

# Estimating Congestion Benefits of Batteries for Unobserved Networks: A Machine Learning Approach

A. Justin Kirkpatrick \*

January 9, 2020

Preliminary - Please Do Not Cite

## Abstract

Energy storage investment in the U.S. is forecast to reach \$2.5B annually by 2020 largely due to state-level mandates and subsidies. The justification for these policies is that energy storage facilitates grid integration of renewable generation by smoothing out the frequency and severity of price spikes due to intermittent renewable supply. While operators of energy storage generate private returns through arbitrage of diurnal price differences, ratepayer benefits are derived from decreases in these price spikes. The result is a transfer from infra-marginal generators to retail utilities and consumers. This paper presents empirical estimates of energy storage price effects in California where the locations and hourly prices for 372MW of energy storage are observed. Results suggest that one megawatt of energy storage decreases evening peak prices by up to 2.2% at the pricing node where the storage is installed, a benefit to ratepayers of \$62,467 per year. This effect coincides with the late-afternoon increase in electricity prices associated with intermittent solar generation. A Double Pooled LASSO-based estimator is used to uncover the unobserved network structure in order to estimate the price effects of storage at other locations across the grid. Together, the combined results suggest that energy storage mandates in California are partially justified by ratepayer benefits.

---

\*Michigan State University, Department of Economics, 110 Marshall-Adams Hall, East Lansing, MI 48824 (email: jkirk@msu.edu). I thank Lori Bennear, Steve Sexton, Chris Timmins, Bryan Bollinger, Frank Wolak, Marty Smith, Michael Caramanis, and participants at Camp Resources for comments and suggestions, as well as Ray Hohenstein at AES for providing context. Generous support for this research was provided by the Alfred P. Sloan Foundation Pre-doctoral Fellowship on Energy Economics, awarded through the NBER.

# 1 Introduction

Since the time of Edison, the absence of economical electricity storage has raised the cost of reliable electricity supply. Compelled to have supply equal demand at all moments in time, utilities must invest in seldom-used generation capacity characterized by low capital costs and high marginal costs. These “peaker” plants are called up to meet diurnal and seasonal peak demand, some for just a few hours on a few days each year.

The inability to store electricity is becoming more problematic amid substantial investments in intermittent renewable generation capacity that are partly a response to federal and state policies intended to achieve environmental and energy security objectives. By 2017, there were more than 1.2 million household solar installations across the US, 30.3 GW of utility-scale solar, and 80.5 GW of wind (U.S. Department of Energy, 2017). This renewable capacity cannot be dispatched by generators or grid operators. It is, instead, governed by sunlight or wind, which vary systematically within days and across seasons and idiosyncratically across days due to weather. This intermittency imposes an additional burden upon the grid operator who must ensure dispatchable generation meets demand net of renewable supply. Moreover, as renewable penetration grows, it can be expected to crowd out dispatchable baseload capacity like coal and gas plants because its marginal costs of generation are essentially nil. Conventional power plants, therefore, increasingly sit idle during much of the day and undertake costly processes to ramp up generation to match fluctuations in renewable supply. The cost of conventional dispatch and, hence, the value of electricity storage are growing with renewable capacity.

Energy storage is expected to facilitate integration of intermittent renewable generation by shifting supply from periods of abundant renewable generation to periods when it is scarce and dispatchable generation must be called up to meet the difference in load and renewables supply. Storage lowers peak prices by reducing the quantity of dispatchable generation demanded during peak periods. Because most electric grids are characterized by convex supply curves, the price reductions achieved by battery discharge at peak demand exceed the price increases induced by the recharging of storage capacity during off-peak periods. Such marginal price changes affect prices paid for every unit of generation because wholesale markets are commonly settled by uniform price auctions. Batteries also lower costs associated with grid congestion. Congestion induces sub-optimal production because electricity trades across areas are constrained by transmission capacity. Reductions in costs of serving electricity demand, or load, are savings to load-serving entities (LSEs) that constitute transfers from non-marginal generators (Walawalkar et al., 2008). In states with cost-of-

service regulation, such cost savings are expected to be passed on to ratepayers. These grid benefits are not appropriated by private storage operators, whose returns from arbitraging high and low prices are expected to decline in storage capacity. It can be shown that prices with storage converge to a price strictly lower than the mean wholesale price without storage.

