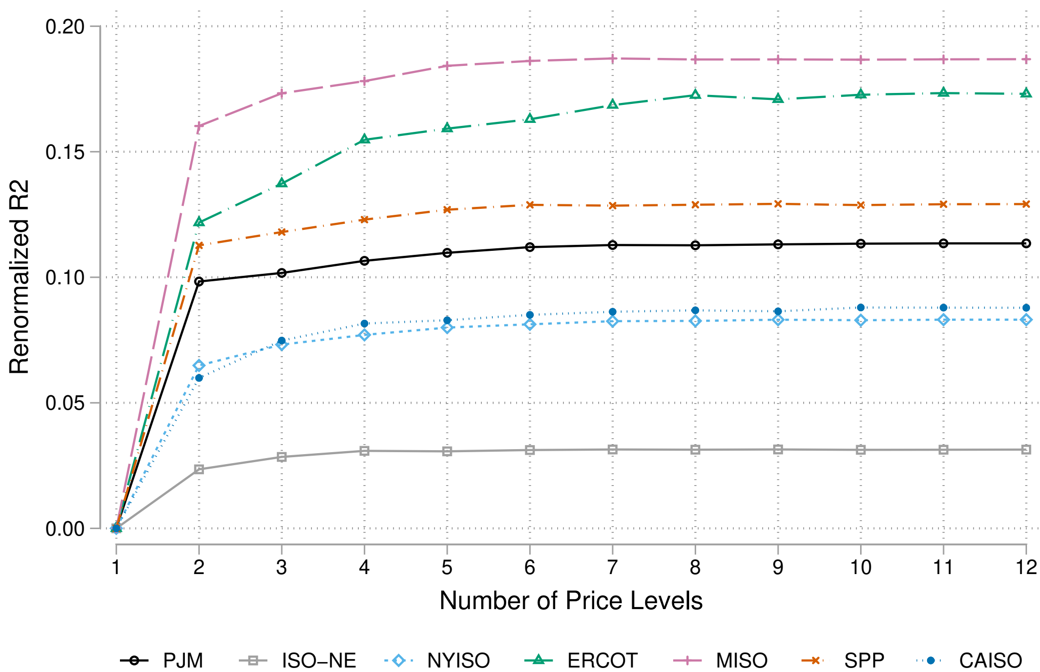
While our analysis has evaluated many potential policies, there are obviously many more potential variations and combinations than we are able to show explicitly. Out of this vast expanse, it becomes difficult to determine which policies are globally optimal. In particular, we may ask how many different price levels we should define, what sources of cyclical variation these should attempt to capture (e.g., daily, weekly, seasonal), and what the specific temporal "cuts" should be for a given policy. As a case study, we apply a variant of the best subsets selection algorithm (BSS) to create a more flexible set of TOU rates. Specifically, we look at systems with as many as 12 different price levels throughout the day and allow the data to choose any contiguous sets of hours to be put into each of the rate categories (with that price pattern then applied to each day throughout the year). In contrast, for our peak/off-peak analysis above, we took as given the peak hour window for each ISO.[26]

The BSS algorithm is an optimization procedure that, in a regression context, selects the best k linear predictors out of a large set of possible predictors. Applied to our context, it begins

[26]Note that the peak hours we took as given were selected using an average over runs of an in-sample version of this same algorithm, as described in Appendix A.2.

by fixing the number of price levels k within the day. Then, we evaluate each of the possible sets of k contiguous price levels within the training data (using the same three-years-prior training data as before). This step is computationally expensive, so we lower the temporal granularity of our policy space to two-hour intervals—this leaves us with between 1 and 12 different price levels. To construct each of the $\binom{12}{k}$ tariff schedules, we define a set of indicator variables $I\{d_{j-1} < \text{hour}_i \leq d_j\}$, where $\{d_j\}_{j=1}^{k}$ is a subset of $\{\text{midnight}, 2\text{AM}, 4\text{AM}, \ldots, 10\text{PM}\}$ with size k. This splits the day into k contiguous periods, potentially with one period maintaining continuity by crossing between days (e.g., a period of 10pm to 2am). For example, when $k = 2$, the set of potential policies are all pairs of $\{\text{midnight}, 2\text{AM}, 4\text{AM}, \ldots, 10\text{PM}\}$ — the hours between the policy pair are classed as "peak" and the remaining hours are "off-peak" (or vice versa). For each policy, the model records the training-data root mean squared error (RMSE) from an OLS regression of simulated equilibrium prices on the set of indicator variables dictated by the policy. Finally, it then applies the policy schedule with the lowest training RMSE to the test data, where we calculate deadweight loss and obtain the renormalized R^2 metric used throughout the rest of the paper. We do this for k between 1 and 12 for every market-year in the data.

Figure 9: Best-Subsets Selection for Within-Day Price Levels



Note: Each point represents the mean equilibrium-basis renormalized R^2 for a given number of price levels in a given market, using the simulation methodology described in Section 4. Means are taken across years between 2009-2020, although CAISO ends in 2015 and ERCOT begins in 2011 due to data availability. Policy structure for each number of price levels in each market-year is selected using a best subsets selection procedure, described further in the text.

Figure 9 shows the results of this exercise, presenting mean renormalized R^2 values for each market by number of price level pairs. By construction, a policy with 1 price group (i.e., a flat price) has a renormalized R^2 of 0. Splitting the day into two pricing periods (i.e., a simple peak vs. off-peak tariff) provides significant efficiency gains in most markets. Optimal peak pricing policies according to the BSS algorithm typically set either an off-peak period throughout the night or an on-peak period in the early evening. However, we find very little evidence for additional efficiency gains beyond this first cut—the efficiency gains flatten out almost immediately, possibly with the exception of ERCOT. We interpret this as another piece of evidence bolstering our theme of relative simplicity: most of the feasible efficiency gains from TOU come from relatively simple schemes.

6.2 Locational Pricing

There is a substantial literature discussing the value of locational marginal pricing in wholesale electricity markets (e.g., Hogan 2002; Wolak 2011b; Triolo and Wolak 2022). However, the link between locational marginal pricing and intermediate dynamic pricing schemes has not been explored as thoroughly. Are intertemporal patterns in wholesale electricity prices predictably different across space?

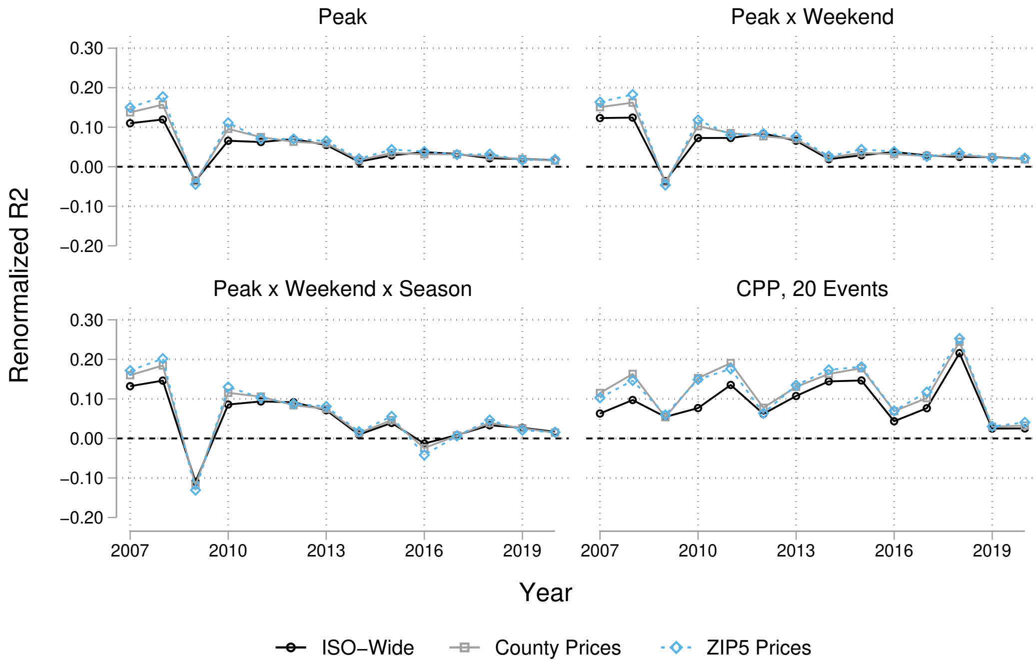
To answer this question, we exploit the spatial information provided by our price data. Recall that, in other sections of this paper, we have used a single hourly average price for each market we study. We now use node-level prices at the same temporal frequency. We estimate the relative out-of-sample efficiency of various TOU and CPP policies interacted with different levels of spatial aggregation. Specifically, we construct policies at the ISO, county, and 5-digit zip code levels for six of our seven markets and calculate the ISO-wide renormalized R^2 for each policy at the node level.[27]

Note that, because we do not have node-level load data, we are unable to simulate equilibria under true real-time pricing. Therefore, as described in Section 4, we now use the "observed" basis rather than the "equilibrium" basis. That is, we are evaluating prospective policies' ability to match observed price patterns *without* accounting for demand response to the implied price changes.

Figure 10 shows the results of this exercise in the PJM market. In general, R^2 values are somewhat below those seen in earlier sections, although keep in mind that the metric is not directly comparable to the other out-of-sample analyses in Section 5 because it uses observed wholesale prices and because those analyses were all carried out on ISO-level data. Collapsing prices to the ISO-level mutes locational price variation, which slightly overstates the performance of TOU policies. Using nodal price data, we see modest gains from locational

[27] CAISO does not make geolocation data available for nodes in its network, so we are unable to replicate this exercise in California.

Figure 10: Locational Pricing – PJM



Note: Each point represents the node-level observed-basis renormalized R^2 for a given TOU or CPP policy in a given year. All policy regressions are of observed real-time locational marginal prices on interactions of policy indicators with spatial indicators. Season splits the year into two six-month periods, beginning in April and October. Versions for all other markets except CAISO are found in Appendix Figures C.1-C.5.

policy complexity. With a single ISO-wide price, mean (median) R^2 values from the four panels above across all markets are 0.044, 0.049, 0.045, and 0.091 (0.035, 0.033, 0.036, and 0.077) for the four pricing schemes, respectively. With county-level pricing, these are 0.052, 0.057, 0.052, and 0.122 (0.033, 0.034, 0.042, and 0.123), respectively. Zip code pricing yields similar results.

The CPP policy (where events are still defined at the ISO-level) generally benefits from locational granularity more than the TOU policies. This is as expected, as binding transmission constraints and local plant outages are a frequent cause of wholesale price spikes, but are difficult to predict far in advance. Flexibly pricing critical peak events in those locations most affected by these price spikes more efficiently targets the differential between wholesale and retail prices. In contrast, the daily, weekly, and seasonal variation captured by TOU policies is less heterogeneous across space, leading to a lower benefit from spatial retail pricing.

We also observe significant heterogeneity by market. Appendix Figures C.1–C.5 show the same figure for all other markets in our sample except for CAISO. The same trends appear on aggregate, although while NYISO and MISO see rather large gains from locational pricing, ISO-NE and ERCOT only experience mild benefits.

6.3 Load-Shifting

We also extend our analysis to consider the role of load-shifting, where price changes lead to substitution in demand across hours. This has an ambiguous effect on deadweight loss since changes in price in any one hour can now either worsen or improve welfare losses in adjacent hours depending on the sign of existing distortions in those hours. The effect of load-shifting on *relative* deadweight loss between constant prices and TOU (the primary measure we focus on in Section 5) will be small if either the absolute effects on deadweight loss are small (i.e., if welfare effects for different hours cancel out), or if the effects on deadweight loss are similar as a proportion for the time-invariant and TOU policies. Jacobsen et al. (2020) show how the presence of substitution alters relative deadweight loss depending on correlation between pricing errors and substitutability.[28]

To investigate load-shifting in our setting we expand on the approach used in Section 5 and now write demand as a linear system of all prices: $\mathbf{q}(\mathbf{p}) = \mathbf{d} - \mathbf{H}\mathbf{p}$. Here $\mathbf{p}$ and $\mathbf{q}$ are vectors of prices and quantities and $\mathbf{H}$ is a positive semi-definite matrix of demand derivatives. When $\mathbf{H}$ is diagonal this demand system nests the analysis done in Section 5. To explore load-shifting we set off-diagonal elements (cross-price derivatives) so that a fraction γ_1 of demand shifts to adjacent hours, spread evenly over γ_2 hours in either direction. Varying γ_1 and γ_2 allows us to approximate a range of results in the empirical literature on load-shifting. The vector of constants $\mathbf{d}$ is set such that prices and demand in the baseline data are reproduced. Equilibrium is solved numerically and is a set of prices in each hour such that demand above matches the piece-wise linear supply curve for each hour described in Appendix A.1.

To construct deadweight losses we observe that linear demand systems in this form can be micro-founded on a quasilinear quadratic model of utility (Spence 1976). We compute consumer surplus in this setting following Choné and Linnemer (2020).[29] Producer surplus over the piece-wise linear marginal cost function is computed at the new equilibrium using the approach as in Section 5 above.

