Soaking Up the Sun: Battery Investment, Renewable Energy, and Market Equilibrium

Preliminary: Please Do Not Cite Without Permission

R. Andrew Butters*
Jackson Dorsey†
Gautam Gowrisankaran‡

February 8, 2021

Abstract

Battery storage offers a potentially valuable complement to renewable energy. Recognizing this, policymakers have recently incentivized and mandated storage as a means to integrate renewable energy and meet climate goals. This paper evaluates the equilibrium value and adoption trajectory of utility-scale batteries using California data, focusing on the impact of falling battery capital costs, complementarities with renewable energy, and market power. We add three key modeling features relative to the literature: (1) a modeling of equilibrium effects from large-scale batteries that includes ramping constraints, (2) a frontier high-frequency forecasting model of load and prices, and (3) linked competitive dynamic equilibrium models of battery adoption and operations. We find that: (1) the value of battery storage increased sharply between 2016-19 as solar generation increased, (2) battery investment exhibits decreasing returns-to-scale—the per-unit value of batteries drops significantly with total installed battery capacity, (3) battery operations increase California's 2018 expected discounted social surplus from the electricity market by $3.8 billion or $2.42 per MWh of solar energy generated, and (4) California would require a 35% subsidy on batteries to meet its 2024 storage mandate of 1.3 GW of power capacity.

JEL Codes: L94, Q40, Q48, Q55
Keywords: energy storage, solar PV, renewable energy, dynamic programming, electricity markets

*Department of Business Economics and Public Policy, Kelley School of Business, Indiana University (e-mail: rabutter@indiana.edu).

†Department of Business Economics and Public Policy, Kelley School of Business, Indiana University (e-mail: jfdorsey@iu.edu).

‡Department of Economics, University of Arizona, HEC Montreal, CEPR, and NBER (e-mail: gautamg2@gmail.com).

We thank Ashley Langer, Derek Lemoine, Erin Mansur, Jim Bushnell, Ken Gillingham, Stan Reynolds, James Mackinnon for helpful comments, discussions, and suggestions.

1 Introduction

Growth in renewable electricity generation has been dramatic over the past 10 years, in the U.S. and worldwide. By displacing generation from fossil fuels, renewables reduce greenhouse gas emissions. However, almost all recent growth in renewables comes from sources such as solar photovoltaics (PV) that are *intermittent*: a solar farm cannot generate electricity after the sun sets, or when a cloud passes overhead. Absent the ability to store electricity, integrating these intermittent sources into the electricity grid requires the capability both to produce electricity at times with low expected renewable production and to adjust production suddenly when renewable production fails. Intermittency adds to the social costs of renewables through the costs of building, maintaining, and operating additional fossil fuel generators (Bushnell and Novan, 2018; Gowrisankaran et al., 2016; Joskow, 2011). Battery storage is a potentially important complement to intermittent renewable energy: it can lower the social costs of renewables by storing energy when renewable production peaks and releasing it when it plummets.

We illustrate these points with data from California, a leader in adding solar generation capacity.[1] Figure 1a displays average electricity demand and Figure 1b displays average solar generation, over the hours of the day and separately for 2015 and 2019. Solar generation increased dramatically over this period, but this generation typically occurs in the middle of the day and not in the evening, when demand is highest. Figure 1c displays *net load*, which is the difference between total demand and renewable generation, and hence the electricity that is supplied by dispatchable generators.[2] Net load in 2019 plummets in the middle of the day but rises again in the early evening to a similar level as in 2015, resulting in a curve with two humps. The double-humped pattern raises the social costs of renewables for two reasons. First, it implies that solar PVs are not producing in the evening when net load, and hence marginal costs, are highest. Second, it increases the ramping costs that generators bear every time they turn on or off (Cullen, 2010; Mansur, 2008; Reguant, 2014).

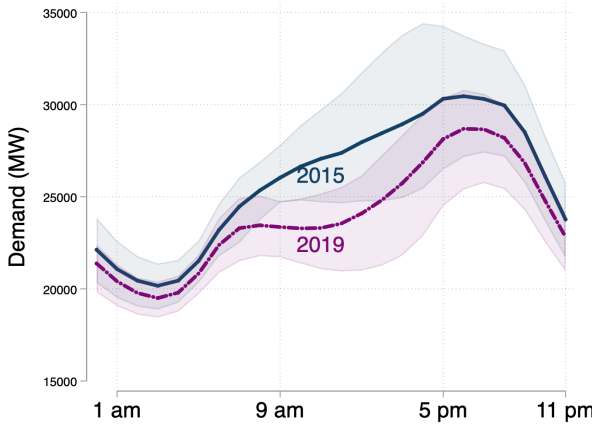
Utility-scale batteries in California can lower the social costs of solar generation by storing energy starting at noon—when the sun is shining—and releasing it from 6PM on. Batteries can help both in using solar to replace fossil fuel generation when marginal costs are highest and in lowering ramping costs. However, the equilibrium value of large-scale batteries is limited because each additional battery, acting as an arbitrageur, will raise prices at noon and lower prices in the evening, thereby flattening the humps and lowering the marginal value of storage.

In conjunction with these trends that affect the revenues from storage, the capital costs of lithium-ion battery cells have dropped 85% from 2010 to 2018 with projections of 50% further cost drops over the next 10 years (Cole and Frazier, 2019; Goldie-Scot, 2019). Not coincidentally, the 2019 Nobel Prize in Chemistry was awarded for the development of lithium-ion batteries. Despite these dramatic cost decreases, the central impediment to utility-scale battery storage

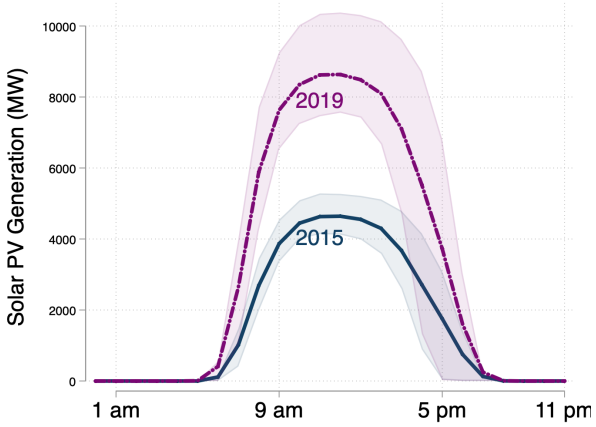
[1] California has a mandate that 60% of generation must be from renewable sources by 2030 and 100% by 2060.

[2] Unlike intermittent generators like wind and solar PV power plants, dispatchable generators can be dispatched on demand at the request of power grid operator. Examples dispatchable generators include natural gas or hydroelectric power plants.

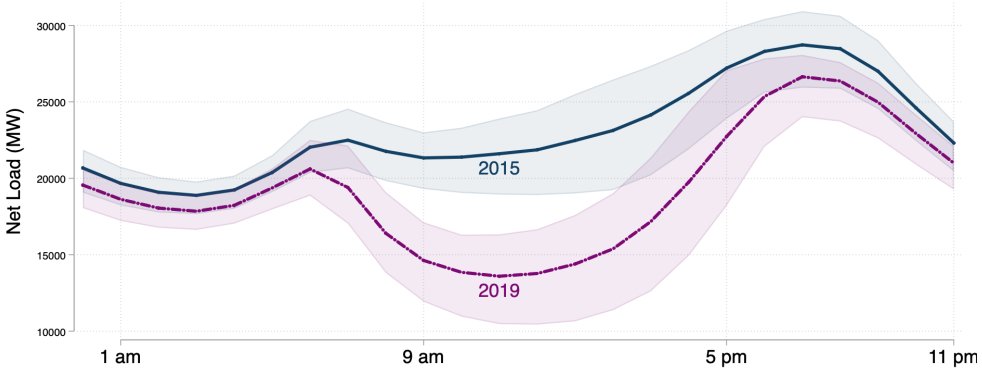
Figure 1: Electricity Demand and Solar Generation by Year in California



(a) Electricity Demand (Load)



(b) Solar Generation



(c) Net Load - "The Duck Curve"

Notes: Panel (a) shows the mean and 25th and 75th percentiles of electricity demand, by year and hour of day. Panel (b) shows the mean and 25th and 75th percentiles of solar generation by year and hour of day. Panel (c) shows the mean and 25th and 75th percentiles of net load, by year and hour of day. Figures calculated by authors from CAISO data.

remains its high capital costs. For this reason, the private market is unlikely to install batteries in the immediate future in the absence of mandates or subsidies.

Recognizing the complementarity between renewable energy and batteries and also these high capital costs, states have paired renewable energy mandates with battery storage requirements. In 2013, California passed a requirement for utilities to procure 1.3 GW of storage power capacity by 2024, with the specific justification that batteries can help optimally integrate renewable energy resources. [3] Other states—notably AZ, MA, NJ, NY, and OR—have also implemented battery procurement requirements as a complement to their renewable energy standards.[4]

This paper has two main goals related to understanding the economics of battery storage. First, we seek to evaluate the equilibrium value and adoption trajectory of utility-scale batteries.

[3]The legislation stated that "additional energy storage systems can optimize the use of the significant additional amounts of variable, intermittent, and offpeak electrical generation from wind and solar energy..."

[4]The Federal Energy Regulatory Commission (FERC) Order 841 requires all electricity markets to remove barriers that would inhibit participation of storage resources in wholesale energy markets.

Specifically, we examine how falling battery capital costs, complementarities with renewable energy, and market power affect wholesale electricity prices, the equilibrium adoption of battery capacity, and the value of batteries given adoption. Second, we seek to understand the role of renewable energy mandates and battery subsidy policies in affecting values and adoption levels. In particular, we quantify the level of subsidies necessary to meet battery mandates and also how much a battery market lowers the costs of meeting renewable energy mandates.

Our study develops a new theoretical and estimation framework that we use to address the above research questions. Our framework contributes three main modeling innovations to the economics and engineering literature on electricity storage, that we believe are crucial for obtaining credible answers. First, we develop a dynamic competitive *equilibrium* battery operations model that allows us to evaluate how much large-scale battery operations would affect the wholesale electricity generation price, and through that, limit the marginal value of additional battery capacity. Second, to solve this equilibrium operations model across different counterfactual battery capacity levels, we develop a high-frequency time-series model of net load and the electricity generation supply curve. Our model allows for the updating of the structural cost parameters over time, which is critical to credibly explaining the relationship between generation and prices in electricity markets. Moreover, our supply curve incorporates ramping constraints, where increases in past generation reduce current marginal costs. Our supply curve estimates are used as an input in our computationally-intensive operations model, implying that both identification and parsimony of the state space are important. Finally, we link our operations model with a battery adoption model. The operations model allows us to understand the value of operating a battery as solar penetration levels and installed battery capacity increases. These values then microfound revenues for our adoption model. The adoption model in turn allows us to understand the social surplus created by having a utility-scale battery market, and how subsidies would affect adoption and social surplus.

Our framework proceeds by developing and solving competitive dynamic equilibrium models of battery capacity adoption and battery operations. Our capacity adoption model specifies an infinite mass of identical potential battery operators. Each potential battery operator can adopt at one point in time and they make the decision of whether to adopt in each period after forming rational expectations regarding future capital adoption costs and equilibrium adoption levels. The gross value of adoption at any point point in time is a function of the battery's ability to arbitrage prices, which depends on the aggregate battery capacity and solar generation. We estimate these values from our operations model.

In our operations model, each battery operator in each five-minute interval of the month observes the current net load and real-time market price, and then decides what percent of its capacity to charge or discharge, up to its technological constraints. To make dynamically optimal decisions, the battery operator must forecast the future price distributions, which are impacted by market-level battery operations decisions. Given that current battery adoption is very low, the equilibrium impact of batteries is not directly observable from market data. However, having estimated electricity generation supply curves, we can derive these equilibrium effects as the

counterfactual of our competitive, dynamic model under optimizing behavior. Our operations model incorporates a number of features that we believe are important in this context: predictable within-day fluctuations in net load; a non-linear marginal cost (or supply) curve of electricity that evolves over time and includes ramping constraints; serial correlation of the shocks to net load and marginal cost curve of electricity due to the dynamics created by weather, transmission congestion, and generator outages; a restriction that battery charge/discharge policies must be based on data that would have been available in real-time to a market participant; a loss in energy from charging and discharging the battery; and the depreciation of batteries from operation, particularly with deep cycles. We estimate the electricity demand process and marginal cost curves using data from the California Independent System Operator (CAISO)—which covers 80% of California's electricity demand—from 2015-19. This setting allows us to assess empirically the complementarity between solar generation and the value of storage. We estimate current and future battery capital costs from data compiled by the National Renewable Energy Laboratory (NREL).

Our results depend crucially on four main identifying assumptions. First, we assume that our market price data allow us to recover the supply curve of electricity for each five-minute interval in our sample. Estimating the supply curve parameters is challenging because we need to decompose price changes into movements along the supply curve and shifts in the supply curve over time. Second, we assume that the net load process and electricity generation supply curves that we identify from the data are structural and hence will continue to hold in the presence of utility-scale batteries. This assumption implicitly rules out the possibility that fossil fuel generators will retire due to large-scale battery storage. We leverage this assumption to evaluate the marginal value of battery operations at counterfactual aggregate battery capacity levels. We let the supply curve be a function of net load and lagged generation (to allow for ramping costs), and allow for serial correlation and heteroskedasticity of the residuals on net load and the electricity supply curve. This rich dependence on observables adds to the plausibility of this assumption. Third, we identify the impact of counterfactual solar adoption with the assumption that the relation between battery storage and solar generation that holds in our data will continue to hold in the future when there is more solar penetration than exists in the data. Finally, to be able to use our data to evaluate the welfare effects of battery operations, we assume that the electricity generation supply curves that we estimate represent the marginal costs of production.

Relation to literature. Our study builds on several literatures. First, our work relates to an engineering and economics literature that investigates the value of storage in wholesale electricity markets. Early engineering papers in this literature modeled the storage decision using a finite-horizon framework and assumed that the storage device operator had perfect foresight about future prices or relied on historical prices when making discharge and charge decisions (Sioshansi et al., 2009; Sioshansi and others, 2011; Walawalkar et al., 2007). More recent engineering studies relax the perfect foresight assumption and model storage decisions given uncertainty

about future prices (Mohsenian-Rad, 2015; Mokrian and Stephen, 2006; Xi et al., 2014).[5] Our operations model uses substantial inputs from the engineering literature and extends this literature by considering the equilibrium effects of large-scale storage and the impact of monopoly or competitive storage markets. A recent economics working paper, Kirkpatrick (2018), empirically estimates the effect of recent utility-scale batteries installations on electricity market prices and transmission line congestion in California. Finally, a contemporaneous economics working paper to ours, Karaduman (2019), also seeks to understand the value of energy storage in an equilibrium setting. Our equilibrium model is complementary to Karaduman (2019), with different approaches to modeling demand and marginal costs which allow for dynamics and equilibrium effects of large battery capacity. Our use of California data, which provides large variation in solar generation, allows us to evaluate directly the impact of solar energy on battery values.

Second, we contribute to an economics literature that explores market impacts of new energy technologies. Bushnell and Novan (2018), Craig et al. (2018), Cullen (2013), Novan (2015) Wolak (2018), and Woo et al. (2016) measure the environmental and market effects of renewable energy generation. Burr (2014), Reddix (2015), Feger et al. (2017), Langer and Lemoine (2018), and De Groote and Verboven (2019), evaluate the impact of solar subsidies on adoption. We add to this literature with a dynamic model of investment and subsidies in battery capacity. While previous work developed dynamic models of investment in a renewable energy, we are the first paper to develop a dynamic model of investment in battery storage. Unlike solar adoption, battery charge and discharge decisions are themselves a difficult, dynamic problem and thus, our adoption model integrates a dynamic operations model.

Third, our work also relates to the literature investigating the time series properties and forecasting of electricity load and real-time prices. In particular, our model for the evolution of prices is motivated by the specific time series features of electricity prices experienced both systematically across most electricity markets (e.g., see Weron (2014) for a comprehensive review) and also for the California CAISO market in particular (Knittel and Roberts, 2005).[6] While no close substitute exists for our model of load and electricity prices, it borrows and combines elements of the model of electricity load and supply used by Kanamura and Ōhashi (2007), together with features of the structural time series approach taken by Pirrong (2012) for other commodities. Given our empirical setting, however, our model in relation to the last two mentioned is at a much higher frequency, accommodates a much richer set of seasonal patterns and dynamic effects, while maintaining a parsimonious set-up that can be estimated in near real-time–both of which are vital for it to serve as the input for dynamic optimization framework.

Finally, our work relates to a literature on the computation of high-frequency dynamic models

[5]Other related work has considered the effects of storage on emissions (Carson and Novan, 2013; Hittinger and Azevedo, 2015; Holladay and LaRiviere, 2018), the value of storage in ancillary service markets (Berrada et al., 2016; Cheng and Powell, 2016; Kazemi et al., 2017), and the role of storage in integrating intermittent renewable power plants (Black and Strbac, 2007; Garcia-Gonzalez et al., 2008; Paatero and Lund, 2005).

[6]Some of the features of wholesale electricity prices including the skewness and volatility are closely connected to the need to constantly balance demand and supply of electricity, and the inability to hold negative inventories (Deaton and Laroque, 1992).

of electricity supply. Following (Cullen and Reynolds, 2017), we simplify the computation burden by solving the social planner single agent dynamic problem, which yields equivalent decisions to the competitive dynamic equilibrium problem. Our computational techniques were developed concurrently with Reynolds (2019).

Findings. We find that the gross social value (not accounting for capital costs) of a small (1000 MWh) utility-scale competitive battery investment would be about $180 per kWh in 2019, where renewable energy holds a 30% share of generation. This contrasts with the reported battery capital costs from NREL of about $355 per kWh in 2019. Thus, battery capacity is not yet at a break-even point. Because of the reduction in battery costs, battery capital costs are expected to reach $180 per kWh by the late 2030s. Thus, a small competitive battery market would not be economically viable in the very near future if California's solar generation were to remain constant.

We find a strong complementarity between the gross social value of utility-scale batteries and renewable energy generation. In particular, if renewable energy share were to equal 44% of generation, the 2019 gross social value of batteries would be $355 per kWh. Combining the mandated increase in renewable energy share in California with the expected decrease in battery costs, the small utility-scale battery installation will be expected to break even by the late 2020s. Due to its role as an arbitrageur, the per-unit value of battery capacity declines significantly in the installed capacity: 1,000 MWh of batteries would add a social value of $180 per kWh while 15,000 MWh of batteries would add a social value of $120 per kWh, both per kWh in 2019. Correspondingly, we find that the 1,000 MWh of batteries would lower evening prices by average of $10 while the figure is $15 for the 15,000 MWh battery fleet.