Though battery storage capacity is expected to grow 1.6GW per year by 2020, reflecting \$2.5B in annual investment (GTM Research, 2017), the impacts of batteries on prices and costs of serving electricity load have not heretofore been empirically estimated in a manner robust to grid congestion, market power, and alternative battery operator objectives. This paper, therefore, examines the effect of storage-induced congestion relief on wholesale electricity prices in California, the largest market for battery storage in the world with 372 MW of installed capacity. Specifically, I develop an empirical model of electricity prices that accounts for intermittent generation and congestion. The model is used to estimate price changes using variation in when and where on the grid 77 unique utility-scale batteries are installed. Price changes are estimated for the nodes at which storage is installed and for other nodes where prices are likely to respond because of grid dependencies. These results are aggregated to estimate the magnitude of grid-level congestion-relief benefits generated by marginal storage capacity. Because the architecture of the electric grid is unobserved due to national security considerations, I use a machine learning technique borrowed from the Research and Development economics literature to determine price dependencies across nodes of the grid.

A 1-megawatt (MW) increase in storage capacity is estimated to reduce afternoon peak prices by up to 2.2% in the day ahead market, primarily in the late-afternoon and early evening, coinciding with the period of highest electricity net demand. The annual cost of serving load is estimated to fall by \$62,647 from a 1MW increase in storage capacity. This annual benefit from congestion relief partly justifies storage mandates like that of California, which are currently estimated to cost \$6,500,000 per MW of capacity (Penna et al., 2016), though costs are dropping rapidly. These grid benefits are approximately 29% to 126% of the private benefits that accrue to battery operators.

Though price effects of storage and its grid-level impacts have not previously been flexibly and empirically estimated, a small engineering literature considers their market impacts using simulation methods, focusing almost entirely on the private returns of storage investments and the battery properties that drive potential profits (Bradbury et al., 2014; Fioravanti et al., 2013; Nottrott et al., 2013; Walawalkar et al., 2007; Hittinger et al., 2012). Only Sioshansi et al. (2009) considers the effect of storage on grid prices, primarily to consider

whether arbitrage opportunities are dissipated when storage is used. The authors estimate the effect of energy storage on grid prices assuming a constant linear relationship between load and prices, and do so by estimating price responses to changes in demand, not storage capacity, limiting the robustness of results to grid characteristics like congestion. Lueken and Apt (2014) estimate potential savings of \$4 billion in the PJM grid that covers Mid-Atlantic and Northeastern states.

Thus, this paper contributes to a growing literature that explores the impacts of new grid technology on market outcomes and grid operations (Woo et al., 2011, 2016; Novan, 2015; Bushnell and Novan, 2018; Craig et al., 2018; Wolak, 2016, 2018). It also introduces a novel empirical approach to recovering the market-relevant network characteristics of the electricity grid when they are unreported by regulators or utilities.

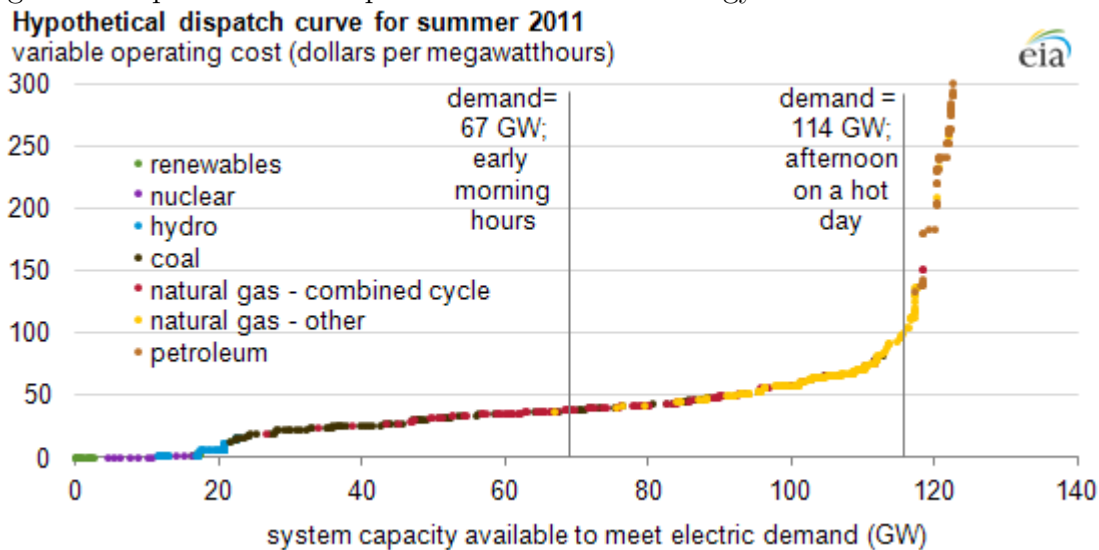
This paper proceeds by introducing a conceptual framework of spatial electricity price determination in Section 2. Data and empirical approach are discussed in Section 3. Section 4 presents empirical results, Section 5 discusses results, and a final section concludes.