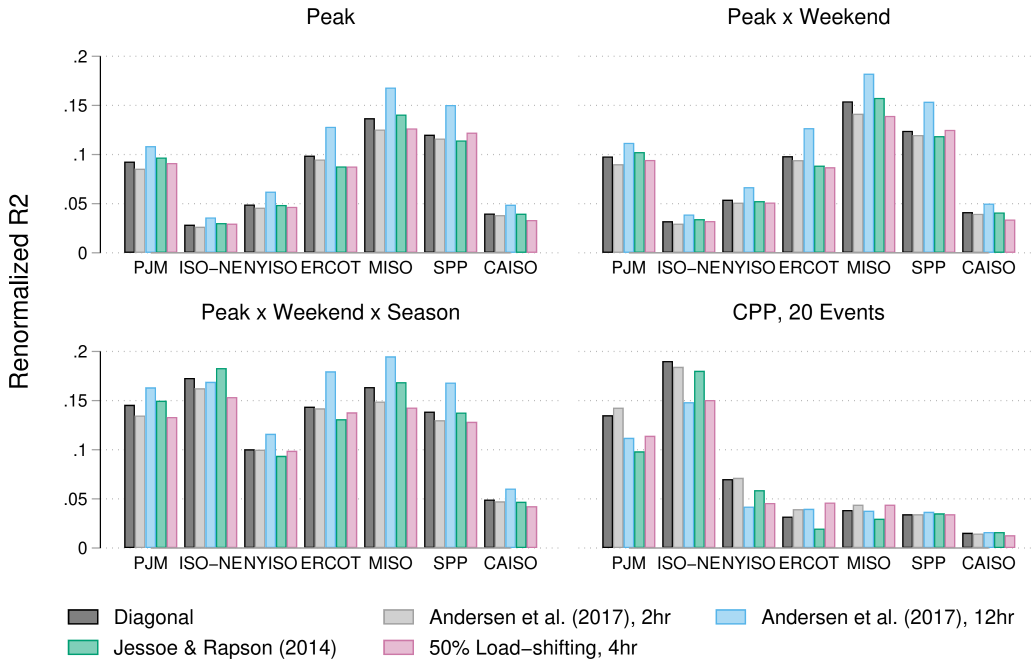
Figure 11 presents welfare results when exploring our model with load-shifting applied to three TOU policies and one CPP policy. For computational reasons we limit the analysis to a single year (2015) where we have complete data. The "Diagonal" benchmark in each panel reproduces out-of-sample R^2 values for 2015 as displayed in Appendix Figure B.1. Note this is a year where CPP performed relatively poorly in many ISOs; we are interested here in comparisons between the diagonal benchmark and the load-shifting cases.

The cases labeled "Andersen et al. (2017), 2 hour" and "12 hour" reflect load-shifting

[28]In the electricity pricing context, correlations between cross-price derivatives and the products of wedges remaining after fitting a TOU policy make the TOU policy more efficient after considering load-shifting. Correlations between the policy itself (i.e., prices in the hours designated as peak) and cross-price derivatives worsen the TOU policy's relative performance. Both types of correlation are likely to be present empirically.

[29]Consumer surplus in this setting is given by $(1/2)(\mathbf{H}^{-1}\mathbf{d} - \mathbf{p})'\mathbf{H}(\mathbf{H}^{-1}\mathbf{d} - \mathbf{p})$.

Figure 11: Efficiency Gains with Various Load-Shifting Patterns



Note: Each bar represents the renormalized R^2 for a given policy and market in 2015. Different colors represent different shoulder substitution patterns, corresponding to different estimates from the literature. Season splits the year into two 6-month periods, beginning in April and October. For the CPP policy, prices are based on and called on day-ahead prices, and each of the 20 CPP events is assigned a unique critical-peak price.

patterns estimated in Andersen et al. (2017) using an experiment in Denmark. They find an average shift in load equal to 29% of the size of the own-price demand response. The shifted load reappears within 2 hours, to as much as 12 hours, on either side of the hours with a price change depending on the setting. Our parameters to span these cases are then $\{\gamma_1 = 0.29, \gamma_2 = 2\}$ and $\{\gamma_1 = 0.29, \gamma_2 = 12\}$. Next we consider an experiment where demand shifts instead reflected complementarities: Jessoe and Rapson (2014) find that information combined with price changes leads to spillovers, making cross-price derivatives negative. In their main treatment there is a reduction in shoulder demand (defined as 2 hours either side) that is 68% as large as the reduction during hours where prices have been increased. We reflect this with $\{\gamma_1 = -0.68, \gamma_2 = 2\}$. We also include an ad-hoc case of $\{\gamma_1 = 0.5, \gamma_2 = 4\}$ to consider even stronger substitution than in the Andersen et al. (2017) study. Other studies that we located from this literature all estimated weaker substitution patterns (and would therefore fall even closer to our primary diagonal case). For example, Ata, Duran, and Islegen (2018) estimate that only about 9% of load response moves to adjacent hours.

Figure 11 shows that when accounting for different types and degrees of load-shifting, and across our datasets, there is relatively little impact on the fraction of deadweight loss recovered by TOU and CPP policies. The mostly-offsetting effects of load-shifting in these scenarios lead to some cases where TOU and CPP look slightly better and some where they look slightly worse. We see average differences from the diagonal R^2 (the mean absolute value of that difference) of -0.004, 0.011, -0.003, and -0.008 (0.005, 0.018, 0.006, and 0.009) for the four alternative substitution patterns, respectively.[30] The largest individual effect in absolute terms is in ISO-NE where broad load-shifting (12 hours either side of peak) leads to a 0.04 lower R^2 under the CPP policy.

Appendix Figure B.9 provides additional detail for a selection of hours, showing first-best equilibrium prices under the various substitution patterns. The relatively small changes in first-best prices (compared to the diagonal case) visible in this figure mirror the small effect we find for load-shifting overall.

7 Conclusion

Wholesale electricity prices vary substantially hour to hour. In theory, there are large efficiency gains that can be realized if this variation in prices can be passed on to final customers, just as they are in nearly all other commodity markets. The question of how to capture some of those efficiency gains has been a perennial topic in energy economics, but it takes on more importance today because the technology to implement more advanced pricing has matured and because the energy transition simultaneously means that prices are likely to become more volatile as renewables play an expanded role in generation, electricity demand will be rising as we electrify transportation and buildings, and much of the new load, be it from electric vehicles or data centers, is likely to be more responsive to variable rates than has been typical in the past. All of this means that it is critical to get prices right, now more than ever.

This paper explores the ability of TOU and CPP pricing policies, which are the lynchpin tools currently available to utilities, to realize these benefits. The paper examines two decades of data from all seven US wholesale power markets, with a focus on measuring goodness of fit statistics associated with alternative rate designs and accounting for out-of-sample forecasting and the equilibrium price effects of alternative pricing policies. These statistics summarize the ability of proposed rates to match wholesale price fluctuations, and as a result they are a simple way to characterize the efficiency gains of alternative tariffs. Our data coverage and our methodological improvements allow us to shed new light on the efficiency potential of time-varying rates.

There is a considerable amount of variation in wholesale electricity prices that is difficult to reflect in TOU and CPP schemes. At the same time, these rates deliver meaningful economic benefits and are largely complementary in nature. Together, simple peak pricing and CPP

[30] In percentage terms, these are -2.8%, 14%, -4.3%, and -6.1% (6.1%, 19.9%, 7.6%, and 10.8%) respectively.

pricing that calls and prices events using day-ahead markets recover about 17% of the efficiency gap with real-time pricing. While a far cry from real-time pricing, regulators and utilities are well-advised to consider them if policies closer to real-time pricing are unavailable. To take full advantage of these efficiency-improving pricing schemes, utilities will need to be sophisticated in how they define peak hours and how they call and price CPP events—current practice is unlikely to capture the benefits we document. Another avenue that utilities might pursue is the introduction of real-time pricing paired with strict price caps. This type of pricing can recover the vast majority of the current efficiency gap while still protecting customers from extreme price fluctuations.

More complex TOU schemes do not improve much on simple peak pricing policies when evaluated out-of-sample and may even backfire. Our results consistently emphasize the importance of more timely rate setting—that is, rates that are able to adjust to current market conditions—rather than having schemes that feature many prices when those prices have to be specified well in advance. Pricing data from past years is hardly informative in predicting detailed price patterns and the risk of overfitting looms large. The patterns of results we document here are largely consistent over time, but the relative efficiency gains of TOU and CPP policies vary substantially across markets.

We believe that these results should be of interest to regulators, utilities, and other stakeholders interested in making electricity markets more efficient. While our analysis calls for significant changes in how utilities operate their CPP programs, technological advances in home automation will make it easier for customers to respond to CPP events that are called at relatively short notice, or to continuous price fluctuations in a system of partial real-time pricing with price caps. This should eventually aid the political feasibility of such changes in policy design.

References

Aigner, Dennis J. 1984. "The Welfare Econometrics of Peak-Load pricing for Electricity." *Journal of Econometrics* 26 (1-2):1–15.

Allcott, Hunt. 2011. "Rethinking Real-Time Electricity Pricing." *Resource and Energy Economics* 33 (4):820–842.

Andersen, Laura M., Lars G. Hansen, Carsten L. Jensen, and Frank A. Wolak. 2017. "Using Real-Time Pricing and Information Provision to Shift Intra-Day Electricity Consumption: Evidence from Denmark." URL http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/into_versus_away_paper.pdf. Manuscript: Stanford University.

Ata, Baris, Serasu Duran, and Ozge Islegen. 2018. "An Analysis of Time-Based Pricing in Retail Electricity Markets." URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2826055. Available at SSRN 2826055.

Badtke-Berkow, Mina, Michael Centore, Kristina Mohlin, and Beia Spiller. 2015. "A Primer on Time-Variant Electricity Pricing." URL https://www.edf.org/sites/default/files/a_primer_on_time-variant_pricing.pdf. Environmental Defense Fund White Paper.

Bailey, Megan R., David P. Brown, Erica Myers, Blake C. Shaffer, and Frank A. Wolak. 2024. "Electric Vehicles and the Energy Transition: Unintended Consequences of a Common Retail Rate Design." URL http://www.nber.org/papers/w32886. NBER Working Paper 32886.

Bailey, Megan R., David P. Brown, Blake C. Shaffer, and Frank A. Wolak. 2023. "Show Me the Money! Incentives and Nudges to Shift Electric Vehicle Charge Timing." URL https://www.nber.org/papers/w31630. NBER Working Paper 31630.

Blonz, Joshua, Karen Palmer, Casey J. Wichman, and Derek C. Wietelman. Forthcoming. "Smart Thermostats, Automation, and Time-Varying Prices." *American Economic Journal: Economic Policy* .

Blonz, Joshua A. 2022. "Making the Best of the Second-Best: Welfare Consequences of Time-Varying Electricity Prices." *Journal of the Association of Environmental and Resource Economists* 9 (6):1087–1126.

Boiteux, M. 1949. "La Tarification des Demandes en Pointe: Application de la Theorie de la Vente au Cout Marginal." *Revue Générale de l'Électricité* 58.

Borenstein, Severin. 2005a. "The Long-Run Efficiency of Real-Time Electricity Pricing." *Energy Journal* 26 (3):93–116.

———. 2005b. "Time-Varying Retail Electricity Prices: Theory and Practice." In *Electricity Deregulation: Choices and Challenges*, edited by Steven Puller and James Griffin. University of Chicago Press: Chicago, IL, 317–357.

Borenstein, Severin and James Bushnell. 2015. "The US Electricity Industry After 20 Years of Restructuring." *Annual Review of Economics* 7:437–463.

Borenstein, Severin and James B. Bushnell. 2022a. "Do Two Electricity Pricing Wrongs Make a Right? Cost Recovery, Externalities, and Efficiency." *American Economic Journal: Economic Policy* 14 (4):80–110.

———. 2022b. "Headwinds and Tailwinds: Implications of Inefficient Retail Energy Pricing for Energy Substitution." *Environmental and Energy Policy and the Economy* 3:37–70.

Borenstein, Severin and Stephen P. Holland. 2005. "On the Efficiency of Competitive Electricity Markets with Time-Invariant Retail Prices." *RAND Journal of Economics* 36 (3):469–493.

Burger, Scott P., Christopher R. Knittel, Ignacio J. Perez-Arriaga, Ian Schneider, and Frederik vom Scheidt. 2020. "The Efficiency and Distributional Effects of Alternative Residential Electricity Rate Designs." *The Energy Journal* 41 (1):199–240.

Cahana, Michael, Natalia Fabra, Mar Reguant, and Jingyuan Wang. 2024. "The Distributional Impacts of Real-Time Pricing." URL `https://mreguant.github.io/papers/Distributional_Impacts_of_Real_Time_Pricing.pdf`. Working Paper.

Choné, Philippe and Laurent Linnemer. 2020. "Linear Demand Systems for Differentiated Goods: Overview and User's Guide." *International Journal of Industrial Organization* 73:102663.