Despite the fact that we are not too far from reaching a break-even point, battery adoption will likely be slow for several years in the absence of a battery mandate or subsidy. The reason for this is that, given that battery costs are declining, there is an option value of waiting for further technological change instead of adopting at the break-even point. In order to reach California's mandate of 5,200 MWh by 2024,[7] a competitive battery market would require a 35% up-front subsidy. Without a battery mandate, we would expect to see virtually no installed capacity by 2024 and would have to wait until 2032 to see this level of installed capacity.

Overall, allowing a competitive battery market in 2018 would increase discounted social welfare in California by $3.8 billion, which amounts to $2.42 per discounted MWh of solar energy that will be produced in California under their renewable energy mandate. In contrast, a monopoly battery market would install at approximately double the rate of a competitive battery market but would less in discounted social surplus.

The remainder of our paper is structured as follows. Section 2 discusses our institutional features and data. Section 3 discusses our model and estimation. Section 5 presents our results and counterfactuals, and Section 6 concludes.

[7] California's original mandate specifies 1,300 MW of battery power. Using a batteries with four-hour storage duration, this translates to 5,200 MWh of energy storage.

2 Data and Institutional Setting

Worldwide investment in utility-scale batteries is expected to expand rapidly in the coming decade. The trajectory of battery investment over time will hinge on both batteries' expected costs and the expected value (i.e., benefits) of deploying battery resources over time. Thus, we employ two primary data sets to estimate and compute our model battery investment. First, we use data on future projections of utility-scale battery costs. Second, we use historical electricity market data to estimate the value of deploying battery storage and how that value changes over time as more renewable power plants come online. In the following subsections, we describe our primary data sources and discuss relevant institutional details.

Capital Cost of Battery Storage

Some energy storage technologies such as pumped hydroelectric storage have been established for decades, but the majority of recent utility storage installations use battery technologies. More specifically, lithium-ion based batteries now dominate the US market—accounting for over 90% of battery storage capacity (EIA, 2020). Nearly all lithium-ion battery resources were installed after 2014. This recent deployment of batteries has been driven by a surge in private R&D efforts and government policies that encourage or mandate battery investment.

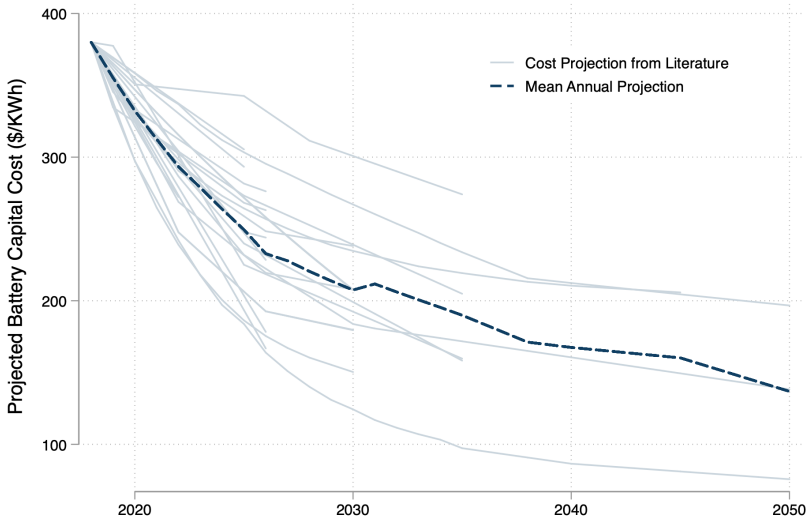
Although the stock of utility-scale batteries is growing quickly, the overall battery fleet remains relatively small. As of 2018, there was only 0.9 GW of aggregate battery power capacity in the US (1.2 GWh of energy capacity), a power capacity similar to that of a typical combined-cycle natural gas power plant (EIA, 2020). Due to the lack of historical data on utility-scale battery capital costs, we use forward-looking projections of battery capital cost from the literature to model the evolution of future battery costs. In particular, we use data from the National Renewable Energy Laboratory (Cole and Frazier, 2019) that compiles cost projections from over 25 publications that consider utility-scale storage costs.

Cole and Frazier (2019) compile data on utility-scale lithium-ion batteries cost projections published between 2016 and 2018. The exact cost of a battery installation will depend on the battery's specifications such as round-trip efficiency and duration. A battery's round-trip efficiency measures the percentage of stored energy that is available for later usage, a more efficient battery typically entails higher costs.[8] The battery's duration indicates the amount of time the battery is able to discharge at its rated power capacity. For example, a 2-hour duration battery could discharge at full power capacity for 2 hours. Although battery systems can be developed with a range of specifications, we follow Cole and Frazier (2019) and focus on the most common type of battery system currently being added in US markets—batteries with 4-hour duration and 85% round-trip efficiency.[9]

[8] Round-trip efficiency is always below 100% because some energy is lost during the charge-discharge cycle.

[9] According to EIA and the DOE Storage Database, the majority of the batteries have a 4-hour storage duration and have round-trip efficiency between 75 and 95%. Batteries with shorter duration (less than 2 hours) are relatively better suited for ancillary service applications, whereas batteries with longer duration (e.g., 4-hour) are relatively better suited for energy arbitrage applications.

Figure 2: Utility-Scale Battery Capital Cost Projections



Notes: Figure constructed by the authors using data from Cole and Frazier (2019). Each transparent line represents a future cost projection from a single publication. The dashed line plots the mean cost projection . The figure reflects all cost projections related to grid battery applications (not electric cars).

Figure 2 demonstrates variation in cost projections for battery storage over time in $/kWh.[10] Each point in the figure represents a normalized cost projection for a particular year from a single publication, and the dashed line plots the mean projection for each year.[11] These data allow us to quantify expectations and uncertainty about future capital cost of storage. In Section 3 we discuss our battery capacity adoption model in more detail and explain how we pair these cost projection data with our model.

The Wholesale Electricity Market

In addition to battery capital costs, another key input to the battery storage investment problem is information about the present discounted value of owning a battery system. To evaluate the value of a battery system, we develop a model of *battery operations*. We then combine our model of battery operations with electricity market data to solve for the optimal sequence of charge-discharge decision and the associated value of storage operations.

The source of data for our operations model is the California Independent System Operator (CAISO).[12] The CAISO, a non-profit independent system operator (ISO), dispatches over 200 million megawatt-hours of electricity to 30 million consumers each year. California deregulated its electricity sector in 1998, and consequently designated the CAISO to manage the state's new

[10] $/kWh costs can be converted to $/kW costs simply by multiplying by the duration (e.g., a $500/kWh, 4-hour battery would have a power capacity cost of $2000/kW)

[11] NREL normalizes the cost projections so that each publications' projection starts at the same baseline cost of $380 in 2018. We only use cost projections related to grid battery applications (not electric cars).

[12] We obtain the data used for the analysis from the CAISO Open Access Same-time Information System (OASIS) portal. OASIS provides real-time data related to the ISO transmission system and its markets, such as system demand forecasts, transmission outage and capacity status, market prices, and market result data.

wholesale electricity market. The CAISO is connected to the Western Interconnection and also regularly imports and exports power to other utilities and power producers across the western US. At each instant, the CAISO balances supply and demand to ensure electricity is delivered to consumers reliably and cost-efficiently. To that end, CAISO runs two distinct market processes: a day-ahead market (DAM) and a real-time market (RTM).

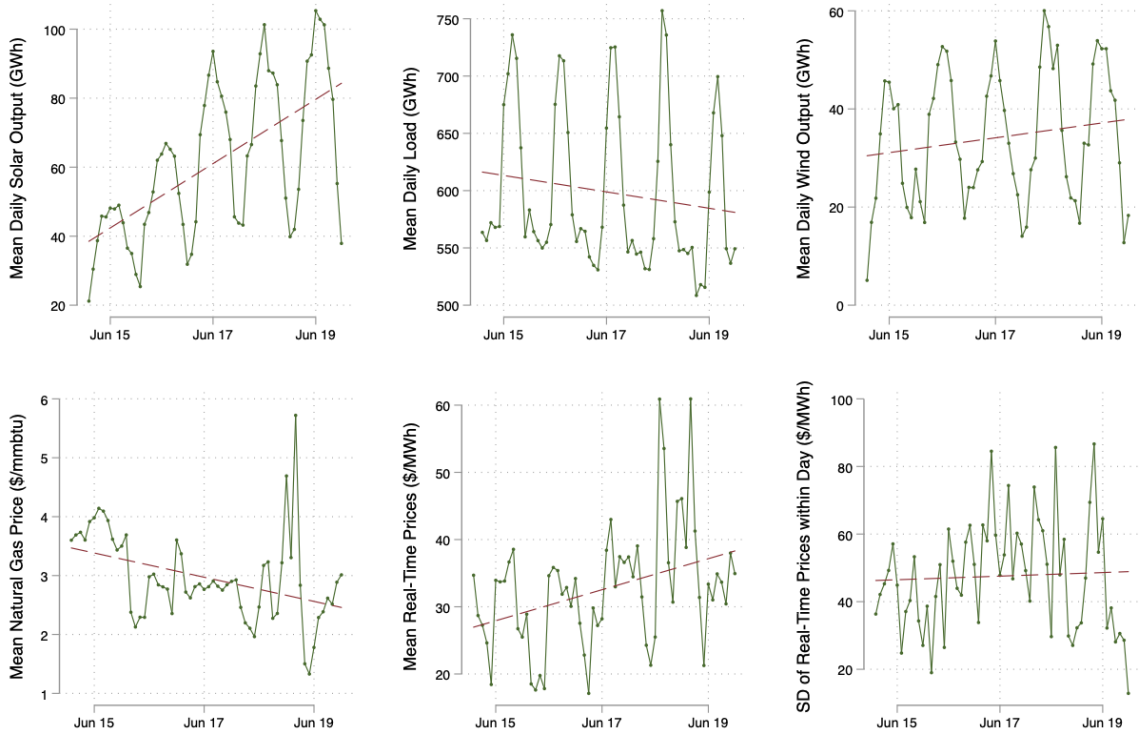
In the day before power is delivered, the system operator conducts 24 energy auctions, one auction for each hour of the day. For each of the 24 energy auctions in the day-ahead market, market participants submit bids to either buy or sell energy and CAISO then computes production quantities and prices to clear each of the hourly auctions. CAISO also uses the day-ahead market to secure energy reserves, these reserves are held ready and available for the ISO to use if needed, and can be called upon to quickly increase or decrease output if there are any unexpected changes in electric supply or demand at the last minute. At the close of the day-ahead market, each power producer is scheduled hourly for a quantity of production (possibly zero), and a capacity allocated for reserves.[13] CAISO allocates production and reserves to meet demand at the lowest cost, subject to reliability and other physical constraints of the system. On the day of energy delivery, CAISO utilizes the real-time market to re-adjust generator production schedules in response to unplanned outages or deviations from the expected day-ahead demand schedule. Market participants can submit real-time market bids until 75 minutes before the delivery hour. During the delivery hour, the system operator continuously updates the demand forecast and dispatches the lowest-cost generators every five minutes.[14] Any unanticipated supply-demand imbalance that occurs within the last 5 minutes before electricity is delivered must be met using reserve generators.

California's grid is currently undertaking a dramatic transition away from fossil-fueled power sources towards renewable resources. The renewable energy transition is, in turn, disrupting market outcomes in CAISO's wholesale markets. The most notable disruption, is the growth of solar PV. As of 2015, California already hosted the largest capacity of solar PV panels in the US, and state lawmakers voted to boost renewable energy further by mandating that 60% of electricity come from renewable sources by the year 2030, and 100% by 2060. The top panel of Figure 3 shows that during the sample period of our study, January 2015 - December 2019, utility scale solar more than doubled from under 40 GWh/day to over 80 GWh/day. At the same time, average demand (load) for electricity remained relatively stable, falling by 7.5%. Average wind power production increased slightly from 28 GWh/day to 36 GWh/day. Figure 3 shows that prices for natural gas, the predominant fossil fuel generation source in CAISO, hovered around $3/MMBtu for much of the sample period. Modest reductions in demand coupled with increasing solar PV investment led to a 13% generation share for solar by 2019. In fact, during some afternoons in 2019, solar PV accounted for over half of the energy supply. While this

[13]Market participants that do not wish to directly participate in the day-ahead market can also submit self-scheduled hourly production plans to the ISO for planning purposes.

[14]Generators can only submit on bid curve to the entire trading hour but the market price can change every five minutes due to demand shocks, transmission congestion, or supply outages.

Figure 3: CAISO Electricity Market Trends



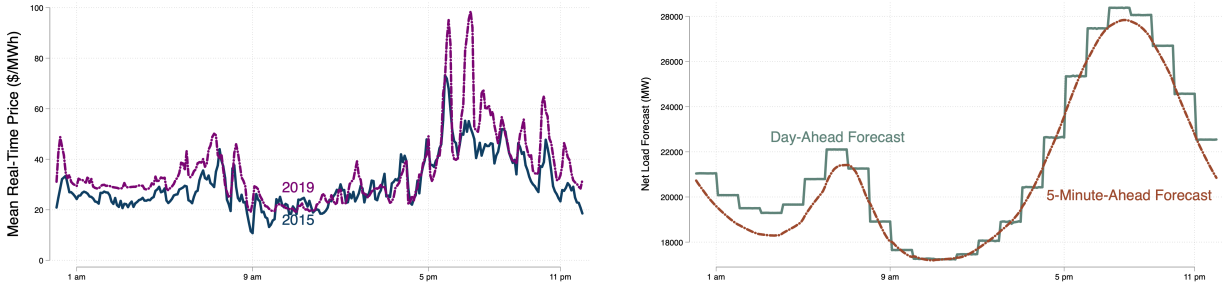
Notes: Each graphic plots the monthly average of a single variable over the sample period. The solar generation measure does not include distributed generation. The reported market prices are for the CAISO South Zone Trading Hub (SP 15). All data collected from CAISO.

unprecedented expansion of solar production has reduced California's carbon footprint, it also introduced new challenges in managing the electric grid. The "duck curve" in Figure 1c illustrates that as solar PV output continues to climb during daytime hours, other power plants must make larger and faster adjustments each evening when the sun sets. For example, on January 1st 2019, fossil fuel generators had to ramp up output by 15,000 megawatts over a three hour period. In the two bottom right panels of Figure 3, we see that as solar generation increased, mean prices in the real-time market have also trended upwards by nearly 20%. Real-time prices have also become more volatile, with the standard deviation of real-time prices increasing through most of the sample period.[15]

Figure 4a demonstrates how market price dynamics have changed as more solar PV has come online in California. In particular, each line shows the mean real-time market price at each five interval of the day for a given sample year. Throughout the sample, real-time prices were relatively low during the middle of the day when solar PV generation hits its daily peak. On the other hand, as total solar generation has risen each year, real-time prices have increased significantly in the evening hours when solar PV generation falls and fossil-fuel plants are called on to quickly ramp up. Table 5 in the appendix shows that for the 6 pm hour, prices rose by

[15] All prices are for the California South Hub Trading Zone (SP15).

Figure 4: Real Time Market Prices and Demand Forecasts



(a) Real Time Market Prices Patterns

(b) Day-Ahead and Real-Time Demand Forecasts

Notes: Panel a shows the average real-time market price (South Hub - SP-15) for each 5-minute interval of the day separately for the years of 2015 and 2019. Panel b plots CAISO's mean load forecast in the day-ahead market and the mean load forecast 5-minutes ahead of operation. In the day-ahead of operation, CAISO publishes a forecast for each hour. CAISO publishes a forecast at the 5-minute frequency on the day of operation.

16% between 2015 and 2019. Bushnell and Novan (2018) and show that wholesale price increases in the morning and evening can be explained by an increase in generation from flexible, high marginal cost natural gas turbines during those hours.[16]

Over the sample period, prices within each hour of the day have also become more sporadic. For instance, mean prices at 5:05 pm barely changed between 2015 and 2019, but prices at 5:55 pm increased by more than 50%. More generally, systematic differences between beginning-of-hour prices and end-of-hour prices have become more pronounced. The increased within-hour price variation was also driven by the new production profile imposed by the "duck curve". The right side of Figure 4 provides intuition about why the "duck curve" could increase within-hour price variation. In particular, Figure 4b shows the average day-ahead forecast of net load (demand minus solar & wind) as well as the five-minute ahead net-load forecast across the day. In the day-ahead market, forecasts are made hourly and CAISO schedules power plants in hour-long blocks to meet the anticipated demand. However, if the next day arrives and net load exceeds the day-ahead forecast value, new generators can be called upon in the real-time market to meet unplanned demand increments or unexpected reductions in solar output. We see in Figure 4b that during the evening ramp up period, the day-ahead forecast is systematically too low at the beginning of each hour and too high at the end of each hour. As a result, during early evening hours generators must reduce output relative to the day-ahead schedule at the beginning of each hour, and they must increase output relative to the day-ahead schedule at the end of each hour. These adjustments lead to larger within-hour price spreads that we observe in Figure 4a. This systematic imbalance in the real-time market has become further exacerbated over time as solar PV's generation share has expanded and faster fossil plant ramp ups are needed each evening.

Higher and more volatile market prices in California are a signal of the strains of operating a grid with significant presence of intermittent renewables. These challenges will persist as Cali-

[16]Early evening prices have risen because there is only a limited amount of plants that are physically capable of ramping up production in very short time frames, furthermore, ramping power plants quickly can impose start-up costs and additional maintenance costs.

fornia and other grids work to reach ambitious renewable energy targets. A key policy question is to understand how to achieve renewable energy targets while avoiding extreme price increases for rate-payers and maintaining a reliable grid that is not prone to regular blackouts. One potential solution is to utilize battery storage to help smooth out adjustment costs and mitigate the risks of price spikes and blackouts.

Battery Operations in the Electricity Market

Realizing the challenges imposed by renewable energy requirements, policymakers have passed new legislation aimed to increase electricity storage investment. In early 2018, the Federal Energy Regulatory Commission (FERC) issued Order 841, which requires independent system operators to remove any existing barriers that would inhibit participation of storage resources in wholesale markets. In addition, the California Public Utility Commission mandated that the state's investor-owned utilities must procure 1,300 megawatts (power capacity) of storage by 2020, with installations required to be operational no later than the end of 2024. Following the announcement of the mandate, California's utilities have rapidly increased investment in storage. By 2019, utilities had enough battery storage to discharge up to 126 MW at a given instant.[17]

The CAISO has made efforts to integrate these new storage technologies into their wholesale markets and ensure compliance with FERC Order 841. CAISO allows batteries to participate in both ancillary service and energy markets as non-generator resources (NGRs).[18] CAISO allows batteries to submit either demand bids or supply bids in both day-ahead and real-time energy auctions. A battery can submit a set of prices and associated quantities at which the battery is willing to discharge energy, with the possibility of submitting negative quantities to indicate battery would like to charge. Batteries also provide the ISO with information about the device's physical constraints such as energy capacity (how much energy the battery can hold in stock), and power capacity (how much the device can charge/discharge at any moment). Under a simple bidding strategy, a battery could submit a single price at which they would be willing to discharge (sell energy), and a lower price at which they would charge (buy energy). Apart from the energy market, batteries also have the option to supply reserve capacity (ancillary services).