## 2 Conceptual model

### 2.1 Electricity markets, congestion, and prices

Electricity prices are governed by two phenomena. The first is an upward-sloping, convex supply curve defined by marginal costs that increase in quantity supplied. In an unconstrained competitive market, an inverse supply curve is constructed by ordering electricity generation capacity from lowest to highest marginal cost, as is shown in Figure 1. This “merit order” defines a supply curve that becomes steeper as quantity increases because of generation technology and the limited frequency at which “peaker plants” are called up. This increasing steepness leads to a large price decrease when storage is discharged during a peak period relative to a small price increase when storage is charged during an off-peak period. The net effect is to lower average prices even though total quantity of electricity demand is unchanged.

Figure 1: Representative dispatch curve. Source: Energy Information Administration



The second phenomenon governing wholesale electricity prices is congestion on the grid caused by transmission constraints that lead to out-of-merit order generation dispatch. Marginal costs of generation diverge across local markets defined by transmission constraints, and in a market with congestion, dispatch occurs over potentially many local supply curves, with more congested nodes facing a steeper supply curve. The introduction of storage capacity at one of these nodes can affect dispatch at other nodes.

In centrally-dispatched energy markets such as those managed by the California Independent Systems Operator (CAISO) and by PJM in the northeast and mid-atlantic states, energy withdrawals are priced at the network node from which the withdrawals are made. A *node* is a physical substation or transmission point where generators may inject power, or load-serving entities (LSE's) may withdraw power. Nodal prices vary spatially and temporally (hourly or in 5-minute increments), and reflect the cost of withdrawing one unit of energy at that node, and injecting one unit of energy at a reference node.<sup>1</sup> Nodal prices determine the revenue that generators receive and the amount paid by LSEs. They are determined by the solutions to a constrained cost-minimization problem in which the grid operator must secure, via a uniform price auction, sufficient generation to meet node-level demand in every instant. The grid operator is constrained by transmission line capacity.

<sup>1</sup>All withdrawals must have a concurrent injection to maintain the equality between supply and withdrawal. In practice, a withdrawal is offset by an increase in supply from a generator located on the grid, though not necessarily at the reference node. Both withdrawal and injection are priced relative to the reference node; the choice of reference node is arbitrary and does not affect nodal price levels.

Prices and generation change according to a non-linear Karush-Kuhn-Tucker (KKT) system of equations where slack conditions are on either generation constraints (e.g. prices change as higher-cost generation is required) or transmission constraints (e.g. prices change as lower-cost generation is unable to serve certain node-level demands) (Bohn et al., 1984).

Constrained optimization is a common tool in the economics literature, and is frequently used to generate equilibrium conditions and to define optimal production or consumption decisions. Here, the constrained optimization is not simply a convenient model for representing the grid. In the CAISO, network dispatch is determined by a numeric solution to the optimization problem run hourly by CAISO staff. Rather than a representation of the optimal dispatch, a constrained optimization *is* the dispatch process. As such, I rely on the properties of the constrained optimization - specifically the shadow values of constraints - to develop intuition about the effects of storage.

Electricity demand is assumed to be exogenous because consumers face retail rates that may vary with time of day but do not vary instantaneously with wholesale prices. Nodal price is primarily determined by the generators' supply bids and by transmission constraints that may preclude least-cost dispatch to serve some nodes. The effects of battery storage on nodal prices vary by congestion status.

## 2.2 Generation in a Single Node Market

To illustrate these price effects, consider first an uncongested market wherein a single, fixed set of generators serves demand. Let  $p(Q) \geq 0$  be the marginal cost of energy at total quantity demanded  $Q$ . I assume the marginal cost of energy is increasing at an increasing rate in  $Q$ , i.e.

$$\frac{dp}{dQ} \geq 0 \tag{1}$$

$$\frac{dp^2}{d^2Q} \geq 0 \tag{2}$$

Equation 2 is a consequence of merit-order dispatch and generator technologies, as shown in Figure 1. Peaking plants used to serve periods of high demand are less capital intensive, but are less efficient and thus more expensive per unit generated.