Chu, Chenghuan Sean, Phillip Leslie, and Alan Sorensen. 2011. "Bundle-Size Pricing as an Approximation to Mixed Bundling." *American Economic Review* 101 (1):263–303.

Cicala, Steve. 2015. "When Does Regulation Distort Costs? Lessons from Fuel Procurement in US Electricity Generation." *American Economic Review* 105 (1):411–444.

Davis, Lucas and Lutz Kilian. 2011. "The Allocative Cost of Price Ceilings in the U.S. Residential Market for Natural Gas." *Journal of Political Economy* 119 (2):212–241.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications." *Quarterly Journal of Economics* 135 (2):561–644.

DellaVigna, Stefano and Matthew Gentzkow. 2019. "Uniform Pricing in U.S. Retail Chains." *Quarterly Journal of Economics* 134 (4):2011–2084.

EFI Foundation. 2024. "Managing Unprecedented Electricity Demand Growth on the Path to Net Zero Emissions." URL `https://efifoundation.org/wp-content/uploads/sites/3/2024/04/Load-growth-April-9-2024.pdf`. EFI Foundation Workshop Report.

Enrich, Jacint, Ruoyi Li, Alejandro Mizrahi, and Mar Reguant. 2024. "Measuring the Impact of Time-of-Use Pricing on Electricity Consumption: Evidence from Spain." *Journal of Environmental Economics and Management* 123:102901.

Fabra, Natalia, David Rapson, Mar Reguant, and Jingyuan Wang. 2021. "Estimating the Elasticity to Real-Time Pricing: Evidence from the Spanish Electricity Market." *AEA Papers and Proceedings* 111:425–29.

Faruqui, Ahmad and Sanem Sergici. 2010. "Household Response to Dynamic Pricing of Electricity: A Survey of 15 Experiments." *Journal of Regulatory Economics* 38:193–225.

Faruqui, Ahmad and Sylvia Tang. 2021. "Best Practices in Tariff Design." URL https://www.brattle.com/wp-content/uploads/2021/06/21941_best_practices_in_tariff_design_-_a_global_survey.pdf. Brattle Group Presentation.

Fowlie, Meredith, Mar Reguant, and Stephen P. Ryan. 2016. "Market-Based Emissions Regulation and Industry Dynamics." *Journal of Political Economy* 124 (1):249–302.

Gillan, James. 2017. "Dynamic Pricing, Attention, and Automation: Evidence from a Field Experiment in Electricity Consumption." URL http://jamesgillan.info/s/JMP_Gillan_most_recent_draft.pdf. Manuscript: UC Berkeley.

Hogan, William W. 2002. "Electricity Market Restructuring: Reforms of Reforms." *Journal of Regulatory Economics* 21 (1):103–132.

Holland, Stephen P. and Erin T. Mansur. 2006. "The Short-Run Effects of Time-Varying Prices in Competitive Electricity Markets." *The Energy Journal* 27 (4):127–155.

Imelda, Matthias Fripp, and Michael J. Roberts. Forthcoming. "Real-Time Pricing and the Cost of Clean Power." *American Economic Journal: Economic Policy* .

Ito, Koichiro. 2014. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review* 104 (2):537–563.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2018. "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *American Economic Journal: Economic Policy* 10 (1):1–29.

———. 2023. "Selection on Welfare Gains: Experimental Evidence from Electricity Plan Choice." *American Economic Review* 113 (11).

Jacobsen, Mark R., Christopher R. Knittel, James M. Sallee, and Arthur A. van Benthem. 2020. "The Use of Regression Statistics to Analyze Imperfect Pricing Policies." *Journal of Political Economy* 128 (5):1826–1876.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2021. *An Introduction to Statistical Learning, 2nd Edition*. Springer.

Jessoe, Katrina and David Rapson. 2014. "Knowledge is (Less) Power: Experimental Evidence from Residential Electricity Use." *American Economic Review* 104 (4):1417–1438.

Joskow, Paul L. 1976. "Contributions to the Theory of Marginal Cost Pricing." *Bell Journal of Economics* 7 (1):197–206.

Kahn, Matthew E. and Frank A. Wolak. 2013. "Using Information to Improve the Effectiveness of Nonlinear Pricing: Evidence from a Field Experiment." URL https://ideas.repec.org/p/pra/mprapa/106089.html. MPRA Paper 106089, University Library of Munich, Germany.

North American Electric Reliability Corporation. 2023. "2023 Long-Term Reliability Assessment." URL https://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/NERC_LTRA_2023.pdf.

Reguant, Mar. 2019. “The Efficiency and Sectoral Distributional Impacts of Large-Scale Renewable Energy Policies.” *Journal of the Association of Environmental and Resource Economists* 6 (S1):S129–S169.

Scheinkman, José A. and Wei Xiong. 2003. “Overconfidence and Speculative Bubbles.” *Journal of Political Economy* 111 (6):1183–1220.

Schittekatte, Tim, Dharik Mallapragada, Paul L. Joskow, and Richard Schmalensee. 2024. “Electricity Retail Rate Design in a Decarbonized Economy: An Analysis of Time-Of-Use and Critical Peak Pricing.” *The Energy Journal* 45 (3).

Spence, Michael. 1976. “Product Differentiation and Welfare.” *American Economic Review* 66 (2):407–414.

Stevens, Luminita. 2020. “Coarse Pricing Policies.” *Review of Economic Studies* 87 (1):420–453.

Timmins, Christopher. 2002. “Measuring the Dynamic Efficiency Costs of Regulators’ Preferences: Municipal Water Utilities in the Arid West.” *Econometrica* 70 (2):603–629.

Train, Kenneth E. and Gil Mehrez. 1994. “Optional Time-of-Use Prices for Electricity: Econometric Analysis of Surplus and Pareto Impacts.” *The RAND Journal of Economics* 25 (2):263–283.

Triolo, Ryan C. and Frank A. Wolak. 2022. “Quantifying the Benefits of a Nodal Market Design in the Texas Electricity Market.” *Energy Economics* 112:106154.

Wolak, Frank A. 2007. “Residential Customer Response to Real-time Pricing: The Anaheim Critical Peak Pricing Experiment.” URL https://escholarship.org/uc/item/3td3n1x1. UC Berkeley: Center for the Study of Energy Markets.

———. 2011a. “Do Residential Customers Respond to Hourly Prices? Evidence from a Dynamic Pricing Experiment.” *American Economic Review* 101 (3):83–87.

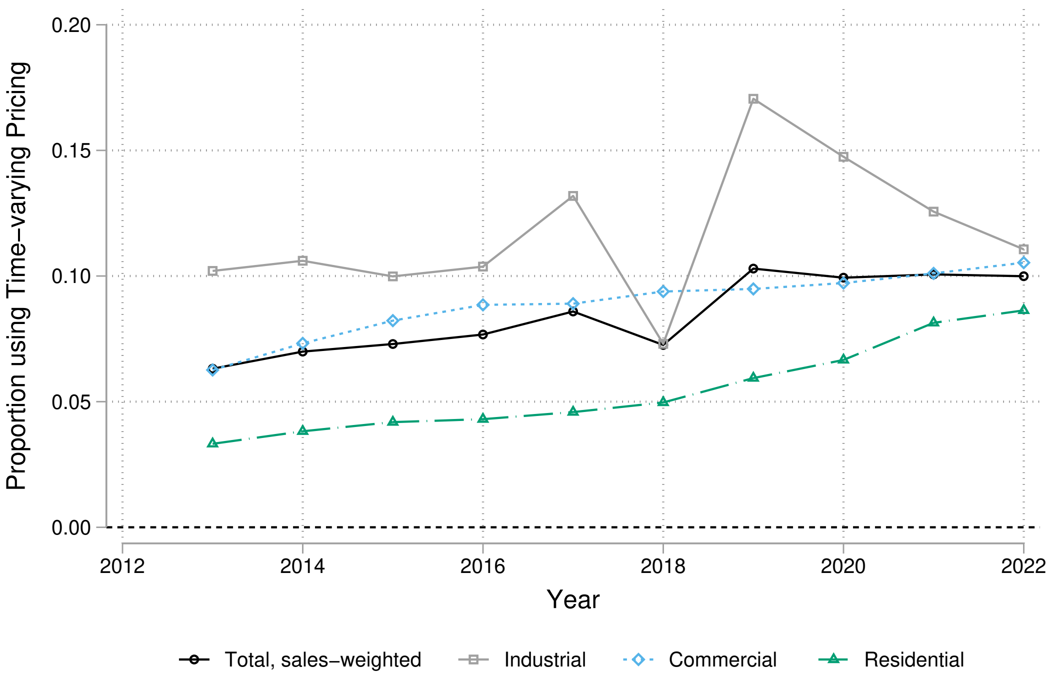
———. 2011b. “Measuring the Benefits of Greater Spatial Granularity in Short-Term Pricing in Wholesale Electricity Markets.” *American Economic Review* 101 (3):247–52.

A Appendix: Additional Methodological Details

A.1 Equilibrium Simulation

To define hourly demand curves we use a linear functional form. We calibrate the slope in each hour using observed average prices/quantities and an assumed elasticity of -0.2 (based on Reguant (2019)), which gives us a slope term of $-0.2(\frac{\overline{q}}{\overline{p}})$ for each hour in the data. We also consider values of -0.1 and -0.3 to assess sensitivity of results. We then calibrate the hourly intercept terms by aligning the observed load in each hour with the average price across the entire market-year. This is implicitly assuming that consumers were facing a flat tariff, set at average price, in the data. In reality most consumers do face flat tariffs, but industrial and commercial consumers are more likely to face time-varying prices than residential consumers. Figure A.1 shows that only 5-10% of load in each year do not face flat prices, although this proportion is increasing modestly over time. Among consumers that do face time-varying prices, industrial consumers are relatively more likely to face true real-time prices and residential/commercial consumers are relatively more likely to face simple time-of-use prices.

Figure A.1: Proportion of Consumers Facing Time-Varying Prices



Note: Each point represents the proportion of consumers in the US that face time-varying prices (including time-of-use pricing, critical-peak pricing, critical-peak rebates, variable-peak pricing, and real-time pricing), calculated by sector. Total is an average of the sectoral estimates weighted by annual usage (in GWh). Source: Form EIA-861 (https://www.eia.gov/electricity/data/eia861/).

In light of this, we interpret our demand curve here as an overall response to retail electricity price changes, which is then an average of the demand response from each segment weighted according to its size and likelihood to face retail prices. We selected values based on estimates of residential, commercial, and industrial demand elasticities and shares. See Reguant (2019) (Table 2 and footnote 27) for a brief overview of these.

To generate hourly supply curves, we begin with market-year-level merit order curves. For ease of computation, we fit linear splines with knots at the 75^{th} and 95^{th} percentiles of quantity along the merit order. This specification was selected via grid search over every possible set of 1, 2, or 3 vigintile knots—this set explained the greatest proportion of the variance in our merit order data, with an average R^2 of 0.959 across all market-years. We then horizontally shift the supply curves such that they intersect demand at the observed price and quantity in each hour. This assumes that hourly shocks to the merit order all occur among the lowest marginal cost technologies (e.g., wind/solar).

Finally, we generate the counterfactual hourly equilibria via the intersection of these two curves. We then simulate hourly policy equilibria by crossing the demand curve with a horizontal "enforced" supply curve set at the policy price. Deadweight losses caused by any given policy are computed by integrating to evaluate consumer and producer surpluses.

A.2 Peak-Hour Definition

To create our standard definition of peak vs. off-peak hours in each market, we find the pair of hours that, as a start and end hour for the peak period, on average explain the greatest proportion of the variation in observed wholesale prices in that market. To do so, we employ an exhaustive search best-subsets selection algorithm, an out-of-sample variant of which we use in Section 6.1.

We define a peak policy by its start and end hour. This is any subset of {midnight, 1AM, 2AM, . . . , 11PM} of size 2, where the peak period is defined as those hours weakly after the start hour and weakly before the end hour (or vice versa). There are $\binom{24}{2} = 276$ possible arrangements. For each market-year of our data, we calculate the in-sample R^2 from each of these possible peak definitions. We then collapse to the market-level median R^2 value for each peak definition (of which there are 276×7), and select the definition with the highest median R^2 value in each market as the 7 peak definitions in our analysis.[31] The results of this exercise (local time) are shown in Table A.1. As is often practiced by utilities, peak hours usually run from sunrise until sundown, with the exceptions of ERCOT (which is more concentrated during the hottest hours of the day) and CAISO (which is shifted slightly later in the day).

[31] Selections using the market-level mean R^2 are qualitatively similar.