We model a battery operating in the real-time energy market. By doing so, we allow storage to respond to price signals both within and across hours. Within-hour adjustments are likely to be an important source of value for storage resources on a grid with high renewable energy penetration. The main data input for our battery operations model is a multi-year time series of real-time prices. In our analysis, we use prices from the CAISO South Zone Trading Hub (SP15) because it currently contains the largest share of utility scale battery installations within CAISO.[19]

[17] 126 MW was the maximum aggregate output by batteries reported by CAISO between May 2018 and December 2019.

[18] An NGR is a resource that a can both inject and withdraw energy from the grid and change back and forth quickly between withdrawal and injection without bearing a start-up cost.

[19] The trading hub prices are calculated by CAISO as a load-weighted average of all the locational marginal prices in the region.

A limitation of our approach is that we do not allow our battery to provide reserve services. Our choice not to model reserve operations warrants some discussion considering that in 2019, more than 75% of battery capacity in CAISO was scheduled to provide reserve services not energy (see Figure 11 in the Appendix). Although many of the earliest battery operators are providing reserve services, it unlikely that reserve services will be the primary battery storage application in the long-run. Sackler (2019) emphasized this point as follows:

> *"Battery storage investors, however, should be wary of building investment cases primarily based upon future ancillary service market value expectations. ...while frequency regulation has historically been one of the most lucrative ancillary services markets (and exemplifying a service that [batteries] can outcompete traditional providers in), most ISOs only require 100-400 MW of the product in any given hour."*

To support this point, Figure 10 in the Appendix shows that CAISO procured less than 800 MW of average hourly regulation reserves[20] in all but five months of our five-year sample. Furthermore, demand for regulation services in CAISO did not grow substantially over time. For this reason, we focus our analysis on battery operations in the energy market to understand long-run market impacts of large storage investments.

3 Model

Our dynamic equilibrium model of battery storage includes two components. The first component, the *capacity adoption model*, solves agents' decisions of whether to make a capital investment in storage capacity in a given year. The second component, the *operations model*, solves agents' short-run dynamic decisions regarding when to charge and discharge their storage systems. More specifically, we allow agents with storage capacity to choose charge and discharge quantities at every five-minute interval throughout each year. We use the outputs from the *operations model* to microfound the payoffs for the *capacity adoption model*. This section first outlines the capacity adoption model and then the operations model. In the following section, we discuss the implementation and estimation of the model including our time series model of electricity supply.

3.1 Capacity adoption model

Our capacity adoption model considers an infinite mass of ex-ante identical potential battery operators, or agents for short. Each agent i has the ability to install a unit capacity of storage technology, $k = 1$, at one point in time. The unit capacity is sufficiently small that each agent takes electricity market prices as given.

[20] This calculation includes both regulation up and regulation down services, which are the two most lucrative reserve products offered by CAISO.

Agents face an infinite horizon dynamic problem with uncertainty and discount the future with factor β. Each year, agents that have not yet adopted storage observe the current state and decide whether to adopt storage, or wait and preserve the option to adopt storage in the future. Adopters bear a fixed cost of obtaining storage but can then use the storage to earn future flow profits, by acting as arbitrageurs in the energy market. Each agent has room for exactly one storage system and cannot replace the system once installed. Hence, agents solve an optimal stopping problem of when to invest.

Agent Decision Problem

We now discuss the decision process at each year y. At this point, agents that have not previously adopted make a binary decision of whether or not to invest in storage capacity. Agents that adopt must pay a fixed cost, c_y, that is the cost net of any subsidy available at year y. At year y, agents observe c_y but do not know future adoption costs. We assume that these costs evolve stochastically, declining over time in expectation due to technological advances. It is common knowledge that agents have rational expectations over future adoption costs and hence form accurate distributions over future trajectories. Accordingly, a benefit of waiting to invest is that capital costs are likely to be lower in the future. As above, adoption costs depend on subsidies that the government offers for battery investments in year y. We consider subsidy paths that evolve deterministically and are known to market participants.

Besides costs, agents must also forecast the expected current and future revenues from their system. The annual per-unit revenues depend on both the aggregate capacity of storage present in the market and the year. Aggregate storage capacity matters because, with additional capacity, storage owners will arbitrage away more of the intertemporal price differentials, reducing the per-unit payoff of storage. The year matters both because the expansion of renewable energy generation over time will likely increase the value of storage by increasing intertemporal price fluctuations and also because the year may affect available subsidies. We model renewable energy market share as evolving exogenously and deterministically over time. We motivate this modeling choice by the policy environment in California, where the state has specified a renewable portfolio standard that rises before reaching a 60% renewable requirement by 2030. Consequently we capture the impact of renewable energy share on the storage payoff through the calendar year, y.[21]

Define K to be the aggregate capacity present in the market at the start of a period. The aggregate state space can then be written as (c, y, K). The individual state includes the aggregate state plus the agent's battery capacity k, which starts out at $k = 1$ upon installation. Define $K^*(c, y, K)$ to be the equilibrium aggregate storage capacity following adoption at state (c, y, K); K^* includes the existing capacity K plus the capacity from the new adopters. Given the rational

[21] We assume the renewable generation share follows the state RPS schedule—and interpolate the share for years without a target—until reaching 60% in 2030. After 2030, we assume that renewable generation share stays constant at 60%. We also assume that this increase is divided between solar and wind proportionally according to the relative increase in capacity that we observe between the two technologies during our sample.

expectations assumption, agents can accurately predict $K^*(c, y, K)$ conditional on the state.

Owners of storage capacity buy and sell energy to maximize expected discounted profits in every small time interval of each year. Define the expected revenues per unit of capacity from storage at year y to be $\pi(y, K^*)$. We approximate $\pi(y, K^*)$ as a log-linear function of the solar generation share and equilibrium battery capacity K^*. For a series of values of y and K^*, we microfound $\pi(y, K^*)$ with our operations model that we discuss in Section 3.2. Namely, using the implied storage values from the operations model, we regress the equilibrium revenues across months and K^* on the above characteristics and controls, and use the fitted values in the capacity adoption model.[22] We allow for capacity investments to depreciate over time from usage at a rate $\delta(K^*)$.[23] Thus, a battery owner that installs a battery system at state (c, y, K) will have $\delta(K^*(c, y, K))$ of capacity at year $y + 1$.

We now formally exposit the agent's decision problem as a Bellman equation:

$$
\begin{aligned}
&V(k, c, y, K) = \\
&\mathbb{1}\{k = 0\}\Bigg[\max\Bigg\{\overbrace{\pi(y, K^*) - c + \beta \int V\Big(\delta(K^*), c', y+1, \delta(K^*)K^*\Big)dG(c'|c, y)}^{\text{Value from adopting}}, \\
&\overbrace{\beta \int V\Big(0, c', y+1, \delta(K^*)K^*\Big)dG(c'|c, y)}^{\text{Value from waiting}}\Bigg\}\Bigg] \qquad (1) \\
&+\mathbb{1}\{k > 0\}\Bigg[\underbrace{\pi(y, K^*)k + \beta \int V\Big(\delta(K^*)k, c', y+1, \delta(K^*)K^*\Big)dG(c'|c, y)}_{\text{Value if invested before } t}\Bigg],
\end{aligned}
$$

where $dG(c'|c, y)$ is the integral over the conditional density of the next period's costs net of subsidies given the year and current period's costs, and where K^* abbreviates $K^*(c, y, K)$.

In (1), an agent that has not already adopted can invest (the second line) or wait (the third line). These agents face an important trade-off in their capacity investment problem. On the one hand, agents that do not invest maintain the option to invest in future years when capital costs will likely be lower (though subsidies may also have expired). On the other hand, agents that wait and do not invest forgo $\pi(y, K^*)$. Finally, agents that invested before y (the fourth line) face no further choices but see their capacity depreciate over time.

Equilibrium of Model

A market equilibrium consists of values of $K^*(c, y, K)$ for all values of the aggregate state such that no potential adopters want to deviate from their strategy given this equilibrium capacity level. The equilibrium condition specifies that potential entrants must be indifferent between

[22] This type of approximation is consistent with a long literature in industrial organization and macroeconomics (e.g. Gowrisankaran and Rysman, 2012; Hendel and Nevo, 2006).

[23] Implicitly, $\delta(K^*)$ and $\pi(y, K^*)$ incorporate the fact that agents will modulate usage of their battery to lower depreciation, and that usage may be lower with a higher K. We discuss this point further in Section 3.2.

adopting and not adopting for all states with positive investment:

$$\overbrace{\pi(y, K^*) - c + \beta \int V\Big(\delta\left(K^*(c, y, K)\right), c', y+1, K^*(c, y, K)\delta\left(K^*(c, y, K)\right) dG(c'|c, y)}^{\text{Value from adopting}} \quad (2)$$

$$= \underbrace{\beta \int V\Big(0, c', y+1, K^*(c, y, K)\delta\left(K^*(c, y, K)\right)\Big) dG(c'|c, y)}_{\text{Value from waiting}}, \quad \forall c, y, K \text{ s.t. } K^*(c, y, K) > K.$$

In addition, in equilibrium, for all states with zero investment, $K^*(c, y, K) = K$, the value from adopting (the left side of (2)) must be less than or equal to the value from waiting (the right side of 2).

Finally, we discuss computation of this model. We compute the social planner's problem, which will generate the same state-contingent adoption rates as the competitive dynamic equilibrium but is easier to compute since it does not require that an equilibrium condition analogous to (2) must be satisfied (Hopenhayn, 1992; Ljungqvist and Sargent, 2012; Lucas and Prescott, 1971; Prescott and Mehra, 1980). For each aggregate state (c, y, K), the planner chooses a non-negative quantity of capacity to add. The planner Bellman equation can be written:

$$\mathcal{V}(c, y, K) = \max_{K^* \geq K} \left\{ K^* \pi\left(y, K^*\right) - c\left(K^* - K\right) + \beta \int \mathcal{V}\Big(c', y+1, \delta\left(K^*\right) K^*\Big) dG(c'|c, y) \right\}, \quad (3)$$

where K^* again abbreviates $K^*(c, y, K)$. Comparing (3) to (1), the planner's faces incentives on its last unit of investment that are equivalent to the those of the agents in the competitive equilibrium.

We compute the planner solution by discretizing both c and K. We choose a range for these values that is sufficiently broad to avoid constraining the solution and a discrete grid that is sufficiently fine to approximate the optimal solution well.

3.2 Operations Model

Our capacity adoption model uses $\pi(y, K^*)$ as an input, for a range of values of y and K^*. Since our data contain essentially no variation in battery capacity, we calculate $\pi(y, K^*)$ by solving π for different values of K^* as counterfactuals from a dynamic competitive equilibrium operations model. Furthermore, we allow the operations model values to vary for each month of our sample to flexibly accommodate the changes in the value of storage arising from (exogenous) fluctuations in renewable penetration seasonally and year to year.

We now describe the modeling that we use to perform this calculation. Within each month, we model each day as homogeneous, for computational tractability,[24] and K^* as fixed. Battery

[24]We assume the days within the month are the same only in our estimation of the optimal policies. When we estimate the value of storage, we apply our policies to the actual time series of net load, and market prices that we observe over our sample. On this dimension, we are likely understating the value of storage to the extent that there is some additional gains that could be obtained by more finely tuning our policy functions.

operators, or agents, buy and sell energy in the real-time electricity market in every five-minute time interval with the goal of maximizing their expected discounted profits from being arbitrageurs. We model agents as solving an infinite-horizon dynamic problem, where the ex-ante equivalent days are repeated in perpetuity. We believe that this is a reasonable approximation because we focus on battery technology with relatively short storage duration—batteries that can completely fill or empty within a few hours—so expectations about prices several days in the future will have relatively little influence on charging decisions today.

Each agent's charge decision at each time interval is a function of its charge level, current market characteristics, as well as the market characteristics last period. Since each agent controls a small capacity of storage, it takes real-time electricity prices as given. Although each agent acts as a price-taker, we allow the aggregate battery charge quantity to impact equilibrium market prices.

We operationalize this by estimating a marginal cost curve of electricity produced by fossil fuel by generators. Any battery charging will increase the quantity of electricity produced by dispatchable generators and thereby increase marginal costs and prices, while discharging will do the opposite. Analogously to the adoption model, each agent forms rational expectations about the evolution of net load and of aggregate quantity supplied by storage owners, which allows it to form rational expectations about the equilibrium price distribution in future periods.

Storage Technology

Our modeling approach captures three critical properties about battery storage technology. First, a battery's power capacity F determines what fraction of the battery can be charged or discharged in each five-minute interval and therefore how *quickly* the battery can transition from full to empty and vice versa. We model battery technology that can fully discharge within a four-hour period. We focus on four-hour duration batteries because they are the most common battery specification currently operating in the CAISO market. For these batteries, $F = \frac{1}{4\times 12}$.

Second, we model the round-trip efficiency of the battery, υ, which determines the percentage of energy that is lost during a charge/discharge cycle from a given level of energy to a higher level and back to the original level. We focus on a round-trip efficiency of $\sqrt{\upsilon} = 0.85$ because the storage capital costs projections from NREL are based on lithium-ion batteries with these parameters.

Finally, we model capacity fading, which occurs when the amount of energy a battery can hold decreases with repeated use. Lithium-ion batteries, as well as most other batteries, will exhibit capacity fading. The extent of capacity fading depends on the number of round-trip cycles the battery has made, the calendar age of the battery, and the characteristics of each cycle.[25] Capacity fading is likely to play an economically important role in a storage operator's decisions. Specifically, a storage operator may not want to engage in arbitrage if the expected difference

[25]For instance, deeper cycling (such as going from 0% to 100% and back to 0% one time) can cause more capacity degradation than shallow cycling (such as going from 40% to 60% and back to 40% five times).

in prices across the charge/discharge periods is not substantially large to justify the additional capacity fading that the battery will endure to complete a cycle.

We model capacity fading using the capacity degradation model of Xu et al. (2016). This model takes each a series of charge/discharge decisions over a time interval, calculates cycle ends and the characteristics of each cycle, and returns a value for how much capacity degradation would occur over this time interval. The appendix provides details on our application of this model. It would be computationally difficult to model agent's optimization over the amount of capacity fading in the operations model. Instead, as we detail in Section 4, we develop a heuristic optimization process to account for capacity depreciation in our model. Our heuristic decreases the agent's perceived round-trip efficiency υ to account for capacity fading. In so doing, we account for capacity fading in the agent's charge decision, though using a functional form that is an approximation to the true functional form as represented by Xu et al. (2016).

Agent Decision Problem

At each five-minute time interval s, each agent seeks to maximize the sum of expected discounted profits, making charge/discharge decisions over the interval.[26] We let S denote the number of time intervals within a day and D the number of days within a year. The per-period discount factor is then $\beta^{\frac{1}{SD}}$.

The agent bases its dynamic operations decisions on a state that is composed of two components. First, the state includes the fraction of the battery's capacity that it has stored as energy. Denote the battery's fraction of charge as $f \in [0, 1]$. Second, the state includes the current market price P_u and the agent's perceived distribution of future market prices at time interval u, which we denote $G^u(P_{u+1}, P_{u+2}, \ldots)$. Because of our perfect competition assumption, the agent only cares about market prices and not the actions of other battery operators, though other agents' actions will affect prices in equilibrium.

We can express the operations Bellman equation as:

$$\begin{aligned} W(f, P_u, G^u(P_{u+1}, P_{u+2}, \ldots)) = \max_q \Big\{ & P_u \times (\mathbb{1}\{q > 0\} q\upsilon + \mathbb{1}\{q < 0\} q/\upsilon) + \\ & + \beta^{\frac{1}{SD}} \int W(f - q, P_{u+1}, G^{u+1}(P_{u+2}, P_{u+3}, \ldots)) d(G^{u+1} | G^u) \Big\}, \\ & \text{s.t. } -F\upsilon \leq q \leq F/\upsilon \text{ and } 0 \leq f - q \leq 1, \end{aligned} \tag{4}$$

where q is the fraction of the battery's capacity that it discharges.[27] Equation (4) states that the agent maximizes its total current profits from charging, which are equal to price times quantity supplied (the first line), plus the expected future value of the position next period, where

[26] CAISO has both a fifteen-minute ahead market (cleared 4 times per hour) and a five-minute ahead market (cleared 12 times per hour). We focus on the five-minute ahead market.

[27] For simplicity, equation (4) defines G^{u+1} to be a function of G^u only, rather than other information available at time u.

the agent's energy held next period is $f - q$ (the second line). This maximization is subject to the battery's power capacity of F and the constraint that it cannot be less than empty or more than full (the third line). The quantity supplied in the first line reflects that some energy is lost through charging and discharging. We assume that these energy losses occur symmetrically across charging and discharging.[28]

Equilibrium of Model

We focus on a symmetric equilibrium, where all battery operators start each period with the same fraction already charged and then choose the same charge/discharge fraction each period. Define $Q(q)$ to be the quantity of electricity supplied to the grid by battery operators at a period where this (common) discharge fraction is given by q:

$$Q(q) = K^* \times \Big(\mathbb{1}\{q > 0\} q \upsilon + \mathbb{1}\{q < 0\} q / \upsilon \Big).$$

To model equilibrium, we need to consider how $Q(q)$ affects equilibrium prices and in turn affects charge/discharge decisions. We do this by developing a time series model of electricity demand and supply.

Our model of demand is relatively straightforward, but justified by the features of the electricity markets. We model the (net) demand for electricity, or *net load*—the electricity demanded by final users net of the amount produced by intermittent renewable sources—as an autoregressive process whose mean depends on the time of day. In the absence of storage, and because electricity supplied must always meet electricity demanded for technologicial reasons, *net load* is the amount of electricity that needs to be supplied by dispatchable generators (which are all generators except intermittent renewable sources). In a world with a fleet of batteries, the amount of electricity that needs to be supplied by dispatchable generators, Z_t, at any time interval is then net load, X_t, minus $Q(q)$, or $Z_t \equiv X_t - Q_t(q)$. It is helpful to note now, that without batteries ($Q_t(q) \equiv 0$), and the amount of electricity that needs to be supplied by dispatchable generators is exactly net load, $Z_t = X_t$.

We assume the price of electricity is a function of the amount supplied by dispatchable generators, the time of day, and an additional shock, ε^P. This additional shock represents other factors that determine the price of electricity conditional on the amount of electricity supplied, factors like weather, generator outages, and transmission congestion. By our perfect competition assumption, wholesale electricity price at interval of day s is equal to marginal cost, yielding:

$$P(s, Z_t, \varepsilon^P, Z_{t-1}) = MC(s, Z_t, \varepsilon^P, Z_{t-1}) \tag{5}$$

where $MC(s, Z_t, \varepsilon^P, Z_{t-1})$ is the marginal cost function, Z_t is the amount of electricity supplied by dispatchable generators in time period t, while Z_{t-1} is the amount of electricity supplied by

[28] If we instead assumed that energy was only lost while charging, the last term of the first line of Equation 4 would be $\mathbb{1}\{q > 0\} q + \mathbb{1}\{q < 0\} q / \sqrt{\upsilon}$).

dispatchable generators in time period $t-1$. By including the amount of electricity supplied last period, we are able to accommodate the possibility of ramping/adjustment costs on the part of electricity generators.