Let  $E$  be the amount of energy a storage system transacts from the grid in an hour, where a transaction is either a discharge to the grid or a withdrawal from the grid. Let  $Q^L$

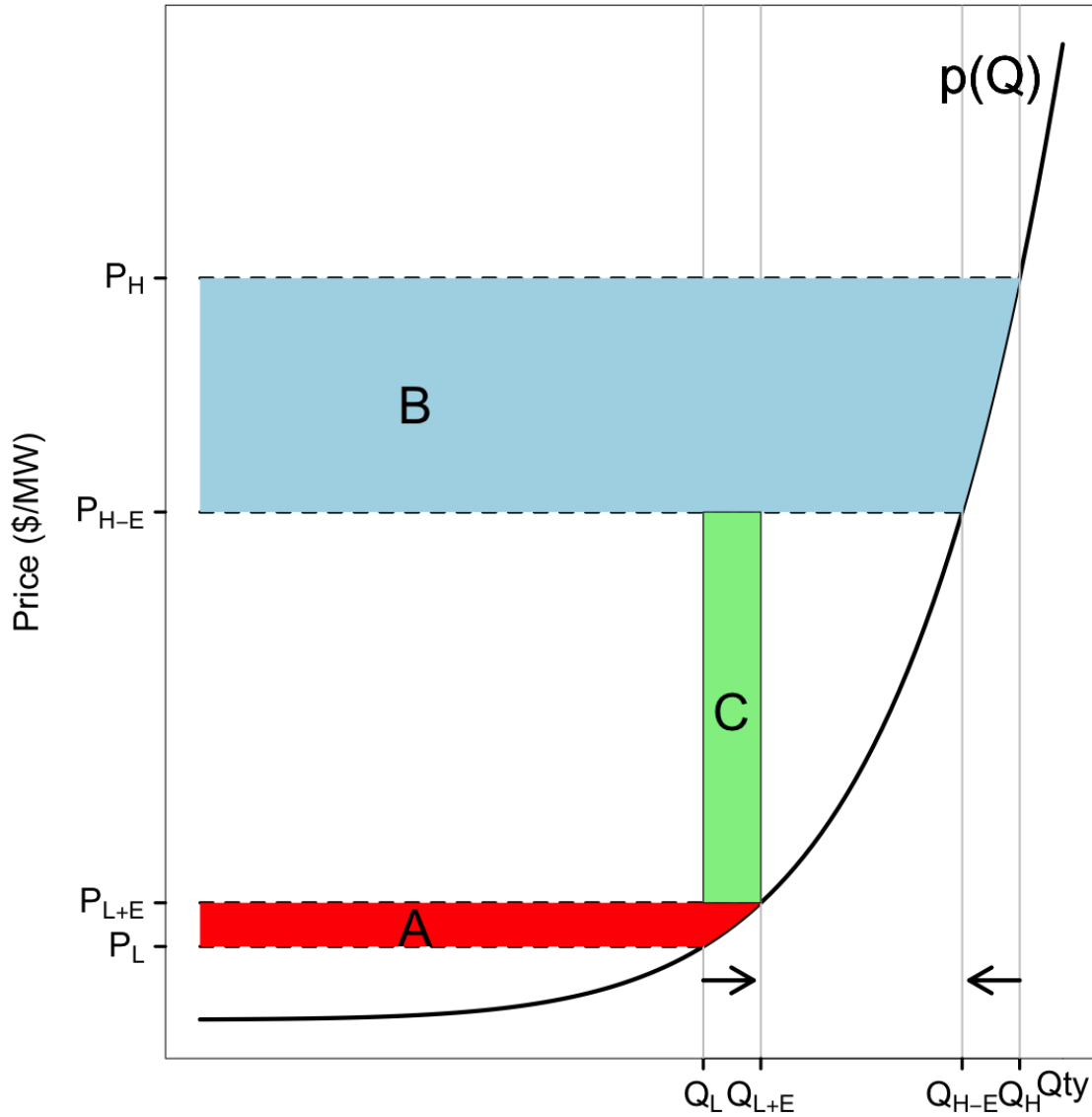
and  $Q^H$  be (nighttime) nadir and (daytime) peak exogenous demand, respectively, such that  $Q^L < Q^H$ . The withdrawal of  $E$  from the grid during the off-peak period increases  $Q^L$  to  $Q^L + E$ . The new price paid for all units of energy is  $p(Q^L + E)$ . Similarly,  $Q^H$  is reduced by  $E$  when energy storage discharges during the peak period, and  $p(Q^H - E) \leq p(Q^H)$ . Not only is  $p(Q^H) \geq p(Q^L)$ , creating a potentially profitable private opportunity for arbitrage (labeled  $C$  in Figure 2), but also:

$$\frac{dp(Q^H)}{dQ} \geq \frac{dp(Q^L)}{dQ}; \quad (3)$$

due to convexity of the supply curve. Hence, the price changes from injection to and withdrawal from the grid may not be symmetric. For strictly convex supply, storage lowers marginal costs at peak more than it raises them at low demand.