Table A.1: Peak Hours for Each Market, Inclusive

Market	Start Hour	End Hour
PJM	6am	8pm
ISO-NE	6am	8pm
NYISO	7am	8pm
ERCOT	12pm	6pm
MISO	6am	8pm
SPP	6am	8pm
CAISO	10am	10pm

A.3 CPP Price Schedule Construction

To define in-sample CPP schedules (such as those in Figure 3 and Table 2), we first order days in the year by their mean of the observed real-time prices during their peak period. Then, we define indicator variables for each of the top-20 peak periods and report the R^2 from the regression of real-time prices on those indicator variables. This creates a schedule consisting of the day-specific mean peak-period price during the top-20 peak periods (i.e., following what we call "event-level pricing" in Section 5.3) and the mean of all remaining real-time price observations in non-CPP hours.

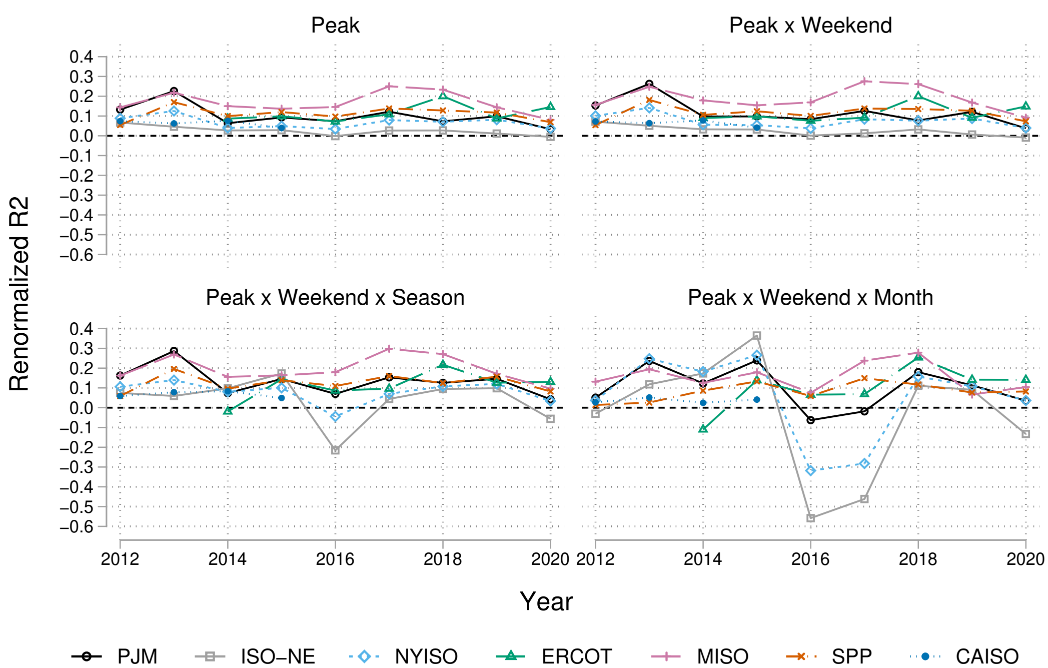
To define out-of-sample CPP schedules, we begin similarly. In most cases, we sort days of the year by their peak period day-ahead price and define the top-n peak periods as critical-peak periods—in Figure 6, we also display results sorted by real-time price. Using the day-ahead market in this way gives our model an imperfect signal of the best CPP events to call given a fixed number of them. However, it is still an *ex post* solution to the true optimal-stopping problem utilities must solve in calling CPP events. We discuss this issue further in Section 5.3.

Then, we define either a single critical-peak price (averaged over all event-hours), event-level critical-peak prices (averaged over hours for each event), or hourly critical-peak prices (each critical-peak hour is priced individually) depending on the structure of the policy. Unless otherwise specified, we report results with event-level prices—Figure 5 compares these structures directly. When using event-level pricing, we define these critical-peak prices using predicted values from a regression of daily peak-period mean simulated equilibrium prices on day-ahead prices. This accounts for risk premia in the forward market and the equilibrium price effects of the critical-peak price. When using single pricing, we use the mean value of these event-level prices. When using hourly pricing, we instead use predicted values from a regression of simulated equilibrium prices on day-ahead prices during peak periods without collapsing to daily peak-period averages (although the regression coefficients are qualitatively similar). In Figure 6, we also display results priced in real-time—in those cases, we use the simulated equilibrium price directly in place of the predicted values based on the day-ahead market.

Finally, we must define prices during off-peak hours. Other than in Section 5.4, we use a flat price for these hours. To define this flat price, we take the mean of the previous three years' simulated equilibrium prices, excluding the $3n$ highest-priced hours (where n is the number of yearly critical peak periods considered). This approximates excluding CPP event hours in the creation of off-peak prices.[32] In Section 5.4, we define TOU schedules underlying CPP policies in much the same way. We construct TOU policies using the simulated equilibrium prices from the three previous years, and exclude the $3n$ highest-priced hours from the training data.

B Appendix: Additional Figures and Tables

Figure B.1: Out-of-Sample Efficiency Gain of Different TOU Designs



Note: Each point represents the equilibrium-basis renormalized R^2 for a given TOU policy in a given market-year, using the simulation methodology described in Section 4. Season splits the year into two 6-month periods, beginning in April and October. CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues. Versions with alternative demand elasticities of -0.1 and -0.3 appear in Appendix Figure B.2. A version using observed-basis R^2 is in Appendix Figure B.3.

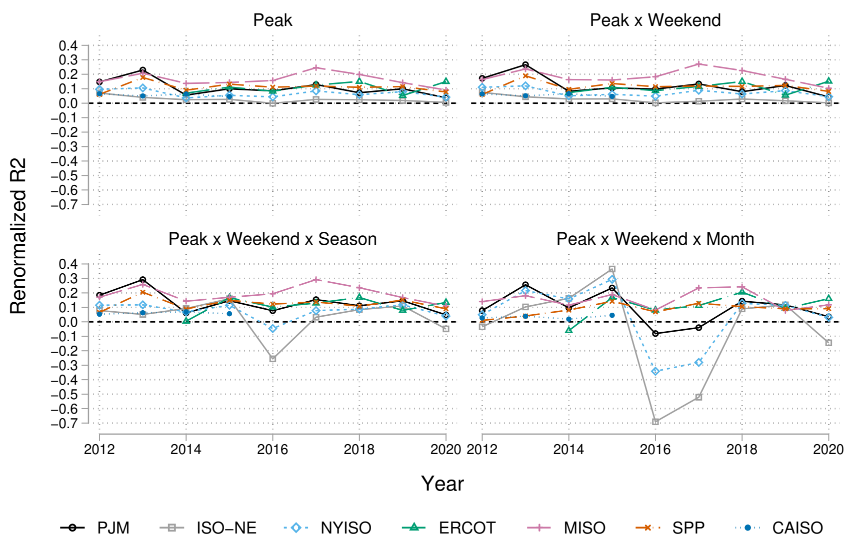
Figure B.1 shows R^2 values for out-of-sample TOU policies. The low R^2 values persist for complex policies. The means (medians) of the four panels in Figure B.1 are 0.095, 0.103, 0.111,

[32] We do not explicitly simulate what peak periods in the training sample would have been critical-peak events because it would require day-ahead data in the training sample, reducing our OoS analysis sample size by three years in each market.

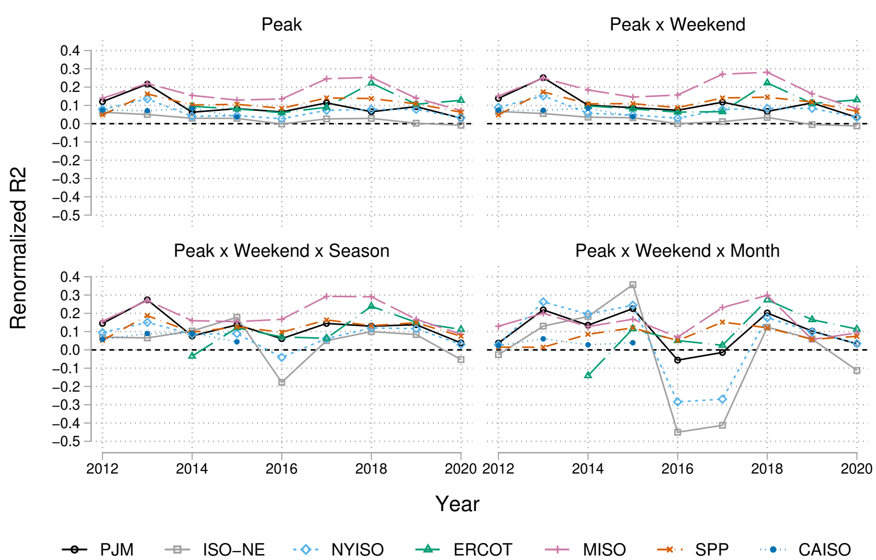
and 0.071 (0.084, 0.092, 0.107, and 0.097) respectively. The standard deviation (interquartile range) of the R^2 values is generally increasing in policy complexity, with values for the four panels of 0.059, 0.066, 0.084, and 0.166 (0.079, 0.085, 0.082, and 0.112) respectively.

Figure B.2: Out-of-Sample Efficiency Gain of Different TOU Designs, Varying Demand Elasticity

(A) −0.1 Demand Elasticity



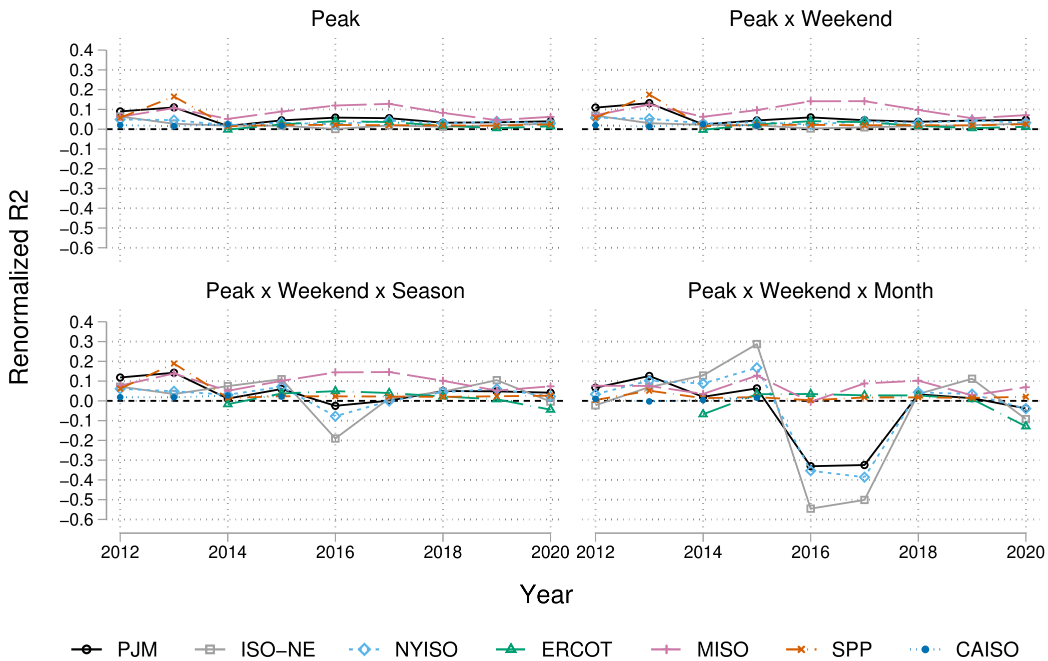
(B) −0.3 Demand Elasticity



Note: Each point represents the equilibrium-basis renormalized R^2 for a given TOU policy in a given market-year, using the simulation methodology described in Section 4. Panel (A) uses a −0.1 demand elasticity and Panel (B) uses a −0.3 demand elasticity. Season splits the year into two 6-month periods, beginning in April and October. CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues. Main version using −0.2 demand elasticity available in Figure B.1.

Figure B.2 shows a variation of our main results in Figure B.1 assuming a demand elasticity of -0.1 or -0.3 (as opposed to our base value of -0.2), in panels A and B respectively. This has little impact on the R^2 results, though it does affect the magnitude of total DWL.