Additionally, we assume that net load can be modeled as the sum of two components: (i) its expected value in any interval of day, X_s, and (ii) a deviation or shock from this expected value ε^L, or $X_t = X_s^L + \varepsilon^L$. We assume both of the error terms ε^P and ε^L are observable to agents at the start of each short interval and have a distribution $d\varepsilon(\cdot,\cdot)$, which governs the likelihood of future values–which we assume is also known to the agents.

Define $f^a \in [0,1]$ to be the fraction of aggregate storage capacity that agents are holding at the start of any period. Any agent can determine the future equilibrium charge/discharge decisions and hence the future price distribution $G(\cdot)$ from f^a, the interval of day s, the amount of electricity supplied by dispatchable generators last period Z_{t-1}, and the additional shocks ε^P and ε^L. Since f^a is the same as f by the symmetry assumption, these five elements indicate the state for an agent at any time interval.

Defining $q^*(f^a, s, \varepsilon^P, \varepsilon^L, Z_{t-1})$ to be the equilibrium quantity discharged at that state, we can rewrite the Bellman equation imposing the equilibrium conditions as:

$$\begin{aligned} W(f^a, s, \varepsilon^P, \varepsilon^L, Z_{t-1}) = \\ \max_q \Big\{ P(s, X_s^L - Q(q^*(f^a, s, \varepsilon^P, \varepsilon^L, Z_{t-1})) + \varepsilon^L, \varepsilon^P, Z_{t-1}) \times (\mathbb{1}\{q>0\} q\upsilon + \mathbb{1}\{q<0\} q/\upsilon) + \\ + \beta^{\frac{1}{SD}} \int W(f-q, s+1-\mathbb{1}\{s=S\}S, \varepsilon^{P\prime}, \varepsilon^{L\prime}, Z_t) d\varepsilon(\varepsilon^{P\prime}, \varepsilon^{L\prime} | \varepsilon^P, \varepsilon^L) \Big\}, \\ \text{s.t. } -F\upsilon \le q \le F/\upsilon \text{ and } 0 \le f - q \le 1. \end{aligned} \tag{6}$$

In a competitive dynamic equilibrium, $q^*(f^a, s, \varepsilon^L, \varepsilon^P, Z_{t-1})$ must be the value of q that maximizes (6) for every state $(f^a, s, \varepsilon^L, \varepsilon^P, Z_{t-1})$.

Analogously to our approach for solving the equilibrium in the investment stage, we recast the battery operations problem as a social planner's problem. We rewrite the problem as the single-agent planner's problem whose allocation is then equivalent to the competitive equilibrium problem. The objective of the social planner is to maximize welfare. Electricity demand is perfectly inelastic in the short-run, therefore the planner will meet demand by choosing the state-contingent battery discharge fraction $q^*(f^a, s, \varepsilon^P, \varepsilon^L, Z_{t-1})$ that minimizes the total expected discounted cost of electricity production. Let $TC(q, s, \varepsilon^P, \varepsilon^L, Z_{t-1})$ denote the total cost of production for any state. It is the integral of marginal cost from zero to net load plus aggregate battery output:

$$TC(q, s, \varepsilon^P, \varepsilon^L, Z_{t-1}) = \int_0^{X_s^L - Q(q) + \varepsilon^L} P(s, \zeta, \varepsilon^P, Z_{t-1}) d\zeta. \tag{7}$$

where we can integrate the price function because of our perfect competition assumption that prices will be equal to marginal cost. Although $TC(\cdot)$ includes an unknown constant, we normalize the constant to zero because the optimal solution to the planner's problem is invariant to

the value of the constant.

We then write the social planner's Bellman equation as:

$$\mathcal{W}(f^a, s, \varepsilon^P, \varepsilon^L, Z_{t-1}) = \max_q \Big\{ -TC(q, s, \varepsilon^P, \varepsilon^L, Z_{t-1}) + \\ +\beta^{\frac{1}{SD}} \int \mathcal{W}(f^a - q, s+1-\mathbb{1}\{s=S\}S, \varepsilon^{P\prime}, \varepsilon^{L\prime}, Z_t) dG(\varepsilon^{P\prime}, \varepsilon^{L\prime} | \varepsilon^P, \varepsilon^L) \Big\}, \\ \text{s.t. } -Fv \leq q \leq F/v \text{ and } 0 \leq f^a + q \leq 1. \tag{8}$$

For a similar model to ours, Cullen and Reynolds (2017) prove that a solution to the planner's problem exists, that the solution and associated state-contingent prices are equivalent to all competitive equilibria, and that a competitive equilibrium exists.

We also use the same framework to compute the optimal operation and the associated value of a monopolist storage operator relative to a competitive battery market. The monopolist's operation problem is also a single agent problem so computation is similar to the competitive problem, the only difference is that the flow payoff for a monopolist is equal to $P(s, X_s^L - Q(q) + \varepsilon^L, \varepsilon^P, Z_{t-1}) \times Q(q)$ instead of $-TC(q, s, \varepsilon^P, \varepsilon^L, Z_{t-1})$. In words, the monopolist values its total revenues as opposed to (the negative of) total cost.

4 Estimation and Implementation

To implement the operations model, we begin by developing a time series model that captures how net load (demand - solar - wind) and the marginal cost curve (supply) evolve. We construct our time series model to achieve several objectives. First, the model should produce credible and well-calibrated forecasts of how both net load and the marginal cost curve evolves at the 5-minute frequency over the day. Such a model enables us to forecast the future evolution of equilibrium prices in the data and also allows us to calculate counterfactual equilibrium prices after a large battery fleet enters the market. Second, the model forecasts should be feasible from the informational perspective of an actual battery operating in the real time market. Third, the model should be able to accommodate the presence of adjustment costs that influence which generator sets the marginal price at each point in time. Lastly, the set of variables that enter our time series model must remain parsimonious enough so that the Bellman equation (Equation 8) remains computationally feasible.

4.1 Time Series Model of Electricity Net Load and Marginal Cost

Given the objectives and constraints outlined above, our model includes the following state variables: (1) time-period-of-day (at the 5-minute frequency), (2) a price residual that influences net load conditional on the other state variables (which jointly determine the current value of net load), (3) a price residual that influences prices conditional on the other state variables, and (4) lagged output from dispatchable generators. This model specification results in very high di-

mensional state space. Consequently, it is not computationally feasible to add additional state variables such as outdoor air temperature or information about the full set of active generators at each moment in time.[29]

We also leverage an important institutional detail of the California electricity market in that it has both a day-ahead market in addition to the real-time (spot) market. For planning and to ensure reliability the day-ahead market serves as the primary scheduling market. While the real-time market serves as the market to settle any last minute adjustments that are required due to unforeseen circumstances. Importantly, each day the system operator produces a projection of net load and prices for each hour of the day in the day-ahead market, which are publicly available to market participants and would be information a potential battery operator could utilize to choose its charge/discharge decisions for the upcoming day. A key feature of our time series model is that we employ information from both the day-ahead market and the real-time market. In particular, we use forecasts of (hourly) net load and prices from the day-ahead market in conjunction with the realized values of net load and prices (at the 5-minute frequency) in the real-time market. By leveraging the informational content from both markets, our model remains feasible to implement in real-time from the perspective of a market participant. In other words, our model could be estimated and employed by an actual battery operator to inform bids in the real-time market. Moreover, the information from both markets allow us to construct a flexible model of the marginal cost curve that can vary over time due to seasonal fluctuations in hydroelectric availability, unexpected generator outages, and changes in other market conditions (e.g., the natural gas price).

Modeling Net Load

We now briefly describe the time series model of net load (demand). The process for net load (X_t^L) at any five-minute period, t, is given by the following equation:

$$\begin{aligned} X_t^L &= E_0\left[X_t^L\right] + \left[X_t^L - E_0\left[X_t^L\right]\right] \\ &= \tau_{s(t)}^L + \varepsilon_t^L \end{aligned} \tag{9}$$

where $E_0[\cdot]$ is the expectation taken at time "zero", or more specifically the net load forecast published by the system operator in the day-ahead market, and $\tau_{s(t)}^L$ captures the interval-of-day specific expectation of net load. This formulation for the process maps directly to the framework provided in the previous section with $X_s^L \equiv \tau_{s(t)}^L$.

To capture the serial correlation in the deviations of the realizations of net load from what is anticipated in the day-ahead market we model the ε_t^L as an AR(1) process given by:

[29]Despite the system operator (CAISO) having information on the current set of generators that are currently available, it is unlikely that an individual battery operator would have access to this information. Thus, the restricted nature of dimensions of our state space is likely to be more consistent with the information set of a battery operator in this market.

$$\varepsilon_t^L = \rho^L \varepsilon_{t-1}^L + \eta_t^L, \quad \eta_t^L \sim N(0, \sigma_L^2) \tag{10}$$

where η_t^L is a mean zero serially uncorrelated shock with variance σ_L^2, and ρ^L governs the persistence of deviations of net load from their forecasted value in the day-ahead market.

A couple of features of the net load model are worth emphasizing. First, the process for net load does not depend on price. Thus, we are assuming the demand (or, net load; which includes generation from solar and wind) for electricity is perfectly price inelastic. While in many other markets and in other segments of this market (e.g., retail) this assumption is false, for the wholesale electricity markets this assumption is justified because end-consumers do not face wholesale price changes, and the generation from renewables is also insensitive to price.

Second, the model of net load flexibly controls for (predictable) systematic fluctuations within the day at the 5-minute frequency (through $\tau_{s(t)}$). Additionally, it can accommodate virtually any lower frequency patterns of seasonal fluctuations that might be present in net load–to the extent that they are predicted by market participants. This is important, as it is well documented, that the (net) demand for electricity is affected by many seasonal factors including weekday effects, and holidays, to seasonal fluctuations in weather, to longer run trends in the penetration of renewable generation. The last of these, of course, is exactly the variation we hope to leverage to estimate the complementarity between solar generation and batteries.

Third, it also captures the serial correlation that is likely to arise in any interim periods (i.e., several adjacent five minute periods) over a forecast horizon (i.e., made a day before). These deviations are likely to arise from the short term deviations in weather patterns. In making its projection of future prices, the battery operator will combine the future trajectory of the predictable component of net load, as well as the current *net load deviation* from what was projected for that period, in forming expectations over the future net load trajectory.

Finally, the model also leads to an error term for net load (ε_t^L) that could be interpreted as "structural"–or, invariant to the (counterfactual) presence of a large fleet of batteries. More specially, it is reasonable to assume that the market participants forecast of the final demand for electricity (and, generation from renewables), and ultimately how reflective that forecast was in what ultimately happened is unlikely to be affected by whether or not batteries were operating in the market.

Under this formulation of the net load process, we require a full set of interval-of-day forecasts (τ_s^L) to compute the optimal policy functions (for each month), and the parameters governing the serial correlation of the deviations from these forecasts (ρ^L, σ_L^2). For the former, we use a temporally disaggregated version of the day-ahead forecasts for net load from the the first day of the month.[30] For the parameters governing the serial correlation of the net load deviations, we

[30] A notable issue with the day-ahead market forecasts of net load and prices is that they are constructed only at the hourly frequency, while the real-time market (and our operations model) is formulated at the five minute frequency. We use the Kalman filter/smoother approach outlined Proietti (2006) to temporally disaggregate the day-ahead forecasts of net load to create a five minute frequency version. For more details, see On-line Appendix.

estimate an AR(1) model using OLS of the deviations of net load from the day-ahead forecasts (also temporally disaggregated) over the year 2015, and hold these parameters fixed over the remaining part of our sample (2016–2019) which is used to compute the value of batteries.[31] In both cases, the approach ensures that the model would be feasible to estimate given the information set of a market participant (i.e., only data from the past is used).

Modeling the Marginal Cost Curve

Next, we describe the time series model of the electricity marginal cost curve. Importantly, it is not sufficient for our model to accurately forecast the historical time series of real-time electricity prices. The model must also credibly predict the prices that would occur in a counterfactual environment with a large fleet of batteries reshuffling the amount of electricity generated at each point in the day, while still using only the data that would be available to a battery operator in real time.

We assume that the real-time market prices, P_t, in period t on day d are determined by the following marginal cost curve for electricity:

$$\begin{aligned} P_t = MC_t &= \delta_d + \theta_d[\mathcal{K}_t - Z_t]^{-\psi_d} \\ \mathcal{K}_t &= e^{\varepsilon_t^P} \kappa_d^{\alpha_d} Z_{t-1}^{1-\alpha_d} \end{aligned} \tag{11}$$

($\mathcal{K}_t$: Available Generation Capacity at t)

recall that Z_t is the quantity produced by dispatchable generators in period t, Z_{t-1} is the quantity produced by dispatchable generators in period $t-1$, and ε_t^P is a mean zero error term. The supply curve is governed by five day-specific parameters: $\delta_d, \theta_d, \kappa_d, \alpha_d, \psi_d$.

We adapt the functional form for supply from a previous literature on commodity storage (Pirrong, 2012). This functional form has several appealing properties for our application. Importantly, the supply function is monotonically increasing in the Z_t, which is critical for solving the operations model.[32] Additionally, the marginal cost function is parsimonious, yet flexible enough to capture the highly-convex (hockey-stick) shape of the electricity marginal cost curve. The supply function achieves convexity though the $\mathcal{K}_t$ term, which represents the generation capacity that is available to produce at time t.[33] As Z_t approaches the capacity constraint $\mathcal{K}_t$, marginal cost increases rapidly, which allows the function to capture high price spikes that occur frequently in the real-time market. Figure 5a illustrates an example of the marginal cost curve during a summer afternoon in 2016 when net load was approaching the constraint on available

[31] We also tried estimating the ρ^L, σ_L^2 parameters separately for each sample year and found that these parameters are relatively constant over time.

[32] The supply curve will always be increasing in Z_t as long as $\psi > 0$ and $\theta > 0$. In contrast, a cubic functional form could capture highly convex costs but may not be monotonic.

[33] Available capacity includes generators that are currently online or those that can quickly become operational without any lead time to start up (e.g., a gas peaker plant).

generating capacity.

We specify the available generation capacity $\mathcal{K}_t$ to be a function of generation last period Z_{t-1}, a mean-zero idiosyncratic shock ε_t^P, parameter κ_d, and parameter α_d. The inclusion of Z_{t-1} in $\mathcal{K}_t$ allows for the possibility of adjustment constraints. Some generating units, such as nuclear plants or combined-cycle gas plants, require an extended period of time to transition from a non-operational state. For this reason, dispatchable generation last period can influence the available capacity in the current period. The κ_d parameter measures the amount of generation capacity that is available at all times throughout the day such as a gas peaker plants that can quickly turn on. The α parameter governs the relative importance of Z_{t-1} versus κ_d in determining the total generation capacity at time t. For $\alpha = 1$, marginal cost is static and does not depend on dispatchable generation output in previous periods. On the other hand, if $\alpha \in (0, 1)$, then an increase in generation last period will reduce marginal cost in the current period.

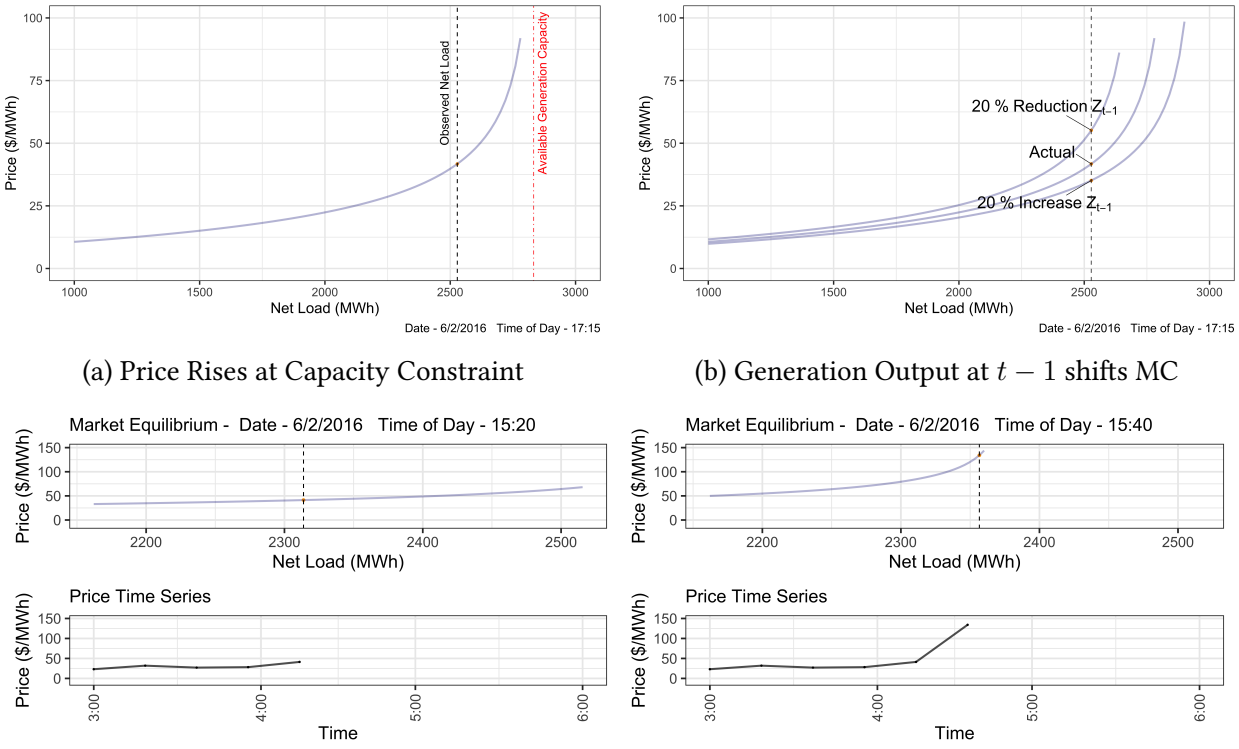
Figure 5b shows an example of how a shift in Z_{t-1} can affect marginal cost; we see that when net load is approaching the available capacity constraint, a shift in Z_{t-1} can lead to a substantial price increase. Finally, available capacity $\mathcal{K}_t$ also depends on a time-varying term ε_t^P which represents random shocks to available generation capacity such as unplanned generator outages or a transmission congestion event, especially those that are likely to be unanticipated at very high frequencies (e.g., at the 5 minute frequency).

The interpretation of the remaining parameters of this supply curve are also transparent. The δ parameter controls the intercept of the marginal cost curve, and is permitted to be negative. The supply curve parameters θ and ψ govern the slope of the supply curve.