While the energy storage operator profits per unit of energy stored, equal to the difference in  $p(Q^H - E)$  and  $p(Q^L + E)$ , load serving entities (LSEs) face decreases in peak prices that are expected to outweigh increases in nadir prices. Because infra-marginal generators (e.g. baseload generators) receive the same prices as those on the margin, batteries reduce the price of every unit sold during peak periods and increase prices of every unit sold during the nighttime nadir. The gain to inframarginal generators (and cost to LSEs) from nadir price increases is depicted in area  $A$  in Figure 2, while the gain to LSEs from peak price reductions is depicted as region  $B$ . Importantly, because the reduction in peak prices is at least weakly greater than the increase in nadir prices and because  $Q^H > Q^L$ , peak cost savings to LSEs dominate nadir cost increases, i.e., area  $B$  is greater than area  $A$ .

Figure 2: Price effects of a discharge of energy from energy storage during daytime peak and nighttime nadir times.  $Q$  is quantity of energy, and  $P$  is price. Area  $C$  is the private benefit accruing to the energy storage operator. Area  $B$  is the surplus gained by the LSE (load serving entity; utility) through the price effect during peak hour discharge. Area  $A$  is the cost to the LSE through the upward price effect during off-peak hour charging



These total net benefits to LSEs are given by:



$$\text{Total Benefits} \cup Q^H \frac{dp(Q^H)}{dQ} - Q^L \frac{dp(Q^L)}{dQ} > 0 \quad (4)$$

Thus, storage is expected to affect a net transfer from infra-marginal generators to LSEs. Such savings to LSEs are passed onto retail electricity customers under cost-of-service regulation.

Net revenues for the battery owners are equal to  $C$ , the quantity stored multiplied by the difference between high and low prices, and do not include the net benefits  $B - A$ .

## 2.3 Congestion

In a multi-node network, Kirchoff’s Law determines energy flows over each transmission line between nodes, or “edge” in network terminology. Edges between nodes have two properties: *susceptance* is the ease with which energy flows over an edge, and *capacity* is the maximum amount of energy that can flow over an edge. Given two different edges between two nodes, for example, a generation node and a withdrawal node, electricity flows in proportion to the susceptance of the edges.<sup>2</sup> If one edge has a susceptance of 2, and another has a susceptance of 1, then  $\frac{1}{3}$  of the flow between nodes will flow over the edge of lower susceptance, and  $\frac{2}{3}$  will flow over the edge of higher susceptance. If the capacity of each edge is identical, one edge will reach capacity before the other line. When this constraint binds, additional energy cannot be moved between nodes without serious consequences for safety. Even though one edge has not reached capacity, no additional power can flow between these two example nodes. This is in contrast to a highway network where congestion on one highway causes motorists to reroute to another highway.

Kirchoff’s law precludes “choosing where electrons flow”—when a transmission line reaches capacity but demand is increasing, the only two options available to the grid operator are to either reduce withdrawals at the withdrawal node, or to increase injection at another node whose path does not include the congested line. A consequence of Kirchoff’s law is that the flow across an electrical grids’ lines is determined *entirely* by the net injections and withdrawals at each node, the susceptance of each transmission line, and the transmission line capacity. The grid operator does not have the ability to adjust or alter the constraints in the short term.

Binding constraints on transmission force deviations from lowest-cost economic dispatch.

---

<sup>2</sup>The inverse of susceptance, *resistance* is sometimes used in power flow modeling

When congestion constraints bind, a node may not be able to acquire electricity from the lowest-cost generator. The marginal cost differential between the lowest cost generator available to serve a given node and the lowest cost generator available at a reference node defines the *congestion cost* of electricity at the given node.

This marginal cost differential is driven by factors beyond the limited frequency of use that “peaker” plants face and congestion. Dynamic constraints on grid operation also restrict the ability of low-cost generation to serve load over short time periods, leading to a steeper supply curve. For instance, fixed startup costs may preclude a lower-cost plant from being dispatched for a short period of time in favor of a more expensive plant that can be dispatched quickly. Or, a lower-cost plant may not be capable (due to design limitations) of ramping quickly to meet demand. Storage may relieve these constraints by providing short-term supply, smoothing demand and allowing the grid optimization to select lower-cost plants. In this manner, storage may affect prices well outside of the periods in which it discharges or is expected to discharge.