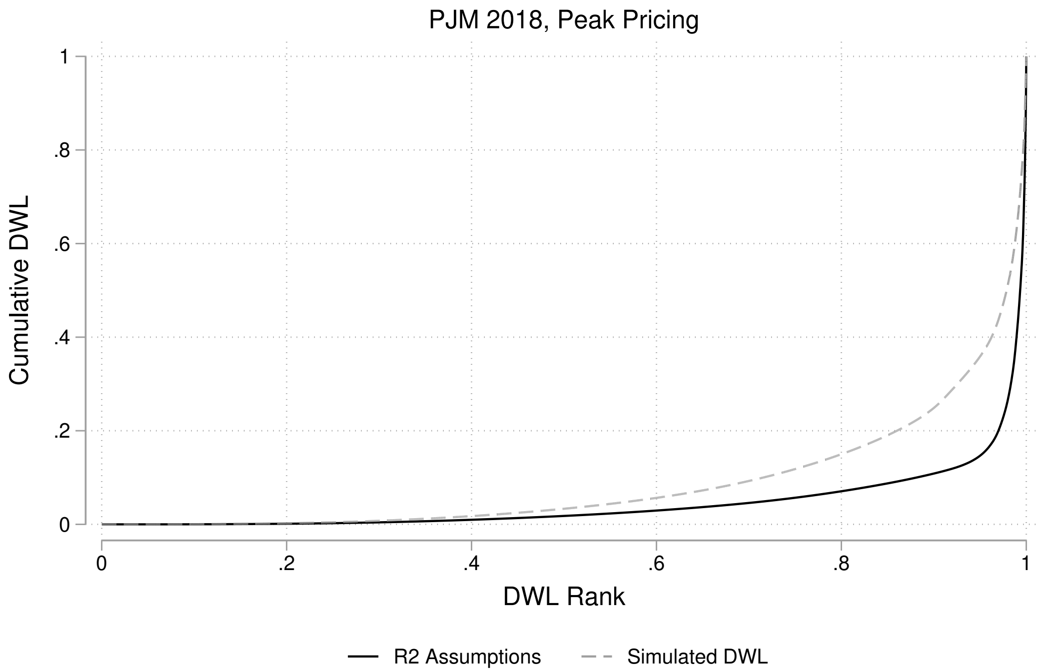
Figure B.3: Out-of-Sample Efficiency Gain of Different TOU Designs, Observed Basis



Note: Each point represents the observed-basis renormalized R^2 for a given TOU policy in a given market-year, using the simulation methodology described in Section 4. Season splits the year into two 6-month periods, beginning in April and October. CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues. Main version using the equilibrium basis available at Figure B.1.

Figure B.3 demonstrates that the result that there are diminishing (and even negative) returns to TOU complexity is similar when using observed wholesale prices instead of simulated real-time prices. Note that, relative to the equilibrium-basis out-of-sample results, we have also changed the functional form of the supply curve: the equilibrium basis out-of-sample results in Figure B.1 assume linear spline supply curves; the observed basis out-of-sample results here assume linear supply.

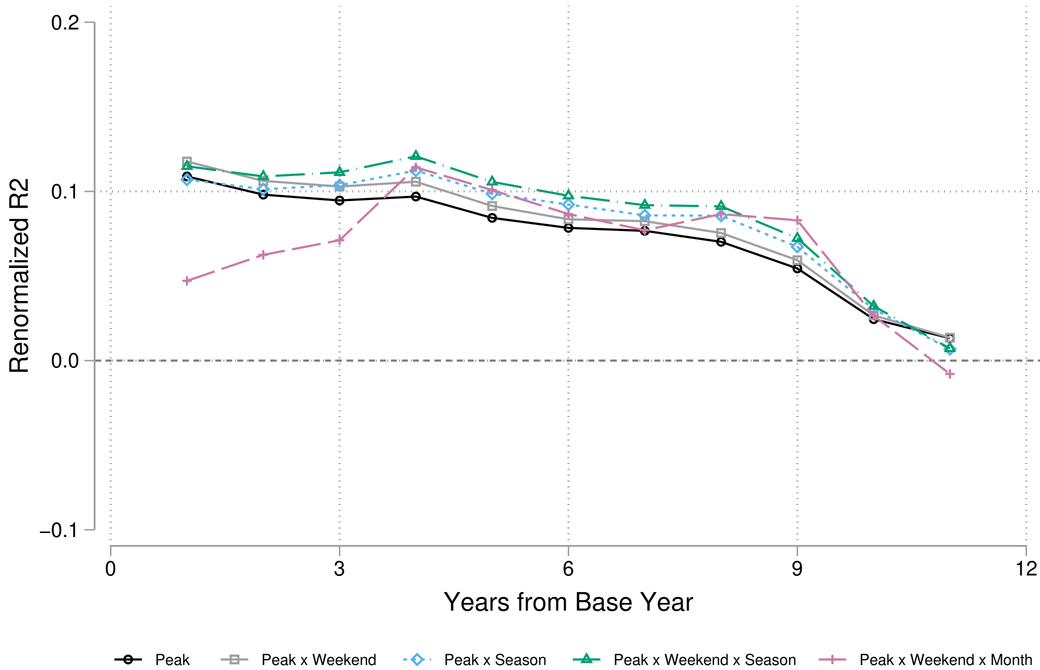
Figure B.4: Cumulative Distribution of In-Sample DWL - PJM - 2018



Note: This figure shows the cumulative distribution function of deadweight loss across hours of the year under linear supply/demand with observed wholesale prices as the benchmark (solid black line) and with spline supply/linear demand with simulated equilibrium prices as the benchmark (dashed grey line), for an in-sample peak/off-peak policy in the PJM market in 2018.

Under the observed wholesale basis, the "target" price schedule is the observed marginal cost of generation. However, the first-best equilibrium price is that which would have been reached under full real-time pricing. When the supply curve is sloped, these two prices are generally not the same. The observed wholesale basis overweights the deadweight loss associated with the most extreme outlier hours, which, under real-time pricing, would have experienced a large reduction in load due to demand response to the higher price. Simulated equilibrium prices have a smaller variance than observed wholesale prices, so the estimated deadweight loss from an intermediary policy is more evenly distributed across hours (Figure B.4).

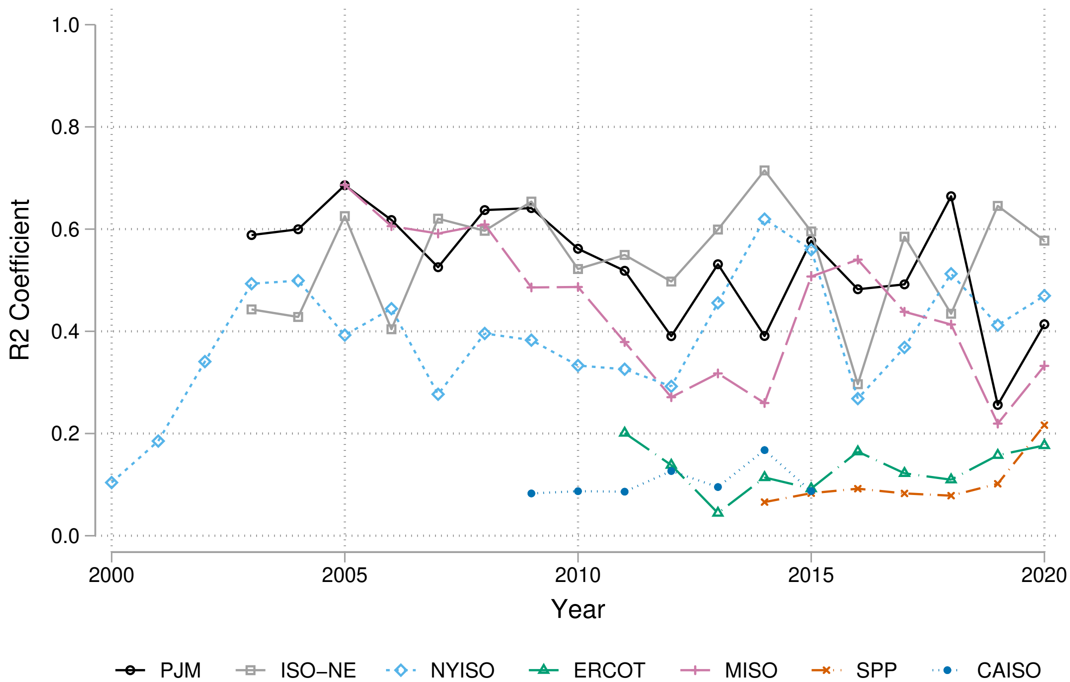
Figure B.5: Varying the Size of the Training Sample



Note: Each point represents the mean renormalized R^2 across all markets with a given policy and number of years of training data, using the simulation methodology described in Section 4. Pre-collapsed data is unbalanced across markets due to varying data availability and unbalanced across levels of the horizontal axis because, with a temporally finite sample, fewer years can support larger training samples.

Figure B.5 presents the mean renormalized R^2 across all market-years of five different TOU policies while varying the size of the training sample. The horizontal axis denotes the number of prior years that constitutes the training sample—note that our inclusion of all possible years of test data causes the panel of results to be unbalanced (with smaller values on the horizontal axis containing more R^2 observations). Our main specification corresponds to a value of 3 along the horizontal axis. We do not see a significant positive or negative trend in R^2 from expanding or contracting the number of years of training data around this point. We do see a decline after 8 years, but this excludes most potential years of test data, especially from CAISO and ERCOT (which can respectively support at most 6 and 9 years of training in our data). Additionally, any reasonable weighting of the training data would prioritize more recent information, so we do not consider this a threat to our empirical strategy. Finally, we also notice that the most complex policy studied (peak $\times$ weekend $\times$ month) improves both in absolute terms and relative to the other policies significantly when moving from 1 to 4 years of pre-period data. This is because, with few years of training data, large spikes in the observed prices can be extremely influential on the policy price on particular tariff levels, causing overfitting. After enough years of data have been added, this policy is “smoothed” significantly, and more closely resembles the policy prices from simpler policies (e.g., peak $\times$ weekend $\times$ season).

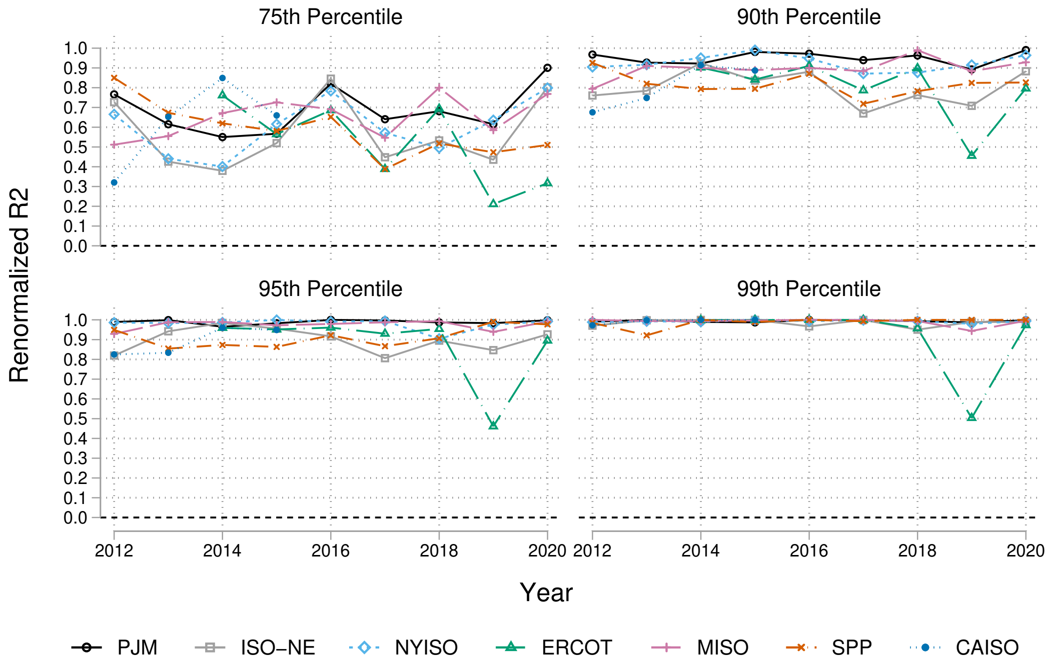
Figure B.6: R^2 Coefficients from a Regression of Real-Time on Day-Ahead Prices



Note: Each point represents the R^2 from a regression of ISO-level real-time prices on day-ahead prices for PJM, ISO-NE, NYISO, ERCOT, MISO, SPP, and CAISO over the period 2000-2020 (with data coverage varying as described in Section 3).

Figure B.6 shows the R^2 from regressions of ISO-level real-time prices on day-ahead prices. These values are significantly below 1, which indicates that there is still significant variation in real-time prices that cannot be predicted even a day in advance. As is visually apparent, certain ISOs' forward markets explain more of the variation in spot prices than others. For example, SPP's base load has a large wind generation share, which is difficult to predict, whereas ISO-NE has larger nuclear and hydroelectric generation shares, which are more predictable. This may partially explain differences in expected efficiency gains from CPP policies. There is typically a small difference in mean between these two market prices—this is a risk premium. Our interest is in whether or not day-ahead prices can accurately *predict* real-time prices, so, as long as this risk premium is consistent in expectation, this is not an issue.