Notably, we allow the supply curve parameters to vary day to day (hence, the d subscripts). In effect, these parameters–and possible movements in them–represent the agglomeration of shifts in natural gas prices, changes in the availability of low cost generation coming from nuclear power plants and hydroelectric sources, as well as temporary generator outages and changes in electricity imports and exports from neighboring states. We avoid adding additional time-varying market characteristics (e.g., fuel prices) directly into the model because doing so would render the dynamic operations model computationally infeasible. However, by allowing the supply curve parameters to vary day-to-day, we still capture a substantial share of cost heterogeneity due to changing market conditions.

There are a few implicit assumptions embedded in our marginal cost formulation that warrant further discussion. First, the inclusion of Z_{t-1} allows for some degree of dynamic adjustment costs, however, our specification is unlikely to capture the full suite of adjustment costs present in the market. Our cost function will account for an increase in cost that arises when the system operator calls on a high-cost generator to produce because a low-cost generator needs more time to ramp up. However, our cost function may not incorporate a fixed cost that a non-marginal generator pays to start up or shut down in a given period. Second, the error term ε_t^P enters marginal cost in a non-linear fashion and contributes to the available capacity $\mathcal{K}_t$. We treat this shock as "structural", that is, we assume ε_t^P is invariant to counterfactual changes in battery operations. To defend this assumption, we appeal to the fact that the model captures the non-

Figure 5: Time-Varying Marginal Cost Curve



(a) Price Rises at Capacity Constraint

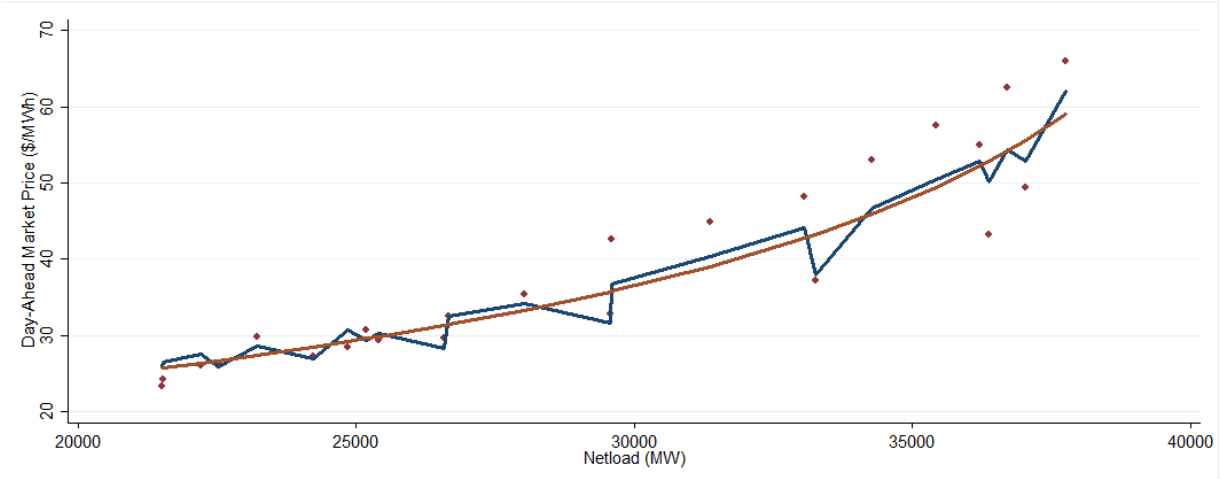
(b) Generation Output at $t-1$ shifts MC

(c) Equilibrium Before Price Spike Event

(d) Equilibrium During Price Spike Event

Notes: Figure 5a shows the market equilibrium and the implied generation capacity available for a single five-minute interval. Figure 5b shows how (20%) changes in last period's dispatchable generation shift the marginal cost curve. Figures 5c and 5d show how both the net load and the marginal cost curve shifts during a period when price increased rapidly over a 20-minute span. Net load is measured as the the energy delivered in MWh over the five-minute period.

Figure 6: Marginal Cost Curve From Day-Ahead Market



Notes: This figure displays the scatter plot of the day-ahead market prices and net load for each hour for a day in June 2016. Additionally, the marginal cost curves estimated for that day's data with adjustment costs (blue line) and without adjustment costs (orange line) are also displayed. The reported market prices are for the CAISO South Zone Trading Hub (SP 15). All data collected from CAISO.

linear relationship between prices and net load in this market, in addition to a large degree of marginal cost heterogeneity over time, and at least partially captures the (known) adjustment costs the face generators.

We recover the structural error ε_t^P as the shock to available capacity $\mathcal{K}_t$ required to rationalize the real-time price observed at time t, conditional on current net load, lagged net load, and that day's supply curve parameters. In particular, at each point in time we can invert the supply curve equation to recover the structural error term using observed data on price and net load:

$$e^{\varepsilon_t^P} = \frac{\left[Z_t + \left[\frac{P_t - \delta_d}{\theta_d}\right]^{-1/\psi}\right]}{\kappa_d^{\alpha_d} Z_{t-1}^{1-\alpha_d}}. \tag{12}$$

Figures 5c and 5d illustrate how our model rationalizes a rapid change in price that occurs in the real-time market, while figure 6 illustrates the fit of the supply curve and the role of accommodating adjustment costs. At 3:20 PM on June 6, 2016 the real time market price was just under \$50/MWh, then at 3:40 PM price nearly tripled to \$140/MWh. We can see in the figures that our model rationalizes this price change through both an increases in net load and an inward shift of the supply curve. In particular, we can perfectly rationalize this price change via a negative shock to the price residual ε_t^P, which we interpret as an unexpected shock to available generation capacity. Due to the high-frequency (5-minute) nature of our application, we allow for the possibility of serial correlation in ε_t^P. In particular, we assume this process is a first order auto-regressive process given by the following equation:

$$\varepsilon_t^P = \rho^P \varepsilon_{t-1}^P + \sigma_t^P \eta_t^P$$

where ρ^P controls the degree of persistence of the price deviations from the supply curve, and η_t^P is a mean zero serially uncorrelated shock with unit variance, and σ_t^P accomodates for possible heteroskedasticity that might exist for different hours of the day (e.g., the evening high ramp up hours of the day).

This formulation of the real-time market prices maps directly to the framework provided in the previous section with a *total cost* function given by the following equation:

$$\begin{aligned} TC(q, s, \varepsilon_t^P, \varepsilon_t^L, Z_{t-1}) = {} & \delta_d Z_t(q, s, \varepsilon_t^P, \varepsilon_t^L) - \frac{\theta_d \left[e^{\varepsilon_t^P} \kappa_d^{\alpha_d} Z_{t-1}^{1-\alpha_d} - Z_t(q, s, \varepsilon_t^P, \varepsilon_t^L)\right]^{1-\psi_d}}{1-\psi_d} \\ & + \frac{\theta_d \left[e^{\varepsilon_t^P} \kappa_d^{\alpha_d} Z_{t-1}^{1-\alpha_d}\right]^{1-\psi_d}}{1-\psi_d} \end{aligned} \tag{13}$$

where the final term in constructed to normalize shut-down costs to equal zero, a normalization

that we maintain throughout, as the overall level of total cost is not identified.

Having derived a parametric form for the total cost function, we still need to estimate the the supply curve parameters $\delta_d, \theta_d, \kappa_d, \alpha_d, \psi_d$, as well as estimate the parameters governing the serial correlation of the price shocks (ρ^P, σ_t^P). Our approach to estimation is motivated by three goals: (1) the supply parameters should be credibly identified off variation in net load, (2) estimation should be feasible in a real-time forecasting environment,[34] and (3) estimation should be computationally feasible considering our multi-year dataset.

Towards those ends, we estimate the (daily) time series of $\delta_d, \theta_d, \kappa_d, \alpha_d, \psi_d$ by non-linear least squares separately for each day using the *day-ahead* market prices on the *day-ahead* forecast of net load (and, lag of net load). More specifically, we find the set of supply curve parameters that minimize the implied sum of squared residuals, or:

$$\min_{\delta_d,\theta_d,\kappa_d,\alpha_d,\psi_d} \sum_{t \in D(d)} \left[\tilde{P}_t - \left(\delta_d + \theta_d \left[\kappa_d^{\alpha_d} \tilde{Z}_{t-1}^{1-\alpha_d} - \tilde{Z}_t \right]^{-\psi_d} \right) \right]^2$$

where $\tilde{P}_t$ is the day-ahead market price in period t, $\tilde{Z}_t$ is the day-ahead forecast of net load for time period t, and $D(d)$ comprises all the time periods belonging to day d. Given the already described hourly frequency of the day-ahead market in the California electricity market, these (daily) non-linear least squares regressions constitute hourly frequency time series regressions (with 24 observations each).[35] Conditional on the full time series of supply curve parameters, we can use the inversion given by equation (12) on the real-time market prices and the realized value of net load to recover the supply shock, ε_t^P.

This approach achieves each of our three objectives. First, the variation in net load that is used to estimate the supply curve parameters primarily comes from plausibly exogenous within-day variation in the demand of electricity as well as weather conditions that determine generation from solar and wind. Therefore, our identifying assumption is that within-day variation in solar generation and electricity demand is uncorrelated with the supply-side factors that will contribute to idiosyncratic deviations in the market clearing price.

Second, because the supply curve parameters are estimated using the *day-ahead* market prices and the *day-ahead forecast* of net load, these supply curve parameters are available to all market participants when they have to make their decisions for the real-time market. This feature of the estimation approach is critical, as it ensures that our estimate of the value of storage uses a model of the evolution of prices that would have been feasible in real-time. Finally, this approach towards estimation is computationally cheap enough to be implemented even in our high frequency and multiple year setting. In estimating the optimal policy, we use the supply

[34] The estimates should only use historical data that would have been available and feasible to implement by an actual market participant.

[35] For this reason, the lagged value of net load in the non-linear least squares step constitutes the lag of net load *last hour*.

curve parameters estimated from the first day of that month.[36] In our assessments of the value of storage, we use the full time series of estimates of $\delta_d, \theta_d, \kappa_d, \alpha_d, \psi_d$, and the implied structural shocks from those estimates, ε_t^P.

The remaining parameters required to estimate the battery's charge/discharge policy are parameters governing the serial correlation and possible heteroskedasticity of the price residual (ρ^P, σ_t^P). Here as well, it is critical that the parameters that are used to estimate the optimal policy functions could have been estimated (forecasted) in real-time.

For the parameters governing serial correlation and possible heteroskedasticity of the price residual (ρ^P, σ_t^P), we estimate an AR(1) model using OLS of the estimated structural shocks, ε_t^P, implied by the structural supply curve parameters estimated from from the day-ahead market prices and forecasts of net load. To capture the possibility of heteroskedasticity, separate estimates of the variation of the resulting residuals from the above mentioned OLS regression are estimated for the late afternoon hours (5pm-9pm) and remaining hours, respectively.[37] To ensure feasibility of the policies being applied to market outcomes from 2016–2019, we estimate these parameters using only 2015 data.

4.2 Computation of the Operations Model

With parameter estimates for the net load and the marginal cost curve in hand, we can now solve the operations model. We discretize the problem in order to be able to compute it. In our base computation, we discretize ε^P, ε^L, Z_{t-1}, and the charge state f^a into 10 dimensions each. As was described above, our model for prices is a function of the 5-minute interval of a day, implying that $S = 288$. Also, our sample consists of 48 months and we allow the operations policies to vary by month and across 5 candidate values of battery capacity K^* ranging from 1,000 to 25,000 MWh. These dimensions result in an overall size of the state space that is $10 \times 10 \times 10 \times 10 \times 288 \times 48 = 69,120,000$ states. We solve the optimization independently for each month-K^* pair which results in 240 dynamic problems with $288,000$ states each.

Probably the most common computational method for single agent dynamic problems is Bellman recursion, which involves recursing (8) until approximate convergence. Many single agent problems with this order of magnitude of states are computable using Bellman recursion. However, Bellman recursion is a contraction mapping with modulus of the discount factor. An important complication in our model is that we are modeling the decision process at the 5 minute interval. Under a conventional annual discount factor of $\beta = 0.95$, $\beta^{\frac{1}{DI}} \approx 0.99999951$. Thus, while a Bellman recursion in our model would be a contraction mapping, the modulus would be

[36] This approach still ensures that the policy estimated for each month uses only information that would have been available to the battery operators at that current period. To the extent that trends in demand and supply conditions lead to extensive movements in the supply curve parameters $(\delta_d, \theta_d, \kappa_d, \alpha_d, \psi_d)$, it is possible that our policy function is sub-optimal as it is estimated off a "stale" estimate of the supply curve. In this regard, one can view our estimates of the value of storage as a conservative estimate that could be increased to the extent that the battery operator could re-estimate its policy function at a finer frequency than once a month.

[37] Due to the kurtosis exhibited in the real-time market prices, both variance estimates are constructed using the median absolute deviation and scaling it by 1.4826.

very close to 1, making computational time infeasibly long.

A popular alternative to Bellman recursion, is policy function iteration. In our context, this method implements the following recursive algorithm. First, it defines an initial policy function vector, $\vec{P}^0$, which is a value of c for every state. This function implies a vector of static payoffs for each state, $\vec{\pi}^0$. Together with the (exogenously determined) transitions for s and $\varepsilon^P, \varepsilon^L$, the policy function defines an initial transition matrix as $\mathbf{Q}_0$. A typical element, $\mathbf{q}_{ij}$ of transition matrix $\mathbf{Q}_0$ indicates the probability of going to state j from state i. Using this notation, we can express the vector of values for every state as:

$$\vec{\mathcal{W}}_0 = \beta^{\frac{1}{DS}} \mathbf{Q}_0' \vec{\mathcal{W}}_0 + \vec{\pi}_0 \Rightarrow \vec{\mathcal{W}}_0 = (I - \beta^{\frac{1}{DS}} \mathbf{Q}_0')^{-1} \vec{\pi}_0. \tag{14}$$

Thus, (14) provides a solution to $\vec{\mathcal{W}}_0$. Policy iteration then proceeds by solving $\vec{P}_1$ as the optimal policies given $\vec{\mathcal{W}}_0$, then solving $\vec{\pi}_1$, $\mathbf{Q}_1$, and $\vec{\mathcal{W}}_1$ analogously, and then repeating this entire process until approximate convergence. A central problem with policy iteration in our context is that a matrix inverse with as many states as we have is computationally infeasible, due to computational time, storage space, and numerical precision.

We develop a variant of policy iteration that exploits the fact that our transition matrix is sparse, since s advances deterministically, and that the state space repeats every S periods. In particular, we modify the step that solves for $\vec{\mathcal{W}}$ in (14) and instead solve for $\vec{\mathcal{W}}$ by performing policy iteration S (instead of 1) periods ahead.

Specifically, define $\vec{\mathcal{W}}_s$ to be the subvector of $\vec{\mathcal{W}}$ for all states at interval s. Define also $\mathbf{Q}_s$ to be the submatrix of transitions from interval s to the following interval (either $s+1$ or 1). Finally, define

$$\vec{\Pi}_1 = \vec{\pi}_1 + \beta^{\frac{1}{DS}} \mathbf{Q}_1' \vec{\pi}_2 + \ldots + \beta^{\frac{S-1}{DS}} \mathbf{Q}_{S-1}' \cdots \mathbf{Q}_1' \vec{\pi}_S$$

Then,

$$\vec{\mathcal{W}}_1 = \vec{\Pi}_1 + \beta^{\frac{1}{D}} \mathbf{Q}_I' \cdots \mathbf{Q}_1' \vec{\mathcal{W}}_1 \Rightarrow \vec{\mathcal{W}}_1 = (I - \beta^{\frac{1}{D}} \mathbf{Q}_I' \cdots \mathbf{Q}_1')^{-1} \Pi_1. \tag{15}$$

Equation (15) allows us to express $\vec{\mathcal{W}}_1$ with a matrix inverse that is of dimension 10,000 in our baseline setting. Using $\vec{\mathcal{W}}_1$, we then solve $\vec{\mathcal{W}}_S, \vec{\mathcal{W}_{S-1}}, \ldots, \vec{\mathcal{W}}_2$ quickly, with one-step backward recursion This results in a computational process where the computation time is linear in the number of intervals.

4.3 Evaluating Storage Value and Capacity Fading Using Policies

In principle, we could use the value functions implied by the optimal policies $\vec{\mathcal{W}}$ to measure the expected gross value of batteries at varying levels of $K^\star$, and to measure how the gross value of batteries evolved across months. A concern with this approach, however, is that our discretization of the price process, as well as the supply curve estimates used to estimate the policies are at best approximations–and, lack capturing the volatility of demand and supply conditions that occur even *within the month* for this market. Recall, that we only solve the Bellman equation

conditional on the estimated supply curve parameters on the the first day of each month. Another concern with this approach would be that it would fail to account for the amount of battery capacity that effectively "fades" away through use.

We overcome both of these issues through a set of heuristics. To ensure that we capture the role of the volatility in demand and supply conditions in the market we compute the gross value of batteries from the observed time series of prices and net load as opposed to the Bellman equations assessment of the expected value. More explicitly, we continually update (at the daily frequency) the supply curve parameters $(\delta_d, \theta_d, \kappa_d, \alpha_d, \psi_d)$ throughout our estimation sample. By continually updating the supply curve parameters, we better capture the likely equilibrium effects on price that would have occurred with a significant battery presence.[38] The full time series of the supply curve parameters together with the day-ahead forecasts of net load and the day-ahead market prices lead to full time series of implied net load $(\hat{\varepsilon}_t^L)$ and price residuals $(\hat{\varepsilon}_t^P)$. Thus, to compute the gross value of batteries we apply our estimated policy functions to the full time series of price and net load residuals, as well as the continually updated time series of the supply curve. This approach is computationally less demanding, as we do not re-compute what the optimal policy decision for every updated value of the supply curve parameters. We view this approach as conservative its assessment of the gross value of batteries to the extent the a more finely updated policy function could achieve higher values.

In order to account for capacity fading we leverage data from the year (2015) prior to our sample period (2016–2019) as training sample. Using the pre-period training data, we solve the model using a sparse grid of different candidate round-trip efficiency levels $\sqrt{\upsilon}$. The idea is that a battery operator may be able to improve their payoff by acting as if the battery has a lower-than-actual round-trip efficiency. A battery with lower round-trip efficiency will be more reluctant to arbitrage prices unless the payoff is sufficiently high. Batteries that arbitrage less will also endure less capacity fading and thus could potentially earn a higher payoff, after considering the costs of capacity fading.

After solving for optimal policies for each of the candidate round-trip efficiency values, we simulate the evolution of the state-of-charge, f, using the actual demand and price data from the training data set for each of the candidate policies (i.e., for each candidate round-trip efficiency). We then feed the simulate series into a capacity degradation model. The capacity degradation model is based off a rainflow cylce counting algorithm developed by Xu et al. (2016), which we explain in more detail in the appendix.