Locational market power may also lead to increased convexity of the supply curve, and thus greater price decreases with the introduction of storage. Binding congestion constraints effectively “island” a generator into its own market and provide an incentive for a generator to shade bids upwards when the operator expects to face little or no competition due to transmission constraints (see Mercadal (2015)). To the extent that storage acts as local competition it may reduce prices by limiting potential generator mark-up.

Price decreases owing to storage are estimated in this paper, but no attempt is made to apportion the shares of estimated price decreases amongst relief of dynamic constraints or locational market power. This matter is left for subsequent work, noting that the overall total price effect of storage is the important figure in the context of this paper.

## 2.4 Price determination in two-node grid

For illustrative purposes, consider the two-node grid depicted in Figure 2.4. The grid consists of two nodes, 1 and 2. There is a generator at each node and a transmission line (edge) connecting them. The generator at Node 1,  $G_1$ , has lower marginal cost, and all load is located at Node 2. Suppose load is 110MW, and each generator has sufficient capacity to serve the full load. In an unconstrained scenario, the lowest-cost generator,  $G_1$ , would serve all load. If transmission line capacity is  $\bar{K}_{12} = 100\text{MW}$ , then it is clear that the more costly  $G_2$  generator must be dispatched. Thus, the clearing price at Node 2,  $P_2$  is the price of the expensive generator, while the price at Node 1,  $P_1$ , remains the price of the less expensive

generator. The congestion cost at Node 2 is the price difference.

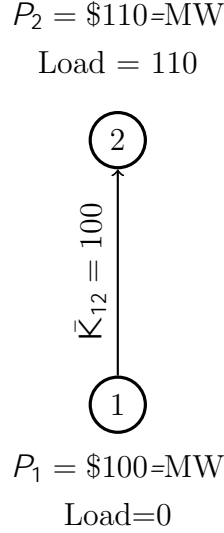


Figure 3: Example of 2-node grid with one line and one potential line constraint. Load is 110MW at Node 2, generation at both Node 1 and Node 2. Note that the lowest-cost electricity at Node 1 will be constrained by the capacity of the line between Node 1 and Node 2, and higher priced generation will be dispatched to meet demand (load) at Node 2.

The optimal power flow in this example can be solved using a Karush-Kuhn-Tucker generalization of the lagrangian to minimize cost of generation subject to total generation and line flow (inequality) constraints. It is assumed for this example that generator constraints do not bind, leaving only two constraints — that the sum of generation must equal demand, and that flow over the line be less than or equal to capacity. Let nodes be denoted as  $n \in \mathcal{N}$ , generation as  $G_n$ , load as  $D_n$ , and capacity as  $\bar{K}$ . Then the cost minimization problem is given by:

$$\mathcal{L} = - \sum_{n=1}^{\mathcal{N}} p_n(G_n)G_n + \underbrace{\left( 0 - \sum_{n=1}^{\mathcal{N}} G_n + \sum_{n=1}^{\mathcal{N}} D_n \right)}_{\text{System-wide power flow constraint}} - \underbrace{\left( \bar{K} - \sum_{n=1}^{\mathcal{N}} [G - D] \right)}_{\text{line flow (slack) constraints}};$$

where  $\mathbf{G}$  is the  $[2 \times 1]$  vector of generation,  $\mathbf{D}$  is the  $[2 \times 1]$  vector of demand, and  $\mathbf{Z}$  is the *shift factor* determined by the susceptances of the lines in the network.  $\mathbf{G}$  is the grid

operator’s choice variable.  $\lambda$  is the familiar KKT (lagrangian) multiplier on the total energy constraint, and  $\mu$  is the KKT multiplier on the line constraint.  $\mathbf{G}$  contains one row for each edge in the network, and one column for each node. Column  $n$  of  $\mathbf{G}$  represents the flow over each edge that will result from the injection of 1 unit of energy at the reference node and the withdrawal of 1 unit of energy from node  $n$ . If Node 1 is the reference node, then  $\mathbf{G} = [0; 1]$ . That is, injecting 1 unit of energy at Node 1 and withdrawing it at Node 2 will increase flow on the line between them by 1 unit. The solution takes the form of optimal (cost minimizing) generator dispatch and associated prices, denoted  $\mathbf{P}$ . It is:

$$\begin{aligned} \mathbf{G} &= [100; 10] \\ \mathbf{P} &= [10; 20]. \end{aligned}$$

Because the objective function minimizes total cost, and because each transmission capacity is included as a constraint, the KKT multiplier  $\mu$  is the incremental reduction in the cost of serving load that would result from relaxing the transmission line constraint by 1 unit. When the constraint is relaxed, it allows one unit of energy to flow from the less expensive generator to the more expensive node, displacing costlier generation there, and lowering the total cost. The shadow value of this constraint is  $\mu = P_1 - P_2 = -10$ . If demand were to decrease to 100MW, or capacity were to increase to 110 MW, the constraint would become “slack”, and  $\mu$  would be equal to zero.