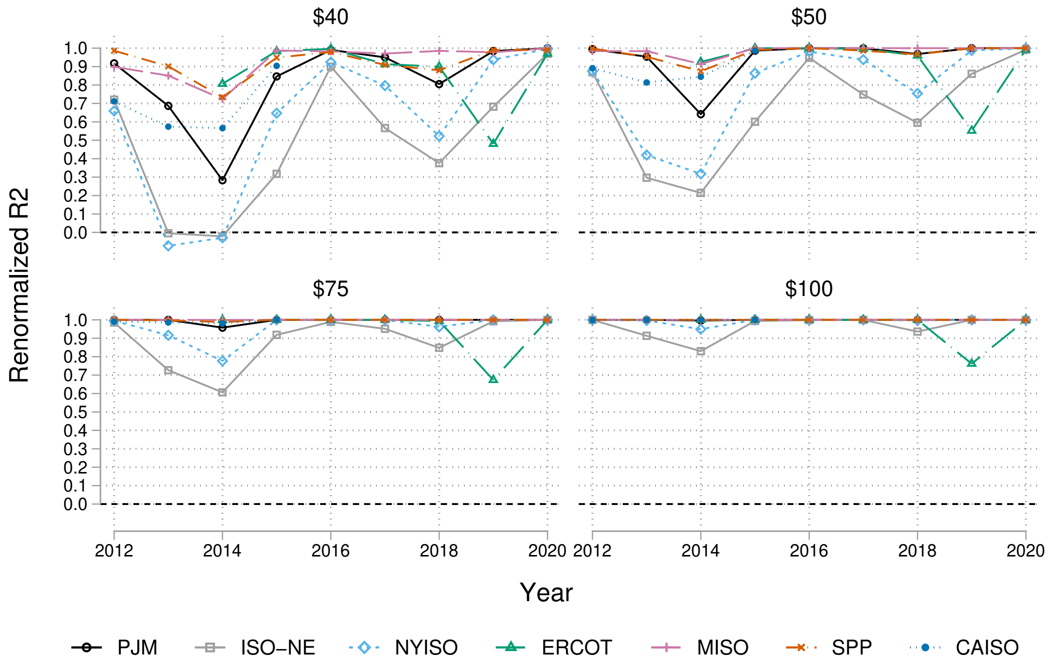
Figure B.7: Real-Time Pricing with Price Caps – In-Sample Caps



Note: Each point represents the equilibrium-basis renormalized R^2 for real-time pricing with a given price cap in a given market-year, using the simulation methodology described in Section 4. CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues. Panels represent real-time pricing schedules with price caps set at the 75^{th}, 90^{th}, 95^{th}, and 99^{th} percentiles of simulated equilibrium prices from the current year, respectively. A version using out-of-sample price caps appears in Figure 8. A version using flat price caps appears in Appendix Figure B.8.

Figure B.7 presents a variation on our analysis of real-time pricing with price caps, instead defining price caps using percentiles of in-sample equilibrium prices. This is not achievable in practice and therefore not as realistic possibility, unlike the results presented in Figure 8—indeed, we are using information that the policymaker does not have access to in order to set these caps. However, the use of this is that we have fixed the proportion of hours affected by the price cap to be equal to exactly $(100-x)\%$ of hours in each panel. These four panels have mean (median) R^2 values of 0.605, 0.863, 0.937, and 0.981 (0.616, 0.888, 0.96, 0.994), respectively. Based on this, we conclude that real-time pricing with price caps binding less than 10% of the time still recover the overwhelming majority of deadweight loss compared to a flat baseline.

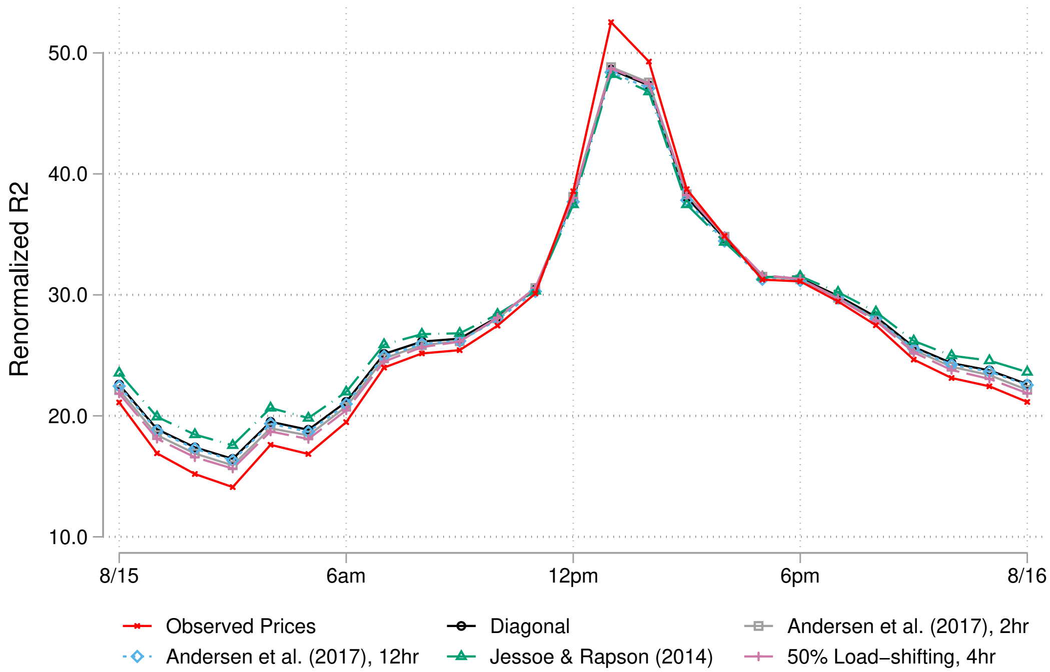
Figure B.8: Real-Time Pricing with Price Caps – Flat Caps



Note: Each point represents the equilibrium-basis renormalized R^2 for real-time pricing with a given price cap in a given market-year, using the simulation methodology described in Section 4. CAISO ends in 2015 and ERCOT begins in 2014 due to data availability issues. A version using out-of-sample price caps appears in Figure 8. A version using in-sample price caps appears in Appendix Figure B.7.

Similarly, Figure B.8 presents a variation of our real-time pricing with fixed-value price caps set at \$40, \$50, \$75, and \$100. As Table B.1 shows, the distribution of prices varies significantly across the years of our sample. Therefore, the proportion of hours in which the cap binds is variable, as in Figure 8. Some market-years using the \$40 cap result in negative R^2 values—this is because the cap is binding in many hours during those market-years (in some cases, the cap is below the median price).

Figure B.9: Simulated Prices with Various Load-Shifting Patterns - PJM - August 15, 2015



Note: Each point represents an hourly simulated equilibrium price of electricity under true real-time pricing from PJM on August 15, 2015. Different colors represent different shoulder substitution patterns, corresponding to different estimates from the literature. Points in red represent observed wholesale prices directly from the data.

Figure B.9 shows a comparison of simulated real-time prices under different shoulder-substitution patterns alongside observed wholesale prices from an example day. We find that these simulated equilibrium prices broadly follow the observed wholesale price, and even more so strongly bunch together with each other. These different calibrations typically diverge in price more during periods of low demand and bunch more tightly in periods of high demand. During those peak hours, any simulation that passes real-time prices onto retail consumers will see a stark response to avoid progressing too far on the steep portion of the merit order curve. Compared to the other calibrations, simulations using Jessoe and Rapson (2014) typically fall on the opposite side of the the no-shoulder substitution (diagonal) case—this is because they find that electricity in adjacent hours are complements, where the other studies found that them to be substitutes.

Table B.1: Percentiles of Simulated Equilibrium Prices

Market	75th	90th	95th	99th
PJM	35.83	47.92	55.13	57.17
ISO-NE	43.86	59.36	69.97	96.47
NYISO	41.11	60.1	71.77	79.07
ERCOT	28.35	36.71	42.27	51.9
MISO	29.82	37.89	44.7	51.95
SPP	27.91	33.8	39.19	55.07
CAISO	38.7	45.96	53.5	78.04

Note: Table shows the 75$^{\text{th}}$, 90$^{\text{th}}$, 95$^{\text{th}}$, and 99$^{\text{th}}$ percentiles of simulated equilibrium prices averaged across the years of our sample for each market in dollars per megawatt hour.

Table B.1 gives the mean across all years of the sample of the 75$^{\text{th}}$, 90$^{\text{th}}$, 95$^{\text{th}}$, and 99$^{\text{th}}$ percentiles of simulated equilibrium prices for each market. This (roughly) gives the average value of the price caps being implemented in Figures 8 and B.7 for each market.

C Appendix: Repeated Exhibits

Table C.1: ISO-level R^2 from In-Sample Electricity Tariff Regressions, Part 1

Panel A: PJM										
Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	0	0	0	0	0	0	0	0
Peak	.	.	.261	.216	.206	.136	.171	.199	.127	.128
Peak x Weekend	.	.	.291	.245	.240	.160	.192	.208	.143	.141
Peak x Season	.	.	.277	.225	.216	.150	.185	.282	.207	.153
Peak x Weekend x Season	.	.	.308	.257	.250	.177	.207	.291	.224	.168
Peak x Weekend x Month	.	.	.365	.335	.437	.282	.263	.391	.370	.293
Hour x DoW x Month	.	.	.601	.601	.665	.481	.490	.611	.558	.493
CPP, 20 Events	.	.	.168	.123	.158	.305	.177	.172	.231	.259
Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.092	.114	.120	.017	.050	.131	.075	.034	.041	.089
Peak x Weekend	.108	.136	.143	.029	.057	.153	.085	.039	.052	.105
Peak x Season	.125	.125	.137	.050	.070	.150	.089	.054	.048	.100
Peak x Weekend x Season	.143	.149	.160	.067	.078	.172	.101	.061	.059	.115
Peak x Weekend x Month	.206	.236	.204	.187	.246	.232	.150	.240	.086	.208
Hour x DoW x Month	.411	.437	.398	.349	.444	.500	.356	.361	.312	.453
CPP, 20 Events	.360	.308	.349	.493	.382	.193	.263	.365	.174	.143
Panel B: ISO-NE										
Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	0	0	0	0	0	0	0	0
Peak	.	.	.131	.130	.130	.099	.189	.119	.092	.097
Peak x Weekend	.	.	.135	.141	.136	.110	.199	.125	.096	.106
Peak x Season	.	.	.147	.146	.134	.103	.201	.230	.236	.102
Peak x Weekend x Season	.	.	.151	.158	.141	.116	.212	.237	.240	.116
Peak x Weekend x Month	.	.	.248	.258	.488	.164	.340	.414	.437	.286
Hour x DoW x Month	.	.	.453	.460	.646	.351	.516	.554	.576	.470
CPP, 20 Events	.	.	.281	.365	.172	.305	.197	.266	.224	.274
Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.067	.076	.031	.02	.029	.029	.032	.015	.029	.050
Peak x Weekend	.074	.081	.034	.028	.032	.036	.033	.019	.032	.053
Peak x Season	.068	.084	.106	.169	.117	.032	.083	.048	.126	.082
Peak x Weekend x Season	.075	.089	.110	.179	.121	.039	.088	.053	.132	.085
Peak x Weekend x Month	.252	.243	.257	.478	.461	.128	.260	.208	.259	.242
Hour x DoW x Month	.402	.443	.385	.572	.581	.334	.382	.349	.401	.412
CPP, 20 Events	.371	.280	.323	.321	.264	.314	.392	.344	.336	.318

Note: Cells of this table present R^2 values for regressions of the hourly ISO-average price of electricity observed in the wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use or critical-peak price schedule (given by the row).

Table C.1: ISO-level R^2 from In-Sample Electricity Tariff Regressions, Part 2

Panel C: NYISO											
Pricing Scheme	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	0	0	0	0	0	0	0	0	0	0	0
Peak	.069	.070	.072	.138	.204	.088	.089	.102	.087	.074	.072
Peak x Weekend	.072	.080	.087	.147	.219	.096	.105	.111	.090	.081	.077
Peak x Season	.073	.087	.081	.140	.212	.098	.095	.105	.157	.145	.078
Peak x Weekend x Season	.076	.098	.099	.150	.227	.106	.113	.116	.161	.153	.085
Peak x Weekend x Month	.142	.184	.235	.307	.295	.295	.179	.147	.296	.261	.184
Hour x DoW x Month	.327	.357	.417	.494	.530	.480	.364	.340	.454	.459	.393
CPP, 20 Events	.303	.386	.279	.218	.211	.247	.408	.291	.207	.168	.209
Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	
Flat Tariff	0	0	0	0	0	0	0	0	0	0	
Peak	.053	.057	.049	.021	.041	.042	.054	.031	.061	.067	
Peak x Weekend	.060	.066	.056	.035	.048	.047	.060	.033	.067	.074	
Peak x Season	.059	.062	.057	.110	.081	.046	.067	.045	.088	.076	
Peak x Weekend x Season	.066	.073	.064	.128	.089	.051	.075	.049	.097	.084	
Peak x Weekend x Month	.139	.125	.164	.373	.401	.117	.136	.219	.197	.234	
Hour x DoW x Month	.337	.338	.314	.482	.551	.334	.345	.351	.378	.433	
CPP, 20 Events	.316	.317	.396	.434	.291	.241	.250	.347	.232	.236	

Panel D: ERCOT										
Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	.	.	.	.	.	.	.	.
Peak	.	.	.	.	.	.	.	.	.	.
Peak x Weekend	.	.	.	.	.	.	.	.	.	.
Peak x Season	.	.	.	.	.	.	.	.	.	.
Peak x Weekend x Season	.	.	.	.	.	.	.	.	.	.
Peak x Weekend x Month	.	.	.	.	.	.	.	.	.	.
Hour x DoW x Month	.	.	.	.	.	.	.	.	.	.
CPP, 20 Events	.	.	.	.	.	.	.	.	.	.
Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.014	.025	.016	.004	.027	.050	.037	.020	.009	.032
Peak x Weekend	.017	.026	.017	.006	.028	.050	.039	.021	.010	.032
Peak x Season	.030	.028	.023	.006	.045	.061	.046	.028	.019	.035
Peak x Weekend x Season	.037	.032	.025	.008	.046	.064	.050	.030	.024	.039
Peak x Weekend x Month	.138	.048	.043	.020	.086	.101	.070	.063	.097	.086
Hour x DoW x Month	.361	.264	.317	.256	.310	.312	.289	.268	.280	.319
CPP, 20 Events	.345	.317	.208	.035	.275	.291	.282	.266	.485	.349

Note: Cells of this table present R^2 values for regressions of the hourly ISO-average price of electricity observed in the wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use or critical-peak price schedule (given by the row).