The above procedure provides two important outputs that we use for the operations model. First, it provides an estimate of how much capacity degradation will occur on average in each month. We use this value to adjust the discount factor β in equation 4 to account for degradation over time due to standard operations. Second, we compare realize payoffs from storage for each of the candidate round-trip efficiency values after accounting for capacity degradation and find the round-trip efficiency value that leads to the highest payoff. We then use the best perform-

[38] Table 6 provides summary statistics of the daily supply curve parameters over varying sub-samples of our analysis sample.

ing round-trip efficiency value as $\sqrt{v}$ when we solve the model on the full data sample.[39] The outcome of this procedure is that we account for capacity degradation without adding a state variable to the operator problem. Our approach is both computational feasible and also ensures that we do not overstate the value of storage by ignoring capacity fading. The approach also allow storage operations to acknowledge the costs of capacity fading using a "rule-of-thumb" when making dispatch decisions.

4.4 Linking the Operations Model and the Adoption Model

Having outlined the approach towards estimating and computing the operations model, we now briefly summarize the steps we take to utilize our operation model results as inputs into the battery adoption model.

Using our approach to account for capacity fading we simulate the policy functions on the realized sequence of supply curve parameters, price residuals, and load residuals. We simulate separate policies for different candidate values K^* and across each sample month. That is, we solve for how different capacity battery fleets would be dispatched and the corresponding equilibrium effects on prices for each month. Therefore, we can also calculate the discounted sum of electricity cost reductions—the value of storage—for each candidate K^* across our sample months.

Estimating the storage value for varying sized battery fleets is critical to determining optimal adoption in storage capacity. Recall that the capacity adoption decision depends on two factors: (1) c the capital cost of storage, and (2) $\pi(y, K^*)$, the expected revenues per unit of capacity. Thus, we use the simulated battery operations to determine $\pi(y, K^*)$.

Solving the storage operations problem for each value of K^* informs how per-unit expected revenues π change as more storage capacity is added to the grid. Consequently, we rely on our battery operations model to identify how changes in K^* affect changes in π. In contrast, we utilize variation in renewable energy capacity over time to estimate how changes in y affect expected revenues, π. Recall that expected battery revenue can change over time, y, due to changes in renewable energy generation capacity. The enormous expansion of intermittent renewable generation over sample, namely, solar PV, has led to unprecedented changes in the dispatch patterns of fossil fuel generators and therefore the marginal cost of electricity. Thus, by solving the operations model for each month, we can evaluate how expected revenues, π change over time as the renewable portfolio standard has ramped up.

Due to the computational burden of the solving the operations model, it isn't feasible to solve the operations model for all possible battery capacity levels K^*. Instead, we solve the operations model for each month of our 48-month sample and across a grid of five candidate values of K^* ranging from 1000 MWh to 25,000 MWh. Thus, we evaluate the per-unit value of storage, $\pi(y, K^*)$ at $5 \times 48 = 240$ points. We then obtain an approximation $\pi(y, K^*)$ by estimating the

[39] We also allow the optimal round-trip efficiency value to depend on the capacity of the battery fleet. Therefore we repeat this procedure separately for each level of aggregate capacity of the battery fleet.

following regression:

$$\pi_{mk} = \gamma_0 + \gamma_1 log(K^*_{mk}) + \gamma_2 RenewableShare_{mk} + \gamma_3 X_{mk} + \nu_t + \varepsilon_{mk} \tag{16}$$

The outcome of variable, π_{mk}, is the present value of battery capacity (per unit) in month m with aggregate battery capacity k.[40] The key explanatory variables are the natural log of aggregate battery capacity and the proportion of total electricity generation that came from renewable energy sources in month t. The same set of candidate of battery capacity values are used to solve the model in every distinct month of the sample, so $log(K^*_{mk})$ will not be correlated with any omitted variables that may affect per-unit storage value. On the other hand, the renewable generation share fluctuates seasonally and is increasing systematically over time. As a consequence, the renewable share variable could be correlated with other factors that are changing over time that are also correlated with the value of storage. To address these endogeneity concerns, we include month-of-year fixed effects, ν_m, in the regression to control for seasonal factors that could be correlated with renewable generation. We also control for the average price of natural gas in month m, the average electricity load (demand) in month m, and the amount of hydroelectric resources available in month m.

Evolution of Battery Capital Costs

The operations model provides the information needed to calculate the benefit of a storage investment, thus, the final input we require to solve the adoption model is a specification for the evolution of battery capital costs. We specify that the dynamic process for the cost of the storage technology, c_y, is given by the following unit root with drift process:

$$c_y = c_{y-1} e^{\tau} e^{\xi_y}, \quad \xi_y \sim N(0, \sigma_c^2) \tag{17}$$

with c_{2018} as the (initial) cost assessment of battery technology in the year 2018, and with τ and σ_c governing the size of the drift and future uncertainty involving the level of costs. To the extent that $\tau < 0$, the costs of storage will generally trend down over time. The uncertainty involved around the size of these future declines in costs is captured by the process ξ_y. We assume that ξ_y are i.i.d. across time.

The motivation for these modeling choices are encapsulated in the cost projections collated by the National Renewable Energy Laboratory (NREL) plotted in Figure 2. These future cost assessments display several patterns critical to our modeling choices including: (i) a non-linear trajectory, (ii) a downward trend in costs, (iii) uncertainty at future horizons, (iv) positive skewness in the distribution of future costs, and (v) a general increase in the uncertainty of the cost assessments throughout the over 30 years of the projection.

Each of these patterns guide the modeling approach for the dynamic process of costs. The (almost) certain downward trend in costs motivates the drift term in our model, the non-linear tra-

[40] More concretely, π_{mk} is the lifetime value of a 1 MWh battery operating in a scenario where month m grid conditions occur in perpetuity.

jectory motivates the exponential formulation. Additionally, the increasing level of uncertainty in the forecast uncertainty motivates the unit-root (in logarithms) formulation of the model, and the positive skewness in the cost assessments motivates the log-normal distribution for the shock process.

Given our cost model, we have two parameters to estimate: the magnitude of the downward drift (τ), and the size of the shock process governing the level of cost uncertainty (σ_c). Importantly, we do not observe actual realizations of the battery capital cost process, but we do observe the set of *projected* cost realizations from Cole and Frazier (2019). Therefore, we treat each cost projection (i.e., each line in Figure 2) as a realization of the cost process to estimate the parameters of the capital cost model. In particular, we use a method of moments approach to recover τ and σ_c using the normalized cost projection data reported in Cole and Frazier (2019). We derive the moment conditions for estimation in the Appendix. Table 1 shows that $\hat{\tau} = -0.044$ and $\hat{\sigma_c} = 0.064$. Following Cole and Frazier (2019) we also assume the the initial condition for capital costs in 2018 is $c_{2018} = \$380/kWh$.

Table 1: Capital Cost Parameters

	Estimate	SE
τ	-0.044	(0.001)
σ_c	0.064	(0.003)

5 Results

Having discussed our modeling approach, we now turn to highlighting the key results from both the operations and adoption models. We begin by discussing several findings from our operations model. In particular, we discuss temporal patterns in equilibrium battery output and how these battery outputs affect equilibrium market prices. We then report our estimated value of battery operations and show how the value of batteries changes with the share of renewable generation and the total amount of battery capacity in the market. Finally, we review our set of results pertaining to battery adoption over time. Namely, we show battery adoption trajectories under several counterfactual policy and competitive environments.

5.1 Operations Model Results

Figure 7a illustrates the mean simulated battery discharge quantity for each hour of the day between January 2016 and December 2019. Each line in the figure shows battery output for a specific aggregate battery fleet capacity K. Batteries discharge the most during the hours that net load is the highest—6-8 am in the morning as well as the evening peak hours of 5-8 pm. We also see that as aggregate battery capacity grows, aggregate batteries discharge increases in the evening, while aggregate charging increases during the day. In Figure 7c, we see how the batteries' storage inventory (i.e., state-of-charge) evolves throughout the day. On average, batteries

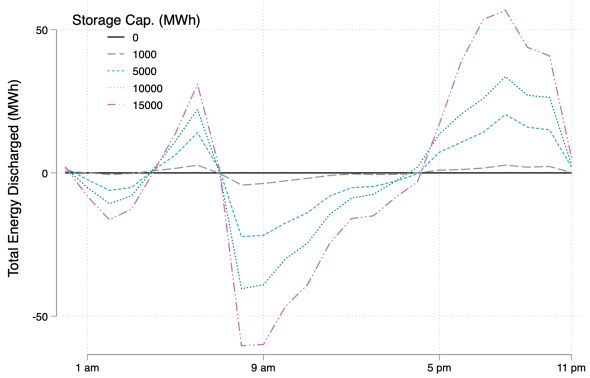
remain between 60-80% charged throughout the day, and mean state-of-charge increases with aggregate capacity. Batteries remain relatively full in order to have sufficient inventory in case of an unexpected price spike event such as a generator outage.

As the fleet expands, Figure 7b shows that batteries operations exert a strong effect on equilibrium prices. Figure 7d zooms in on the evening hours, we see that during the 6-7 pm hours—the hours with the highest average net load—a relatively small 1,000 MWh battery fleet would reduces average prices by over $10 per MWh. Figure 7b also illustrates that batteries' charging during the middle day have a relatively small affect on prices because marginal cost is relatively low and flat during those hours. Figure 7b also illustrates that the first few units of battery investment would have the largest impact on equilibrium price, whereas incremental storage investment has a smaller impact on prices. The first batteries that enter the market will reduce the occurrence of extreme pricing events by discharging during periods when net load approaches the available generation capacity. By doing so, the batteries will reduce prices and also move to the equilibrium to flatter regions of the marginal cost curve, thus reducing the marginal impact of subsequent battery entry on prices. Table 3 emphasizes this result, we find that the first 1,000 MWh of storage capacity would reduce load-weighted average price by almost 10% from $33.67 per MWh to $30.62 per MWh, on the other hand, the next 14,000 MWh would only reduce mean price by 3% to $28.91 per MWh.

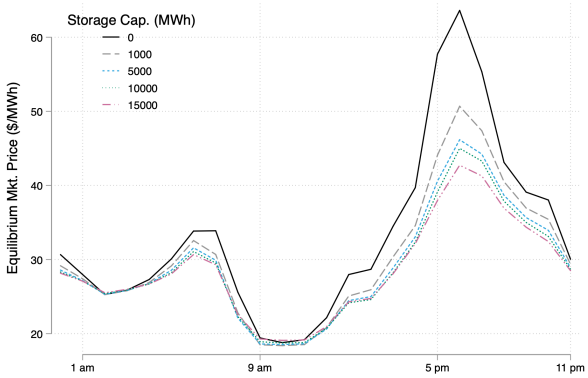
Welfare and Distributional Effects of Storage Operations

We next turn to investigating the distributional consequences of storage operations in the electricity market. We have already seen that storage operations would have substantial impacts on equilibrium prices. Consequently, storage will impact total welfare and may have important distributional effects. Recall that we assume that demand is perfectly inelastic and is held fixed under counterfactual changes in storage operations. Therefore, a change in welfare is equal to the change in the total cost of electricity generation. The first column of Table 2 shows the mean effect on battery operations of welfare in thousands of dollars per hour. We find that a 1,000 MWh storage fleet would increase welfare (reduce the total costs) by $1,190 per hour on average. A larger fleet with 15,000 MWh would further reduce costs by $10,910 per hour. The other columns of Table 2 report how battery operations would affect the economic rents earned by different types of market participants. Column 2 indicates that batteries would significantly reduce the total price (price$\times$load) that load serving entities need to pay to meet demand. In particular, a 15,000 MWh battery fleet would reduce mean hourly expenditures for utilities by over $140,000 per hour. Relatedly, batteries would substantially reduce the the revenues of both dispatchable generators and intermittent generators. More specifically, hourly revenue would fall by $125,000 for dispatchable generators (price$\times$ net load) and $16,000 for solar and wind generators. Perhaps surprisingly, intermittent renewable generators are harmed by battery operations because batteries–by reducing ramping constraints–reduce prices during the early afternoon (2-4pm) when many solar plants are producing.

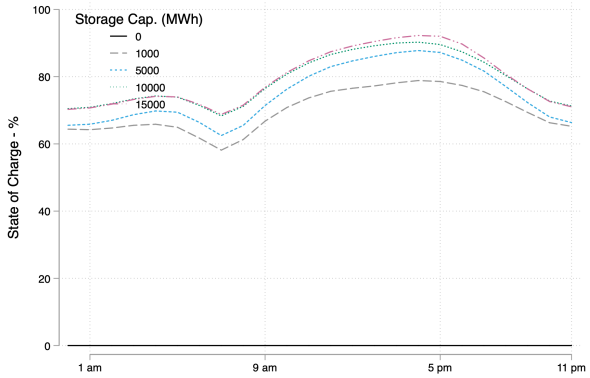
Figure 7: Battery Output and Equilibrium Prices Effects



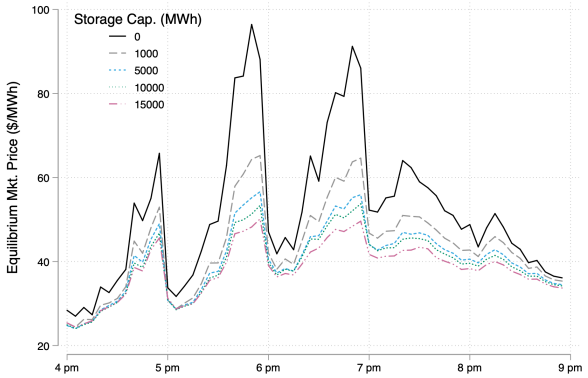
(a) Mean Hourly Battery Output Across Day



(b) Mean Hourly Equilibrium Prices



(c) Mean State-of-Charge



(d) Peak Equilibrium Prices (5-min. Freq.)

Notes: Each line plots the mean counterfactual outcome across 2016-2019.

Table 2: Mean Hourly Welfare and Revenues Across Aggregate Battery Capacity Levels

K	Welfare	$ to Serve Load	Dispatchable Gen. Rev.	Solar+Wind Rev.	Battery Rev.
0	0.00	885.22	769.07	116.10	0.00
1000	1.19	796.28	691.45	104.79	1.14
5000	5.03	765.29	663.75	101.51	4.32
10000	8.08	754.78	654.07	100.68	5.92
15000	10.91	744.82	644.43	100.36	6.98

Notes: All variables are hourly means in thousands of dollars. Welfare is the change in the mean hourly total cost of generation relative to the case where battery capacity (K=0). "$ to Serve Load" equals the the equilibrium price times the the total load in that hour (i.e., the cost to load serving entities). "Dispatchable Gen. Rev.", "Solar + Wind Rev.", "Battery Rev." are the mean hourly revenue for dispatchable generators, renewable generators, and battery operators respectively.

Monopoly Battery Operation

(In Progress)

Table 3: Equilibrium Prices and Aggregate Battery Capacity

K	Price (All hours)	Price (6-9 AM)	Price (10 AM - 3 PM)	Price (4-8 PM)
0	33.67	30.58	23.42	50.18
1000	30.62	28.63	21.87	40.86
5000	29.58	27.94	21.60	37.96
10000	29.22	27.71	21.58	36.96
15000	28.91	27.59	21.83	35.95

Notes: Prices reported are in $/MWh and are the load-weighted mean across all five minute intervals in our main sample covering 2016-2019.

Measuring Solar and Storage Complementary

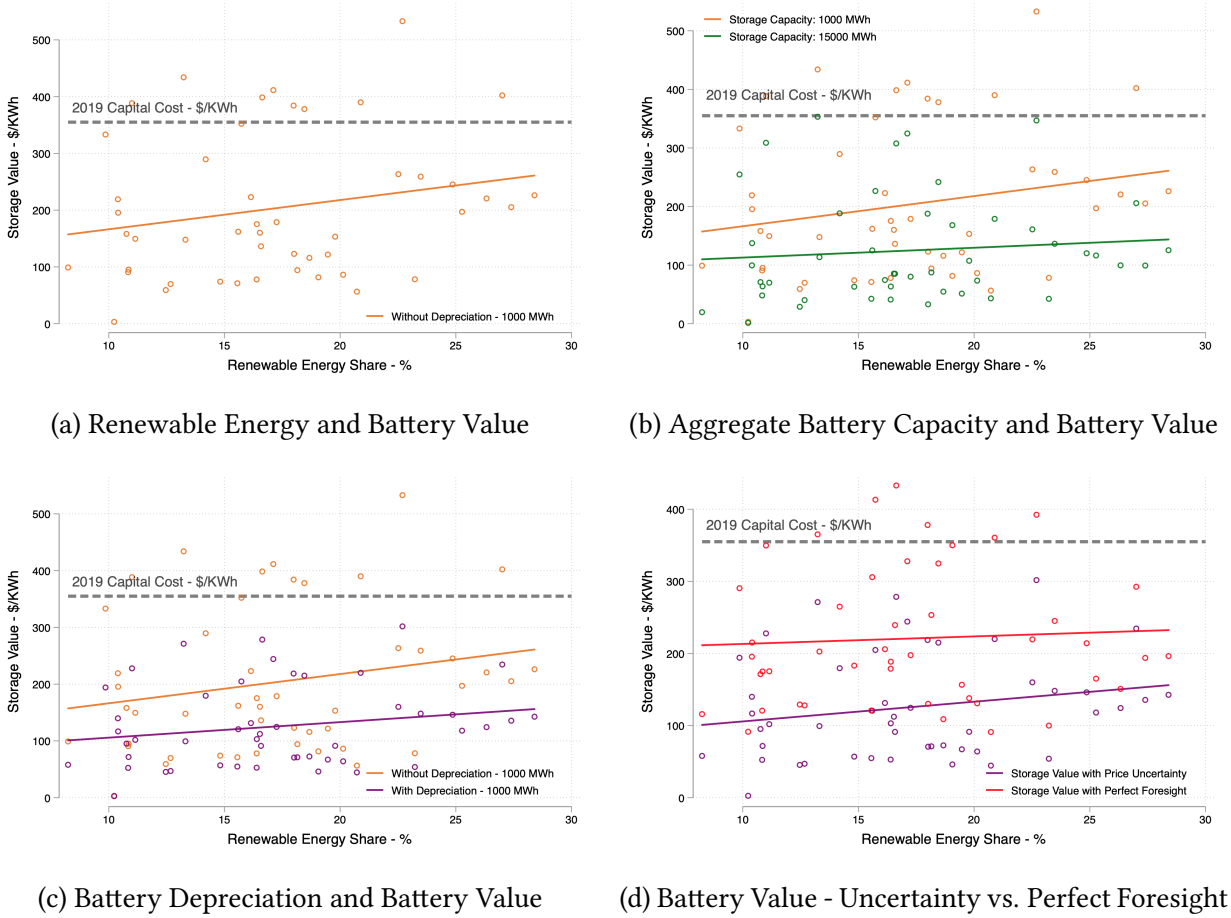
We next use the outputs from the operations model to calculate the discounted social value of battery storage. In reporting our results, we first discuss how the value of storage estimated over the 2015–2019 period relates to the share of generation coming from intermittent renewables—a set of results that speak directly to the complementarity of storage and solar generation.