In the simple example, the weight on the transmission constraint is 1, as a unit of energy withdrawn at Node 2 (and injected at Node 1) resulted in an increase in flow on the line of 1. In a more complex network, multiple paths will exist between the withdrawal node and the injection node. Each transmission line will have a constraint which, if binding, will have a corresponding non-zero KKT multiplier. The differential between prices at the withdrawal node and the injection node will be the sum of these corresponding KKT multipliers, weighted by the amount of flow that passes over each line.

If storage is located at Node 2 which becomes congested during peak demand (e.g., in order to defer a transmission upgrade), then transmission constraints bind less frequently with the storage capacity, and the nodal price will be weakly closer to the grid price, even on days where grid-wide demand is high. In this case, storage at the congested Node 2 does not influence the price at Node 1, and prices at Node 1 would be unchanged. If there were a *third* node located in the congested part of the grid near Node 2, served also by the same single, congestable transmission line, then prices at Node 3 would also respond to the introduction

of storage capacity assuming lines between Nodes 2 and 3 are uncongested. Prices at the two nodes in the congested portion of the grid would decline while prices at Node 1 would be unchanged. Because the network is not a perfect lattice where all physical neighbors are connected, these congested parts of the grid could be physically disconnected, but linked in the network. Without knowledge of the physical network structure (as in Chen et al. (2009)), estimating price effects is difficult.

## 2.5 Price determination in 4-node grid

A final example demonstrates price dependencies amid congestion. Consider the 4-node network shown in Figure 2.5. Demand is located at Node D1, and generation is located at nodes G2, G3, and G4, with prices as shown. Capacities are denoted  $\bar{K}$ . The susceptance of each line determines the flow when power is injected at one node and withdrawn at another. With some algebraic manipulation, the susceptances form the shift factor matrix shown in 5, with Node D1 chosen as the reference node<sup>3</sup>.

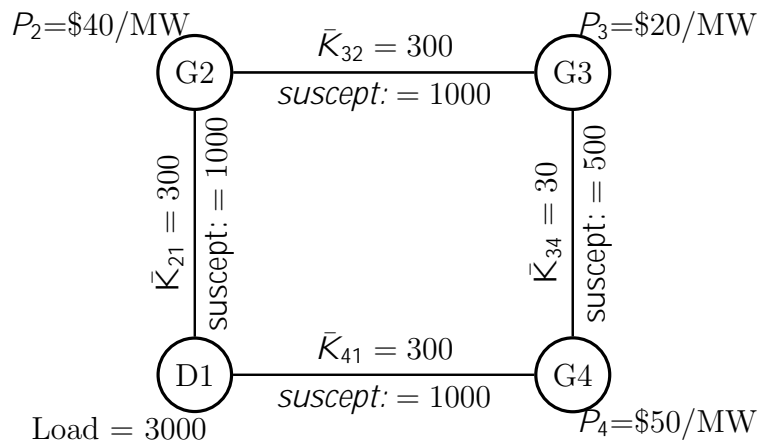


Figure 4: Example of 4-node grid with line constraints. Note that the lowest-cost electricity at  $G3$  will be constrained by the low capacity of the edge  $K_{34}$ , and higher priced generation will be dispatched to meet demand at  $D1$ .

The shift factor matrix for this grid, is derived using a power flow model with Node D1