Table C.1: ISO-level R^2 from In-Sample Electricity Tariff Regressions, Part 3

Panel E: MISO

Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	.	.	0	0	0	0	0	0
Peak	.	.	.	.	.240	.203	.220	.221	.209	.197
Peak x Weekend	.	.	.	.	.298	.244	.265	.259	.230	.217
Peak x Season	.	.	.	.	.247	.211	.228	.248	.253	.211
Peak x Weekend x Season	.	.	.	.	.309	.253	.276	.287	.275	.232
Peak x Weekend x Month	.	.	.	.	.418	.321	.320	.385	.323	.315
Hour x DoW x Month	.	.	.	.	.623	.554	.567	.630	.533	.557
CPP, 20 Events	.	.	.	.	.199	.208	.132	.150	.124	.148

Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.140	.077	.114	.058	.144	.164	.132	.090	.052	.098
Peak x Weekend	.157	.086	.132	.072	.164	.192	.148	.106	.063	.113
Peak x Season	.157	.086	.127	.068	.153	.188	.152	.100	.055	.107
Peak x Weekend x Season	.173	.097	.147	.082	.173	.217	.169	.116	.066	.122
Peak x Weekend x Month	.225	.154	.174	.149	.259	.363	.198	.174	.094	.194
Hour x DoW x Month	.463	.370	.397	.398	.470	.583	.415	.389	.313	.422
CPP, 20 Events	.203	.243	.136	.268	.184	.243	.198	.205	.204	.153

Panel F: SPP

Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	.	.	.	.	0	0	0	0
Peak	.	.	.	.	.	.	.204	.189	.096	.145
Peak x Weekend	.	.	.	.	.	.	.214	.205	.108	.155
Peak x Season	.	.	.	.	.	.	.212	.301	.133	.160
Peak x Weekend x Season	.	.	.	.	.	.	.223	.317	.146	.170
Peak x Weekend x Month	.	.	.	.	.	.	.260	.508	.202	.296
Hour x DoW x Month	.	.	.	.	.	.	.560	.731	.451	.558
CPP, 20 Events	.	.	.	.	.	.	.127	.114	.116	.099

Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.163	.176	.170	.019	.029	.029	.034	.024	.028	.065
Peak x Weekend	.170	.184	.182	.021	.030	.030	.034	.025	.030	.068
Peak x Season	.229	.196	.188	.021	.031	.031	.038	.025	.033	.068
Peak x Weekend x Season	.236	.204	.203	.023	.032	.032	.038	.028	.035	.072
Peak x Weekend x Month	.332	.312	.257	.036	.047	.061	.050	.044	.041	.114
Hour x DoW x Month	.568	.547	.512	.289	.326	.286	.273	.263	.284	.328
CPP, 20 Events	.112	.144	.094	.113	.075	.113	.077	.097	.083	.125

Note: Cells of this table present R^2 values for regressions of the hourly ISO-average price of electricity observed in the wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use or critical-peak price schedule (given by the row).

Table C.1: ISO-level R^2 from In-Sample Electricity Tariff Regressions, Part 4

Panel G: CAISO										
Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	.	.	.	.	.	.	0	0
Peak	.	.	.	.	.	.	.	.	.036	.039
Peak x Weekend	.	.	.	.	.	.	.	.	.037	.040
Peak x Season	.	.	.	.	.	.	.	.	.056	.045
Peak x Weekend x Season	.	.	.	.	.	.	.	.	.058	.047
Peak x Weekend x Month	.	.	.	.	.	.	.	.	.080	.064
Hour x DoW x Month	.	.	.	.	.	.	.	.	.345	.306
CPP, 20 Events	.	.	.	.	.	.	.	.	.155	.101
Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	.	.	.	.	.
Peak	.029	.027	.027	.035	.016	.	.	.	.	.
Peak x Weekend	.030	.027	.029	.038	.016	.	.	.	.	.
Peak x Season	.034	.032	.037	.043	.021	.	.	.	.	.
Peak x Weekend x Season	.036	.032	.039	.046	.023	.	.	.	.	.
Peak x Weekend x Month	.060	.058	.061	.090	.047	.	.	.	.	.
Hour x DoW x Month	.310	.271	.292	.331	.273	.	.	.	.	.
CPP, 20 Events	.082	.152	.103	.101	.128	.	.	.	.	.

Note: Cells of this table present R^2 values for regressions of the hourly ISO-average price of electricity observed in the wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use or critical-peak price schedule (given by the row).

Table C.2: Node-Level R^2 from Electricity Tariff Regressions, Part 1

Panel A: PJM

Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	0	0	0	0	0	0	0	0
Peak	.	.	.228	.166	.153	.103	.113	.130	.087	.080
Peak x Weekend	.	.	.254	.189	.179	.121	.127	.136	.098	.088
Peak x Season	.	.	.244	.173	.160	.113	.122	.185	.140	.096
Peak x Weekend x Season	.	.	.270	.199	.186	.133	.136	.191	.152	.105
Hour x DoW x Month	.	.	.527	.451	.487	.362	.324	.400	.378	.306
County	.	.	.000	.016	.004	.001	.003	.006	.005	.003
Pricing Scheme	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.067	.081	.059	.013	.031	.062	.042	.022	.023	.025
Peak x Weekend	.079	.096	.071	.023	.035	.072	.047	.026	.028	.030
Peak x Season	.093	.088	.068	.038	.041	.070	.049	.035	.027	.028
Peak x Weekend x Season	.105	.105	.079	.051	.046	.081	.056	.039	.033	.033
Hour x DoW x Month	.300	.308	.195	.271	.262	.235	.196	.230	.172	.128
County	.002	.001	.001	.000	.000	.000	.000	.000	.000	.000

Panel B: ISO-NE

Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	0	0	0	0	0	0	0	0
Peak	.	.	.120	.123	.111	.083	.148	.113	.091	.093
Peak x Weekend	.	.	.124	.133	.116	.092	.156	.119	.094	.102
Peak x Season	.	.	.134	.138	.115	.086	.158	.219	.230	.097
Peak x Weekend x Season	.	.	.139	.149	.121	.097	.167	.225	.234	.111
Hour x DoW x Month	.	.	.416	.435	.552	.294	.405	.527	.564	.451
County	.	.	.001	.000	.001	.001	.002	.001	.000	.000
Pricing Scheme	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.065	.068	.030	.020	.028	.028	.030	.014	.028	.048
Peak x Weekend	.071	.072	.034	.028	.031	.036	.031	.019	.031	.051
Peak x Season	.066	.074	.104	.164	.115	.031	.077	.047	.123	.080
Peak x Weekend x Season	.072	.079	.108	.174	.119	.039	.082	.052	.129	.083
Hour x DoW x Month	.386	.393	.378	.563	.575	.328	.357	.340	.392	.399
County	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note: Cells of this table present in-sample R^2 values for regressions of the hourly price of electricity observed in each node in the wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use price schedule (given by the row).

Table C.2: Node-Level R^2 from Electricity Tariff Regressions, Part 2

Panel C: NYISO

Pricing Scheme	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	0	0	0	0	0	0	0	0	0	0	0
Peak	.042	.051	.055	.087	.117	.054	.058	.064	.038	.041	.045
Peak x Weekend	.044	.058	.067	.092	.125	.059	.068	.070	.040	.045	.048
Peak x Season	.044	.064	.062	.088	.121	.060	.062	.066	.070	.080	.049
Peak x Weekend x Season	.046	.073	.076	.094	.130	.065	.074	.072	.072	.085	.053
Hour x DoW x Month	.199	.263	.320	.311	.304	.294	.236	.213	.202	.254	.246
County	.001	.002	.001	.001	.000	.001	.003	.001	.000	.001	.001

Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.035	.036	.028	.017	.026	.020	.031	.021	.024	.027
Peak x Weekend	.040	.043	.032	.028	.031	.023	.034	.023	.027	.029
Peak x Season	.039	.040	.032	.088	.052	.023	.039	.031	.035	.030
Peak x Weekend x Season	.044	.047	.037	.103	.057	.025	.043	.034	.039	.033
Hour x DoW x Month	.223	.217	.178	.387	.355	.161	.199	.240	.151	.171
County	.001	.001	.001	.000	.001	.000	.001	.000	.001	.002

Panel D: ERCOT

Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	.	.	.	.	.	.	.	.
Peak	.	.	.	.	.	.	.	.	.	.
Peak x Weekend	.	.	.	.	.	.	.	.	.	.
Peak x Season	.	.	.	.	.	.	.	.	.	.
Peak x Weekend x Season	.	.	.	.	.	.	.	.	.	.
Hour x DoW x Month	.	.	.	.	.	.	.	.	.	.
County	.	.	.	.	.	.	.	.	.	.

Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.014	.017	.012	.004	.019	.025	.012	.013	.010	.016
Peak x Weekend	.016	.018	.014	.005	.019	.026	.013	.013	.011	.017
Peak x Season	.030	.019	.017	.005	.031	.031	.015	.018	.020	.018
Peak x Weekend x Season	.035	.022	.019	.007	.032	.033	.016	.019	.023	.020
Hour x DoW x Month	.341	.179	.242	.217	.212	.158	.093	.165	.269	.162
County	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note: Cells of this table present in-sample R^2 values for regressions of the hourly price of electricity observed in each node in the wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use price schedule (given by the row).

Table C.2: Node-Level R^2 from Electricity Tariff Regressions, Part 3

Panel E: MISO										
Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	.	.	0	0	0	0	0	0
Peak	.	.	.	.	.166	.126	.138	.164	.107	.099
Peak x Weekend	.	.	.	.	.207	.152	.166	.191	.118	.108
Peak x Season	.	.	.	.	.171	.131	.143	.186	.129	.106
Peak x Weekend x Season	.	.	.	.	.214	.157	.173	.214	.141	.116
Hour x DoW x Month	.	.	.	.	.430	.344	.355	.466	.274	.279
County	.	.	.	.	.000	.001	.001	.000	.000	.000
Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.070	.038	.041	.028	.044	.057	.044	.041	.028	.028
Peak x Weekend	.078	.043	.048	.034	.050	.066	.049	.048	.034	.032
Peak x Season	.078	.043	.046	.032	.047	.065	.050	.045	.03	.031
Peak x Weekend x Season	.086	.048	.053	.039	.053	.075	.056	.053	.036	.035
Hour x DoW x Month	.231	.183	.143	.188	.144	.200	.138	.177	.168	.121
County	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Panel F: SPP										
Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	.	.	.	.	0	0	0	0
Peak	.	.	.	.	.	.	.056	.097	.037	.066
Peak x Weekend	.	.	.	.	.	.	.059	.106	.042	.071
Peak x Season	.	.	.	.	.	.	.058	.156	.052	.073
Peak x Weekend x Season	.	.	.	.	.	.	.061	.164	.057	.077
Hour x DoW x Month	.	.	.	.	.	.	.154	.379	.175	.255
County	.	.	.	.	.	.	.000	.000	.000	.000
Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	0	0	0	0	0
Peak	.102	.084	.092	.016	.023	.021	.022	.016	.019	.031
Peak x Weekend	.107	.087	.098	.018	.024	.022	.022	.018	.020	.033
Peak x Season	.144	.093	.102	.018	.024	.022	.025	.018	.023	.033
Peak x Weekend x Season	.148	.097	.110	.019	.025	.023	.025	.019	.024	.035
Hour x DoW x Month	.356	.259	.276	.244	.248	.206	.179	.181	.193	.159
County	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note: Cells of this table present in-sample R^2 values for regressions of the hourly price of electricity observed in each node in the wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use price schedule (given by the row).