Figure 8 provides an illustrative visualization of some key associations. First, Figure 8a plots the per-unit discounted present value of batteries for each month in our four-year sample, assuming a small battery fleet with capacity of 1,000 MWh. In calculating the battery value for each month, we assume that battery will operate in perpetuity and the market conditions that occurred in that month repeat forever. The orange line plots a simple linear fit of the relationship between storage value and the share of electricity generated with solar PV. The dashed-grey line shows estimated break-even capital cost for a 1 kWh of storage in 2019 from Cole and Frazier (2019). In this panel, the strong positive association between the prevalence of solar generation and the value of storage is clear. Additionally, by comparing the current assessment of the capital costs to the relationship between solar generation share and the value of storage–it is evident a modest increases in the share of solar would be associated with increases in the social value of storage that would cover current assessments of capital costs.

Figure 8b contrasts the average per-unit value of storage for a smaller battery fleet (1,000 MWh) and a larger battery fleet (15,000 MWh). The per-unit value of storage falls by 20-50% with 15,000 MWh of capacity, compared the per-unit value with 1,000 MWh capacity. This pattern highlights the decrease in returns-to-scale from storage investment. As more batteries enter the market, battery operations will eventually reduce prices in peak periods. By mitigating price fluctuations, batteries reduce the total cost of electricity generation but also reduces the value of subsequent battery investment. Figure 8b shows that this relationship is particularly pronounced at higher levels of renewable energy penetration.

In the lower panel, Figure demonstrates how battery depreciation (i.e., capacity fading) influences storage value. Depreciation from cycling reduces the implied value of a storage investment

Figure 8: Value of Battery Investment



(a) Renewable Energy and Battery Value

(b) Aggregate Battery Capacity and Battery Value

(c) Battery Depreciation and Battery Value

(d) Battery Value - Uncertainty vs. Perfect Foresight

Notes: Each point in the scatter plots (a-b) represents the implied value of storage for a single month during the sample (2016-2019). The solid lines in (a-b) plot the linear trend for each group.

by between 25% to 50% depending on the level of renewable energy generation. In particular, we find that batteries tend to cycle more as renewable generation increases and therefore depreciation will also increase. Due to the economically important affects of battery depreciation, we account for depreciation when calculating the payoff of a storage investment in our adoption model.

Finally, Figure 8d compares the value of storage under uncertainty about load and the electricity supply (the base case) relative to a scenario where battery operators could perfectly predict future realizations of load and the supply curve. We see that a battery operators could provide between 50% to 100% higher value under perfect information. This result underscores the volatility and unpredictable nature of real-time market prices. Importantly, our results that allow for uncertainty should be interpreted as a lower bound for storage value that could be further improved through better forecasting.

Linking the Operation Results to the Adoption Model

To provide a more systematic evaluation of the relationship between renewable energy penetration and per-unit storage value, as well as storage fleet size and per-unit storage value, we estimate regression Equation 16. Recall, that this regression equation forms the bridge between our results from the operations model to our model of battery capacity adoption. The dependent variable for these regressions is the present discounted value per MWh of storage capacity accounting for capacity fading from operations.

We report the regression results in Table 4. Our preferred specification (3) include controls for the mean load for the month, the mean natural gas price over the month, and the Sacramento Valley hydroelectric water year index (WYI) associated with that month. We show the results both with and without month-of-year fixed effects.

Table 4: Per-Unit Value of Storage as a Function of Storage Capacity and Renewable Penetration

	(1)	(2)	(3)
Log(K)	-28.761***	-28.761***	-28.761***
	(7.445)	(6.911)	(5.484)
Renewable Share [0,100]	2.956**	5.825***	8.567**
	(1.483)	(1.692)	(3.583)
Month of Year FE	N	N	Y
Load, NG Price, Hydro Controls	N	Y	Y
Observations	192	192	192

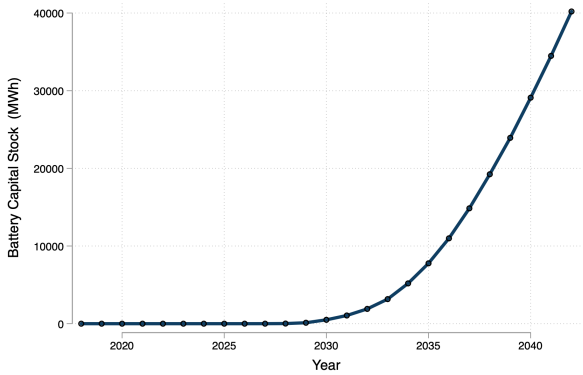
Notes: The dependent variable is the present discounted social value per kWh of storage capacity after accounting for capacity fading from operations. Each observation represents a single month of the sample for a single storage capacity (K). Specifications with controls include the mean load for the month, the mean natural gas price over the month, and the Sacramento Valley hydroelectric water year index (WYI) associated with that month.

Consistent with 8, the regression results paint a clear picture of a close connection between the growth in solar generation and the value of storage. By leveraging the extended time series variation during a stretch of considerable growth in solar generation, we estimate that each percentage point increase in renewable energy market share corresponds $8.56 per kWH increase in the value of a storage investment in a competitive storage market. The results also show that each 1% in battery capacity decreases the per-unit value of storage by $28.76 per KWh.

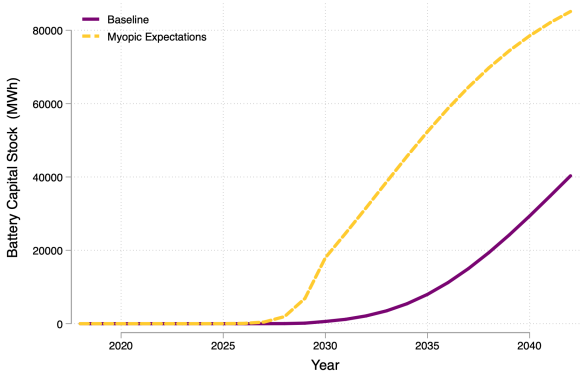
5.2 Battery Adoption Results

We next turn to discussing the results from our battery adoption model. We first use our adoption model to solve for the equilibrium adoption path under a competitive battery market and under a monopoly battery market. Recall that the agents' adoption decisions depend on the future

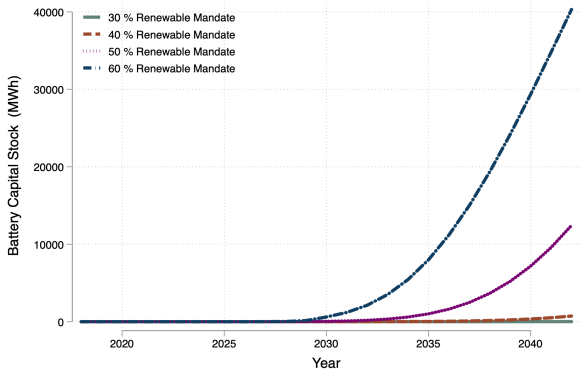
Figure 9: Battery Adoption Paths



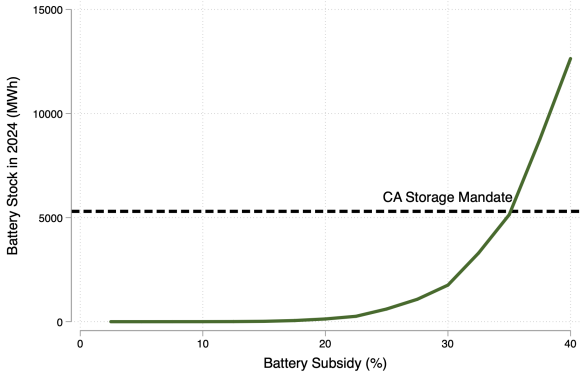
(a) Adoption without Subsidies



(b) Myopic vs. Forward-Looking Adoption



(c) Renewable Mandates and Battery Adoption



(d) Battery Adoption by 2024 with Subsidies

Notes: Figure (a) figure plots the competitive equilibrium adoption. The other figures plot battery capacity adoption under different counterfactuals.

trajectory of battery costs and the future path of battery values. Therefore, the adoption choice is a function of our estimates of the cost process (Equation 17) and predictions of future battery values from regression equation 16. In particular, we use regression equation 16 to predict the value of a battery investment as function of the renewable generation share and the amount of battery capacity in the market in each year. We take the level of renewable generation in each future year as exogenously determined by California's renewable portfolio standard (RPS). The California RPS requires an incrementally higher share of renewable generation each year before reaching 60% in 2030.[41] To predict future battery value for adoption model, we also assume that batteries have rational expectations about the path of battery investment by other market participants.

Figure 9a depicts the mean battery capacity over time without any subsidies or mandates.[42] We see that for the competitive battery market, battery capacity reaches about 40,000 MWh by

[41] We assume that renewable share remains at 60% in the years following 2030.

[42] The exact adoption path will depend on the realizations of the cost process, we therefore report the mean capacity level for each year implied by by taking the average outcome across many simulated realizations of the cost process.

2040, with almost all investment coming after 2030. There are two important reason why battery adoption is essentially non-existent before 2030. First, the value of a storage investment will be higher in the future as more renewable generation is added to the grid. Second, battery investors obtain a large option value in delaying investment until costs are lower in the future.

Figure 9b shows the difference in the competitive adoption path between the cases of forward-looking agents and myopic agents. Myopic agents adopt storage as soon as the current value of storage exceeds the capital cost, while myopic agents consider the potential benefit of waiting longer to invest after further capital cost declines. We see that myopic agents invest heavily in storage between the years of 2025 and 2030, reaching over 20,000 MWh by 2030. The large differences between the myopic and forward-looking investment paths demonstrates the strong option value of delaying investment due to future expectations regarding the rapidly falling capital cost of storage.

We next consider how the competitive adoption path would change in response to changes in renewable energy policy—the California RPS. Figure 9c shows the competitive market's investment in battery capacity for a 30% RPS by 2030, a 40% RPS by 2030, a 50% RPS by 2030, and a 60% RPS by 2030 (the current policy). The result indicate that almost no storage investment would occur under an RPS below 40%. With the more aggressive renewable energy mandates, storage investment greatly increases. The 50% RPS would result in 15,000 MWh of expected storage capacity by 2040, and the 60% RPS would result in 40,000 MWh. These results suggest that battery storage investments are not likely to be economically viable unless intermittent renewable penetration is relatively high.

In Figure 9d, we investigate how government subsidies for battery storage can steer the battery adoption path for both the competitive and monopoly cases. We find that moderate subsidies can speed the rate of adoption. In particular, we find that with a 35% subsidy for batteries, California could achieve 5,200 GWh (1300 GW) of batteries by 2024— the amount required under California's AB 2514 battery mandate.

Finally, we use the adoption model to evaluate the present discounted value of batteries investment in the California. Our welfare analysis indicates that storage investments will be a highly valuable tool for California to meet its ambitious renewable energy goals. In particular, we find the total expected discounted surplus from the battery market is $3.6 billion billion as of 2018. To put in other terms, the presence of battery storage increases the average value of renewable generation by $2.42/MWh.[43]

[43] In present value terms, the California is expected to generate 1.5 billion MWh from renewable energy sources.

6 Conclusion

A significant challenge to meeting the world's growing demand for energy is that utilities cannot typically store electricity for later use.[44] As the majority of renewable generation come from intermittent resources, the interest and potential role for battery storage technology has grown substantially. This paper develops a dynamic, competitive equilibrium model of battery adoption and operations. The model includes a number of key features that we believe are critical for understanding the battery adoption capacity and value created by batteries under different policies. This includes modeling the equilibrium effects of large-scale battery adoption smoothing out price peaks and valleys, ramping costs, depreciation from battery use, and uncertainty as faced by agents in the market. We evaluate the predictions of our model using data on the California electricity market. We develop frontier time-series forecasting algorithm to model future marginal cost curves for electricity.

Our results indicate that complementarity between solar and battery technology is economically large. We are also currently not very far from a point where a small amount of battery storage could break even. Utility-scale battery storage would contribute significantly to the social welfare of the California electricity market, adding $2.42 per discounted MWh of renewable energy generated, as of 2018, of $3.8 billion in total. In the absence of battery subsidies or mandates, we would likely see large-scale battery adoption by about 2030, spurred by declining capital costs and increasing renewable energy penetration. Installation subsidies of 35% would be necessary to meet California's battery storage mandate by 2024. If the battery storage were owned by a monopolist rather than a competitive market, more adoption would occur earlier but the adoption would create less social surplus.

[44] The Energy Information Administration (EIA) projects world energy consumption will increase by an average of 1.4% per year through 2040 (Conti et al., 2016). As of 2011, less than 3% of all delivered electricity used storage (Dunn et al., 2011).

References

Berrada, A., Loudiyi, K., and Zorkani, I. (2016). Valuation of energy storage in energy and regulation markets. *Energy*, 115:1109–1118.

Black, M. and Strbac, G. (2007). Value of bulk energy storage for managing wind power fluctuations. *IEEE transactions on energy conversion*, 22(1):197–205.

Burr, C. (2014). Subsidies, tariffs and investments in the solar power market. *University of Colorado-Boulder Working Paper*.

Bushnell, J. and Novan, K. (2018). Setting with the sun: The impacts of renewable energy on wholesale power markets. Technical report, National Bureau of Economic Research.

Carson, R. T. and Novan, K. (2013). The private and social economics of bulk electricity storage. *Journal of Environmental Economics and Management*, 66(3):404–423.

Cheng, B. and Powell, W. B. (2016). Co-optimizing battery storage for the frequency regulation and energy arbitrage using multi-scale dynamic programming. *IEEE Transactions on Smart Grid*, 9(3):1997–2005.

Cole, W. J. and Frazier, A. (2019). Cost projections for utility-scale battery storage. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Conti, J., Holtberg, P., Diefenderfer, J., LaRose, A., Turnure, J. T., and Westfall, L. (2016). International Energy Outlook 2016 With Projections to 2040. Technical report, USDOE Energy Information Administration (EIA), Washington, DC (United States). Office of Energy Analysis.

Craig, M. T., Jaramillo, P., Hodge, B.-M., Williams, N. J., and Severnini, E. (2018). A retrospective analysis of the market price response to distributed photovoltaic generation in california. *Energy policy*, 121:394–403.

Cullen, J. (2013). Measuring the environmental benefits of wind-generated electricity. *American Economic Journal: Economic Policy*, 5(4):107–33.

Cullen, J. A. (2010). Dynamic response to environmental regulation in the electricity industry. In *Industrial Organization Seminar*, page 50.

Cullen, J. A. and Reynolds, S. S. (2017). Market dynamics and investment in the electricity sector. Technical report, Working paper.

De Groote, O. and Verboven, F. (2019). Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems. *American Economic Review*, 109(6):2137–72.

Deaton, A. and Laroque, G. (1992). On the behaviour of commodity prices. *The Review of Economic Studies*, 59(1):1–23.

Dunn, B., Kamath, H., and Tarascon, J.-M. (2011). Electrical energy storage for the grid: A battery of choices. *Science*, 334(6058):928–935.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods: Second Edition*. Number 9780199641178 in OUP Catalogue. Oxford University Press.

EIA, U. (2020). Battery storage in the united states: An update on market trends. *Washington, DC: US EIA*.

Feger, F., Pavanini, N., and Radulescu, D. (2017). Welfare and redistribution in residential electricity markets with solar power.

Garcia-Gonzalez, J., de la Muela, R. M. R., Santos, L. M., and Gonzalez, A. M. (2008). Stochastic joint optimization of wind generation and pumped-storage units in an electricity market. *IEEE Transactions on Power Systems*, 23(2):460–468.

Goldie-Scot, L. (2019). A behind the scenes take on lithium-ion battery prices. *BloombergNEF*. Available at https://about.bnef.com/blog/behind-scenes-take-lithium-ion-battery-prices/.

Gowrisankaran, G., Reynolds, S. S., and Samano, M. (2016). Intermittency and the value of renewable energy. *Journal of Political Economy*, 124(4):1187–1234.

Gowrisankaran, G. and Rysman, M. (2012). Dynamics of consumer demand for new durable goods. *Journal of political Economy*, 120(6):1173–1219.

Harvey, A. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Hendel, I. and Nevo, A. (2006). Measuring the implications of sales and consumer inventory behavior. *Econometrica*, 74(6):1637–1673.

Hittinger, E. S. and Azevedo, I. M. (2015). Bulk energy storage increases United States electricity system emissions. *Environmental science & technology*, 49(5):3203–3210.

Holladay, J. S. and LaRiviere, J. (2018). How does welfare from load shifting electricity policy vary with market prices? evidence from bulk storage and electricity generation. *The Energy Journal*, 39(6).

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):1127–1150.

Joskow, P. L. (2011). Comparing the costs of intermittent and dispatchable electricity generating technologies. *The American Economic Review P&P*, 101(3):238–241.

Kanamura, T. and Ōhashi, K. (2007). A structural model for electricity prices with spikes: Measurement of spike risk and optimal policies for hydropower plant operation. *Energy Economics*, 29(5):1010–1032.

Karaduman, O. (2019). Economics of grid-scale energy storage. Technical report. Available at: https://economics.mit.edu/files/18357.

Kazemi, M., Zareipour, H., Amjady, N., Rosehart, W. D., and Ehsan, M. (2017). Operation scheduling of battery storage systems in joint energy and ancillary services markets. *IEEE Transactions on Sustainable Energy*, 8(4):1726–1735.

Kirkpatrick, A. J. (2018). Estimating congestion benefits of batteries for unobserved networks: A machine learning approach.

Knittel, C. R. and Roberts, M. R. (2005). An empirical examination of restructured electricity prices. *Energy Economics*, 27(5):791–817.

Langer, A. and Lemoine, D. (2018). Designing dynamic subsidies to spur adoption of new technologies. Technical report, National Bureau of Economic Research.

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive Macroeconomic Theory*. The MIT Press.

Lucas, R. E. and Prescott, E. C. (1971). Investment under uncertainty. *Econometrica*, 39(5):659–681.

Mansur, E. T. (2008). Measuring welfare in restructured electricity markets. *The Review of Economics and Statistics*, 90(2):369–386.

Mohsenian-Rad, H. (2015). Optimal bidding, scheduling, and deployment of battery systems in California day-ahead energy market. *IEEE Transactions on Power Systems*, 31(1):442–453.

Mokrian, P. and Stephen, M. (2006). A stochastic programming framework for the valuation of electricity storage. Technical report, 26th USAEE/IAEE North American Conference.

Novan, K. (2015). Valuing the wind: renewable energy policies and air pollution avoided. *American Economic Journal: Economic Policy*, 7(3):291–326.

Paatero, J. V. and Lund, P. D. (2005). Effect of energy storage on variations in wind power. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, 8(4):421–441.

Pirrong, C. (2012). *Commodity Price Dynamics: A Structural Approach*. Cambridge University Press.

Prescott, E. C. and Mehra, R. (1980). Recursive competitive equilibrium: The case of homogeneous households. *Econometrica*, 48(6):1365–1379.

Proietti, T. (2006). Temporal disaggregation by state space methods: Dynamic regression methods revisited. *The Econometrics Journal*, 9(3):357–372.

Reddix, K. (2015). Powering demand: Solar photovoltaic subsidies in california.

Reguant, M. (2014). Complementary Bidding Mechanisms and Startup Costs in Electricity Markets. *The Review of Economic Studies*, 81(4):1708–1742.

Reynolds, S. (2019). Notes on computation for policy function iteration. Technical report, University of Arizona Mimeo.

Sackler, D. (2019). New battery storage on shaky ground in ancillary service markets. *Utility Dive*. Available at https://www.utilitydive.com/news/new-battery-storage-on-shaky-ground-in-ancillary-service-markets/567303/.

Sioshansi, R., Denholm, P., Jenkin, T., and Weiss, J. (2009). Estimating the value of electricity storage in PJM: Arbitrage and some welfare effects. *Energy Economics*, 31(2):269–277.

Sioshansi, R. and others (2011). Increasing the value of wind with energy storage. *Energy Journal*, 32(2):1–29.

Walawalkar, R., Apt, J., and Mancini, R. (2007). Economics of electric energy storage for energy arbitrage and regulation in New York. *Energy Policy*, 35(4):2558 – 2568.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081.

Wolak, F. A. (2018). The evidence from california on the economic impact of inefficient distribution network pricing. Technical report, National Bureau of Economic Research.

Woo, C.-K., Moore, J., Schneiderman, B., Ho, T., Olson, A., Alagappan, L., Chawla, K., Toyama, N., and Zarnikau, J. (2016). Merit-order effects of renewable energy and price divergence in california's day-ahead and real-time electricity markets. *Energy Policy*, 92:299–312.

Xi, X., Sioshansi, R., and Marano, V. (2014). A stochastic dynamic programming model for co-optimization of distributed energy storage. *Energy Systems*, 5(3):475–505.

Xu, B., Oudalov, A., Ulbig, A., Andersson, G., and Kirschen, D. S. (2016). Modeling of lithium-ion battery degradation for cell life assessment. *IEEE Transactions on Smart Grid*, 9(2):1131–1140.

Appendix

A Capital Cost Estimation

As shown in Section 4 , capital costs evolve according to the following process:

$$c_y = c_{y-1} e^{\tau} e^{\xi_y}, \quad \xi_t \sim N(0, \sigma_c^2) \tag{18}$$

Taking the natural log of each side of the equation and subtracting yields:

$$log(c_y) - log(c_{y-1}) = \tau + \xi_y \tag{19}$$

We can then rewrite costs in terms of the initial cost c_0 in the year 2018.

$$log(c_y) - log(c_0) = \tau \cdot y + \sum_{1}^{y} \xi_y \tag{20}$$

Finally, we use Equation 20 to derive the following moment conditions:

First Moment

$$E[log(c_y) - log(c_0)] = \tau \cdot y \tag{21}$$

Second Moment

$$\begin{aligned} Var[log(c_y) - log(c_0)] &= Var[y\tau + \sum_{1}^{y} \xi_y] \\ Var[log(c_y) - log(C_0)] &= Var[y\tau] + Var[\sum_{1}^{y} \xi_y] \\ Var[log(c_y) - log(c_0)|y] &= y \cdot Var[\xi_y] \\ SD[log(c_y) - log(c_0)|y] &= \sqrt{y} \cdot SD[\xi_y] \\ SD[log(c_y) - log(c_0)|y] &= \sqrt{y} \cdot \sigma \end{aligned} \tag{22}$$

We find the parameters τ and σ_c that solve the two moment conditions by estimating two univariate regressions. For the first regression the dependent variable is $log(c_y) - log(c_0)$ and the dependent variable is y. For the second regression, the dependent variable is the standard deviation of all the cost realizations $(log(c_y) - log(c_0)$ conditional on y, and the independent variable is $\sqrt{y}$. In the second regression, we only have one observation for each year, so we weight the regression by the number of cost projections that were made for that year. Figure 2 shows that years that are further in the future tend to have fewer cost projections.

Without loss of generality we also normalize the initial capital cost to $c_0 = 1$ before estimating the model. Therefore, $1 - c_y$ represents the percentage reduction in costs at year t relative to

2018. We then re-scale the cost process before solving the investment problem so that $\tilde{c}_t = 380 * c_t$ consistent with Cole and Frazier (2019) that assume capital costs in 2018 are $380/kWh.

B Modeling Battery Capacity Degradation

We model capacity degradation using the model from Xu et al. (2016). In the model, the degradation rate of a battery cells depends on the following factors: (1) temperature, (2) depth-of-discharge, (3) state-of-charge, (4) calendar time, (5) number of cycles. Battery degradation is therefore a nonlinear process with respect time and stress cycles. For our application, we assume that the battery is operated at 25 C (77 F) throughout the year.

Let the K denote the batteries capacity this period and let K' denote the batteries capacity next period. Furthermore, let g_d be a function that determines degradation between the current period and next period.

$$K' = Ke^{-g_d} \tag{23}$$

Battery degradation, g_d, consists of calendar degradation, g_t, and cycle degradation, g_c. Define N as the total amount of cycles that the battery undertakes, and let n_i be a variable that indicates if cycle i was a full cycle ($n_i = 1$) or a half cycle ($n_i = 0.5$):

$$g_d = g_t + \sum_{i}^{N} n_i g_c \tag{24}$$

Calendar degradation refers to degradation that occurs over time regardless of how much the battery is charged or discharged. Calendar degradation is a function of time as well as the battery's mean state-of-charge. Battery capacity will degrade more if the battery is left idle at full state-of-charge relative to if the battery is left idle at 50% SOC. More concretely, calendar degradation is a function of time elapsed in seconds, $t^{(s)}$, as well as mean state-of-charge during the time elapsed, $\bar{\sigma}$

$$g_t = 0.000000000414 * t^{(s)} * e^{1.04(\bar{\sigma}-0.5)} \tag{25}$$

The second component of the degradation function is the portion attributed to cycling activity, the capacity reduction due to cycling depends on the mean state-of-charge during cycle i , σ_i, as well as the depth of discharge of the cycle, δ_i. A larger depth of discharge causes more capacity degradation. Cycling from 0% to 100% once is more damaging than cycling from 25-75% twice.

$$f_c = e^{1.04(\sigma_i-0.5)} * (140000\delta_i^{-0.501} - 123000)^{-1} \tag{26}$$

By substituting equations 2,3,4 into Equation 1, we see that capacity (next period) is a function of $t^{(s)}$, N, $\bar{\sigma}$, n_i, δ_i, and σ_i.

We perform the following algorithm to simulate battery degradation:

1. First solve for the optimal policy, using annual discount factor of $\beta = 0.95$
2. Use the optimal policy from (1) to simulate the charge/discharge actions using the realized stream of price residuals ε^P, load residuals ε^L, and supply curve parameters across all time periods in month m.
    - For a 1 month simulation $t^{(s)} = 60 * 60 * 24 * 30 = 2592000$
    - Record the batteries' state-of-charge for each time interval (5 min) of the simulation
3. Use the simulated path of state-of-charge levels to calculate the mean state-of-charge over the simulation period $\bar{\sigma}$
4. Feed the simulated path of charge levels into a rainflow cycle counting algorithm.
    - See https://www.mathworks.com/matlabcentral/fileexchange/3026-rainflow-counting-algorithm
    - This rainflow counting algorithm will return N, $\bar{\sigma}$, n_i, δ_i, σ_i. In words, it will count the number of cycles (half and full), and determine the mean state of charge for each cycle and the depth-of-discharge for each cycle.
5. Calculate the total degradation rate e^{-g_d} and for each month-long simulation. Then calculate the lifetime value of the battery by dividing the sum of the flow profits by the sum of the monthly discount rate and the capacity degradation rate:

$$\text{Storage Value} = \frac{\text{Sum of Simulated Flow Profits During Month Simulation}}{(1 - e^{-g_d}) + (1 - 0.95)\frac{1}{12}} \tag{27}$$

* Note that this formulation technically assumes that both power and energy capacity are being reduced. The engineering literature says that only energy capacity should degrade, so this calculation should provide an upper bound on the value of storage.

C Tables & Figures

Table 5: Summary Statistics

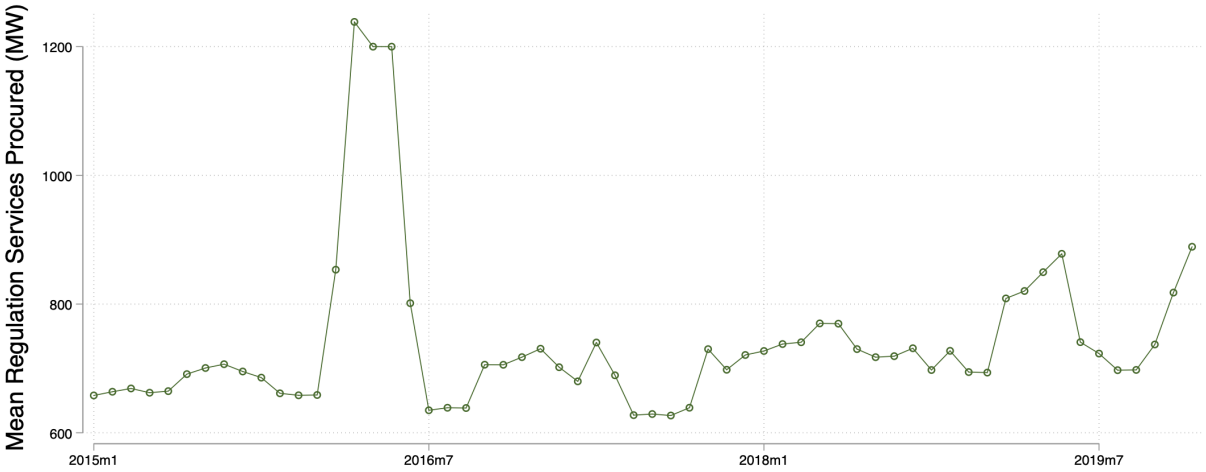
Daytime 10 AM - 3 PM

	2015	2016	2017	2018	2019	Total
Real-Time Price ($/MWh)	25.57	24.45	23.49	27.20	25.54	25.25
	(67.06)	(80.68)	(69.44)	(61.27)	(54.62)	(67.20)
Day-Ahead Price ($/MWh)	27.80	22.47	24.14	32.13	23.75	26.06
	(8.848)	(11.10)	(15.73)	(32.32)	(14.67)	(18.82)
Battery Discharge (MW)	.	.	.	-2.908	-5.038	-4.146
	(.)	(.)	(.)	(25.65)	(35.34)	(31.66)
Solar Generation (MW)	4373.3	5634.9	7264.2	7802.7	8219.6	6658.4
	(1033.5)	(1480.1)	(1949.3)	(1939.2)	(2406.1)	(2323.1)
Imports (MW)	6659.8	6246.6	5182.9	4742.0	3295.4	5314.3
	(1106.2)	(1266.2)	(1494.9)	(1811.8)	(2068.9)	(1934.4)
Hydro Generation (MW)	1306.1	2324.4	3272.6	2114.7	2886.3	2357.6
	(568.3)	(843.3)	(932.1)	(569.6)	(809.8)	(1018.5)
Net Load (MW)	22442.0	19760.4	17745.0	16452.8	14698.5	18220.6
	(4840.2)	(4440.3)	(5368.5)	(5223.0)	(5152.0)	(5686.9)

Evening Peak 4 PM - 8 PM

	2015	2016	2017	2018	2019	Total
Real-Time Price ($/MWh)	44.14	42.59	54.51	58.90	51.58	50.34
	(95.11)	(100.4)	(135.3)	(129.1)	(105.9)	(114.5)
Day-Ahead Price ($/MWh)	41.63	40.03	56.13	67.30	52.38	51.49
	(9.889)	(13.47)	(46.49)	(70.53)	(27.94)	(41.70)
Battery Discharge (MW)	.	.	.	9.239	3.436	5.866
	(.)	(.)	(.)	(27.98)	(34.18)	(31.86)
Solar Generation (MW)	1070.7	1442.9	1882.7	2078.6	2229.8	1740.8
	(1475.2)	(1997.0)	(2679.7)	(2896.0)	(3149.8)	(2551.7)
Imports (MW)	7838.8	8296.1	7981.6	8126.7	6969.8	7882.9
	(1290.2)	(1406.6)	(1583.5)	(2079.5)	(2444.5)	(1831.1)
Hydro Generation (MW)	2190.5	3422.3	4222.8	3128.0	3930.9	3353.6
	(724.2)	(948.9)	(898.3)	(714.6)	(807.9)	(1089.5)
Net Load (MW)	27659.4	26383.9	25983.7	25285.4	24180.3	25898.8
	(4796.5)	(4823.0)	(5666.0)	(5477.4)	(5782.1)	(5449.0)

Figure 10: Regulation Service Quantity Procured in CAISO



Note: The figure plots the mean hourly quantity of regulation services procured by CAISO each month. Regulation quantity is calcualted the sum of "regulation up" and "regulation down" quantities in the day-ahead market.

Figure 11: Storage Applications in the Day-Ahead Market

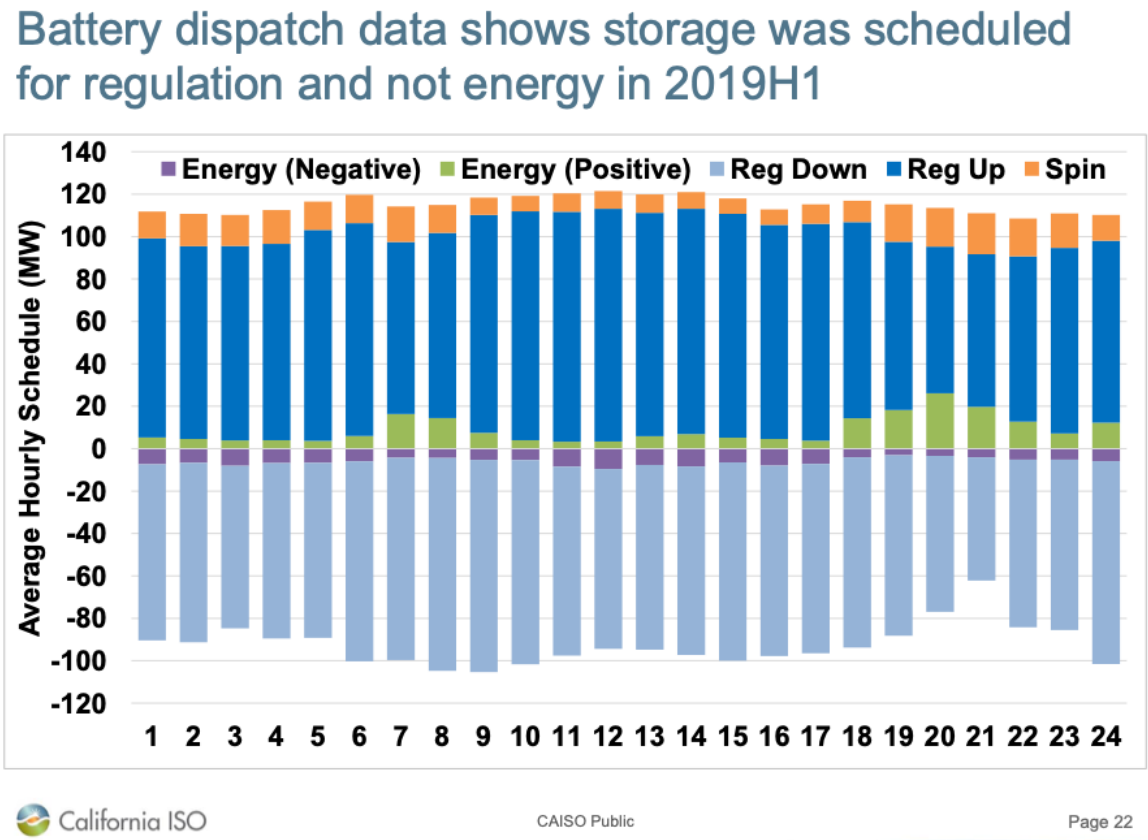


Table 6: Summary Statistics of Daily Marginal Cost Curve Parameters

Parameter	2015	2016	2017	2018	2019	2016-19
δ						
Mean	-47.160	-59.863	-38.812	-26.875	-30.305	-38.978
Std.	56.789	75.965	58.459	45.561	49.610	59.922
θ						
Mean	64.299	90.093	56.125	35.629	42.349	56.072
Std.	94.598	130.004	98.374	74.279	82.789	100.817
κ						
Mean	2.575	3.088	2.563	2.108	2.326	2.522
Std.	2.015	2.851	2.468	2.040	2.347	2.468
α						
Mean	0.722	0.850	0.816	0.815	0.796	0.819
Std.	0.238	0.114	0.137	0.137	0.133	0.132
ψ						
Mean	0.431	0.277	0.341	0.325	0.315	0.314
Std.	0.305	0.163	0.240	0.237	0.221	0.218

Notes: This table summarizes the means and standard deviations of the daily marginal cost curve parameters over distinct parts of our analysis sample.

D Brief Discussion of Kalman Filter/Smoother

In section 4.1, we described our steps for estimation of the time series model of electricity demand and marginal costs. Importantly, our estimation procedure used information on prices and quantities that are both set (forecasted) in the day-ahead market, as well as the prices and quantities that are ultimately realized in the real-time market at each five minute interval.

One complication that arises with using the day-ahead market is that the California system operator (CAISO) sets the day-ahead market only at the hourly frequency, reporting prices and forecasts for net load that are constant over the 12 five minute intervals over the hour. As our operations model is based on the five minute frequency, our estimation procedure must be able to accommodate the mixed-frequency nature of the quantities and prices set in both the day-ahead and real-time markets.

We overcome this hurdle by incorporating the Kalman filter/smoother in an effort to temporally disaggregate (e.g., or interpolate) the forecasts of netload to yield a forecast at the five minute frequency.

Specifically, assume that the a series F_t is observed only every h periods, and what is observed is the average of the interim h periods of the latent process f_t, or $F_t = \sum_{j=0}^{h-1} f_{t-j}$. The objective is to take a series F_t and construct an estimate of the latent process f_t such that the implied values for F_t match what's observed. Casting the problem in a state space model and using the

Kalman filter/smoother to estimate the latent process is well documented in the literature (e.g., Proietti (2006)).

Casting the process as a state space model can be done as follows:

$$\begin{aligned} F_t &= Z_t f_t \\ f_t &= f_{t-1} + \eta_t \end{aligned}$$

where Z_t is a time varying selection matrix designed to handle the possibly missing observations during the interim periods before the temporally aggregated version, (F_t) is observed, and η_t is a serially independent error term that contributes to the time series variation in the latent process of interest f_t.

Using techniques outlined in Durbin and Koopman (2012); Harvey (1989) to build the appropriate augmentation to the state space system matrices consistent with the net load forecast set in the day-ahead market representing an *average* of CAISO's implicit forecast for net load over each five minute period that hour, we apply the Kalman filter/smoother recursions to recover an estimate of the latent f_t for each day in our sample. These estimates can then be used as the deterministic portion of net load, X_s^L, where deviations from these forecasted values are captured by ε_t^L.