Table C.2: Node-Level R^2 from Electricity Tariff Regressions, Part 4

Panel G: CAISO										
Pricing Scheme	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Flat Tariff	.	.	.	.	.	.	.	.	0	0
Peak	.	.	.	.	.	.	.	.	.024	.033
Peak x Weekend	.	.	.	.	.	.	.	.	.025	.034
Peak x Season	.	.	.	.	.	.	.	.	.039	.037
Peak x Weekend x Season	.	.	.	.	.	.	.	.	.040	.039
Hour x DoW x Month	.	.	.	.	.	.	.	.	.236	.256
County	.	.	.	.	.	.	.	.		.
Pricing Scheme	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Flat Tariff	0	0	0	0	0	.	.	.	.	.
Peak	.024	.013	.015	.025	.010	.	.	.	.	.
Peak x Weekend	.025	.013	.017	.027	.010	.	.	.	.	.
Peak x Season	.029	.016	.021	.032	.014	.	.	.	.	.
Peak x Weekend x Season	.030	.016	.023	.034	.015	.	.	.	.	.
Hour x DoW x Month	.258	.132	.168	.241	.174	.	.	.	.	.
County	.		.	.		.	.	.	.	.

Note: Cells of this table present in-sample R^2 values for regressions of the hourly price of electricity observed in each node in the wholesale market in a given year (given by the column) and a given set of independent variables which define a time-of-use price schedule (given by the row).

Figure C.1: Locational Pricing - ISO-NE

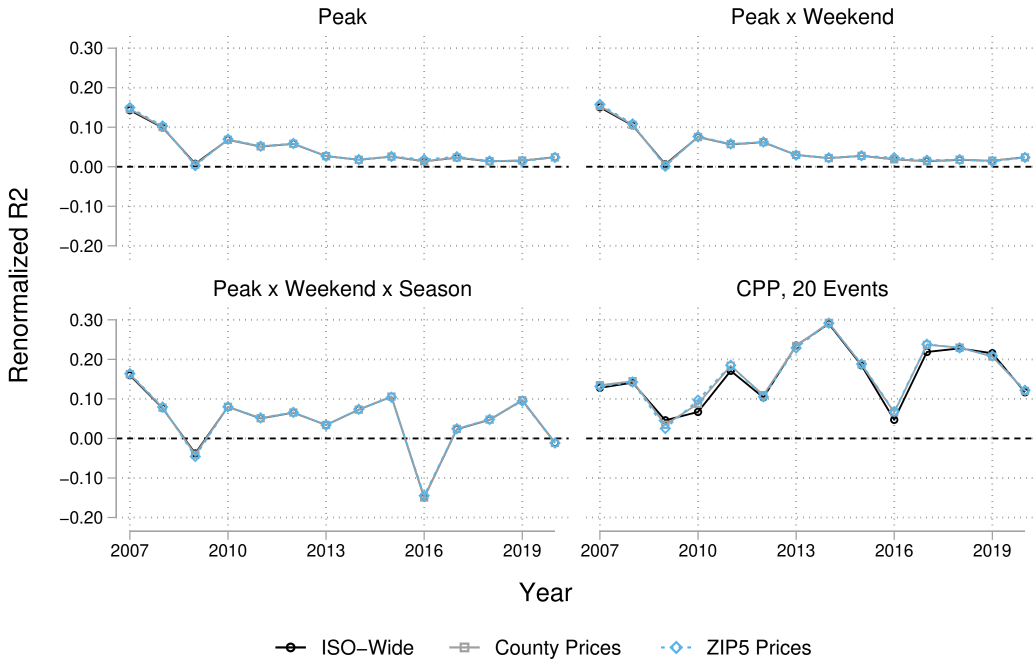


Figure C.2: Locational Pricing - NYISO

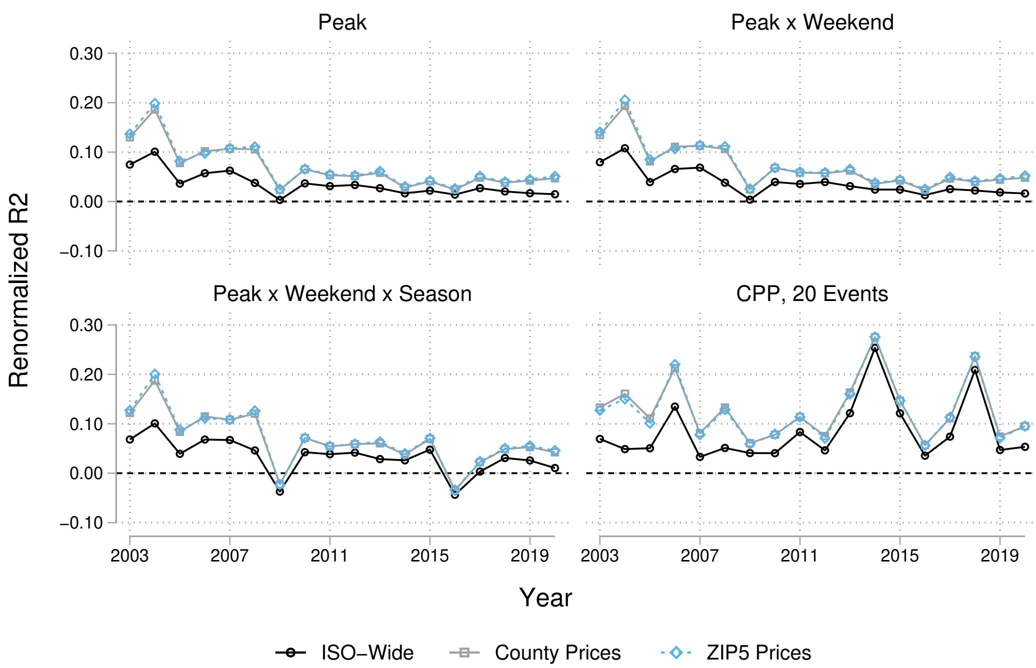


Figure C.3: Locational Pricing - ERCOT

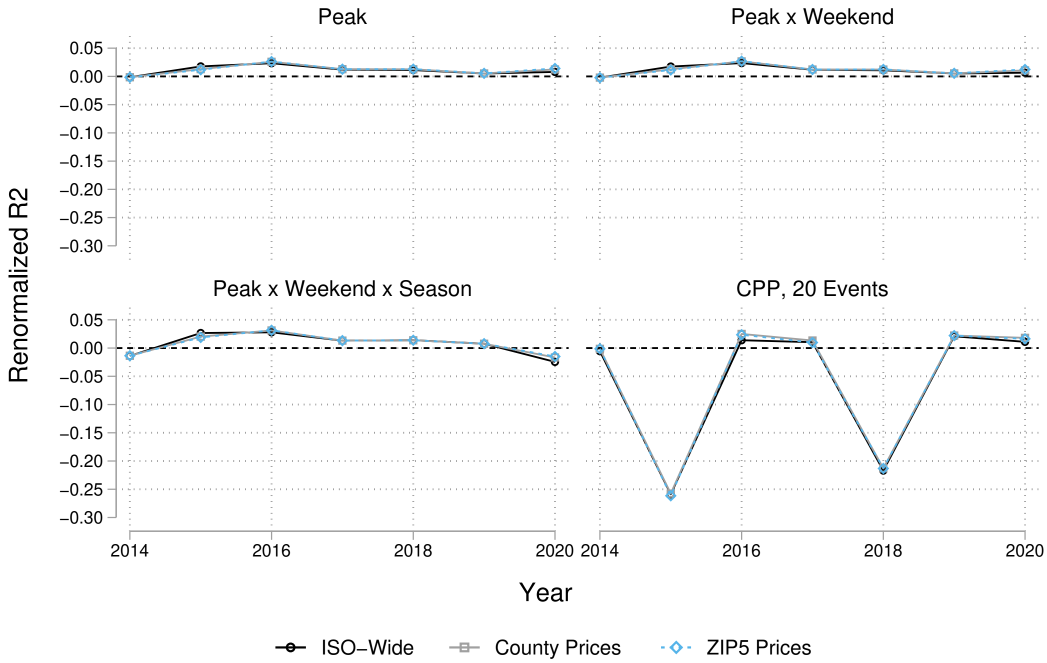


Figure C.4: Locational Pricing - MISO

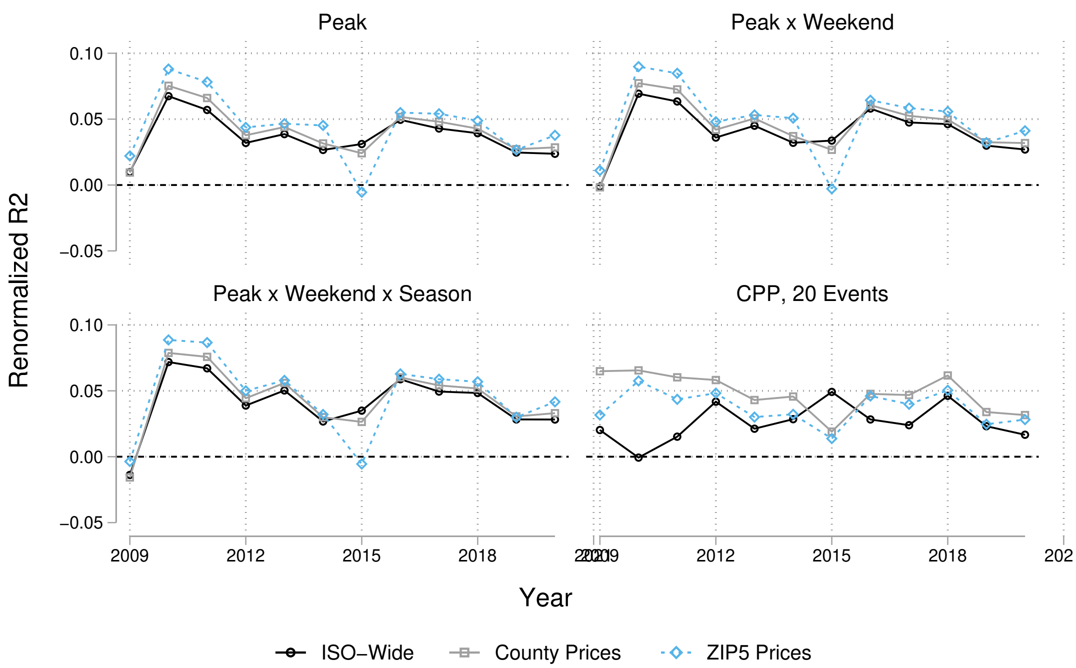
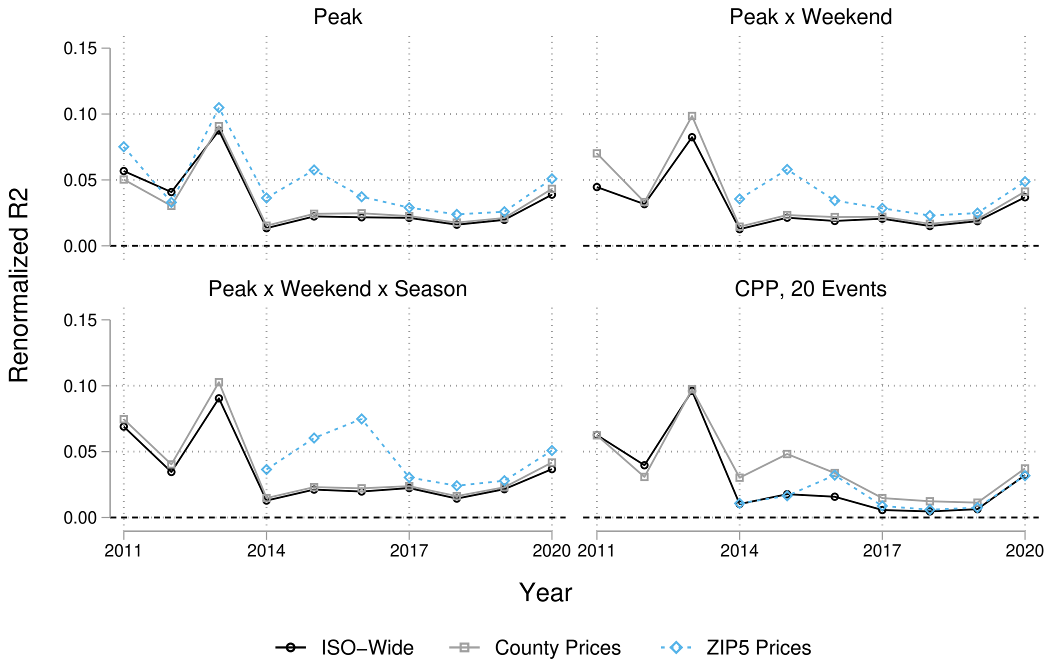


Figure C.5: Locational Pricing - SPP



Contact.

MIT CEEPR Working Paper Series
is published by the MIT Center for Energy and Environmental Policy Research from submissions by affiliated researchers.
For inquiries and/or for permission to reproduce material in this working paper, please contact:

General inquiries: ceepr@mit.edu
Media inquiries: ceepr-media@mit.edu

MIT CEEPR
Center for Energy and
Environmental Policy Research

MIT Center for Energy and
Environmental Policy Research
Massachusetts Institute of Technology
77 Massachusetts Avenue, E19-411
Cambridge, MA 02139-4307
USA

ceepr.mit.edu

MASSACHUSETTS INSTITUTE OF TECHNOLOGY