<sup>3</sup>The choice of which node is the “reference” node is arbitrary. In practice, an increase in withdrawal at one node is met with an increase in supply at another node, or many other nodes, not necessarily the reference node. In this case, the change in the power flow equations is the sum of the effect of withdrawing one unit of energy at the withdrawal node and injecting at the reference node, and withdrawing one unit at the reference node and injecting at the generation node(s)

as the reference node. The power flow model is an algebraic manipulation of the susceptances and edges of the network.<sup>4</sup> Based on the susceptances, the shift factor matrix is:

$$\begin{aligned}
 & \begin{matrix} & 2 & & 3 \\ & 0 & 0.8 & 0.6 & 0.2 \\ \begin{matrix} 6 \\ 6 \\ 6 \\ 4 \\ 0 \end{matrix} & & -0.2 & 0.6 & 0.2 \\ & 0 & 0.2 & 0.4 & -0.25 \\ & 0 & 0.2 & 0.4 & 0.8 \end{matrix} \\
 = & \begin{matrix} & 2 & & 3 \\ & 0 & 0.8 & 0.6 & 0.2 \\ \begin{matrix} 6 \\ 6 \\ 6 \\ 4 \\ 0 \end{matrix} & & -0.2 & 0.6 & 0.2 \\ & 0 & 0.2 & 0.4 & -0.25 \\ & 0 & 0.2 & 0.4 & 0.8 \end{matrix} \tag{5}
 \end{aligned}$$

Each column in  $\mathbf{S}$  represents a node, while each row represents a line. I index lines between nodes by  $e \in E$  where  $E = \{21;32;34;41\}$ . Each column reports the flow of energy across each edge resulting from injecting at the reference node and withdrawing at node  $n \in \{D1;G2;G3;G4\}$ . For instance, 1MW injected at  $D1$  and withdrawn at  $G3$  will increase the flow over  $e_{21}$  by 0.6,  $e_{32}$  by 0.6,  $e_{34}$  by 0.4, and  $e_{41}$  by 0.4 (all units in MW). Similarly, injecting at  $G3$  and withdrawing at  $D1$  yields the opposite —  $e_{21}$  by -0.6,  $e_{32}$  by -0.6,  $e_{34}$  by -0.4, and  $e_{41}$  by -0.4. The lower relative susceptance on  $e_{34}$  causes the lower flow relative to other lines. The shift factors in column 1, the reference node, are zero because injecting 1MW and withdrawing 1MW at the reference node does not increase flow anywhere on the network.

Given the prices and net demand vector  $\mathbf{D} = [300;0;0;0]$ , one can solve a constrained optimization for the generation vector that results in the lowest-cost power flow subject to the line constraints. The solution is:

$$\begin{aligned}
 & \begin{matrix} & 2 & 3 & & 2 & 3 & & 2 & 3 \\ & 0 & & & 45 & & & 0 & \\ \begin{matrix} 6 \\ 6 \\ 6 \\ 4 \end{matrix} & & 750 & & 6 & & 40 & & 6 & & 0 \\ & & & & 4 & & 35 & & 4 & & 25 \\ & & & & 1250 & & 50 & & & & 0 \end{matrix} \\
 \mathbf{G} = & \begin{matrix} & 2 & 3 & & 2 & 3 & & 2 & 3 \\ & 0 & & & 45 & & & 0 & \\ \begin{matrix} 6 \\ 6 \\ 6 \\ 4 \end{matrix} & & 750 & & 6 & & 40 & & 6 & & 0 \\ & & & & 4 & & 35 & & 4 & & 25 \\ & & & & 1250 & & 50 & & & & 0 \end{matrix}; \mathbf{P} = \begin{matrix} & 2 & 3 & & 2 & 3 & & 2 & 3 \\ & 0 & & & 45 & & & 0 & \\ \begin{matrix} 6 \\ 6 \\ 6 \\ 4 \end{matrix} & & 750 & & 6 & & 40 & & 6 & & 0 \\ & & & & 4 & & 35 & & 4 & & 25 \\ & & & & 1250 & & 50 & & & & 0 \end{matrix}; = \begin{matrix} & 2 & 3 \\ & 0 & \\ \begin{matrix} 6 \\ 6 \\ 6 \\ 4 \end{matrix} & & 0 \\ & & & & 4 & & 25 \\ & & & & 1250 & & 0 \end{matrix}
 \end{aligned}$$

The least expensive generator is not serving load at all nodes. In fact, the generator with greatest supply in the solution is the most expensive. The constraint,  $\mathbf{S} \mathbf{G} = \mathbf{D}$ , determines the price difference between the reference node,  $D1$  and the other nodes. For instance, at  $G3$ ,  $0 + 0 + (25 \times 0.4) + 0 = \$10$  less at  $G3$  relative to  $D1$ . *If storage at any node were to eliminate the congestion on  $e_{34}$ , the price at nodes  $G2$ ,  $G3$ , and  $G4$  would decrease.*

Entries in the shift factor matrix  $\mathbf{S}$  may be negative, as is the case here — an increase in injection at some node may *reduce* flow across some lines. This can be particularly beneficial

---

<sup>4</sup>I abstract away from reactive power and angles to illustrate the nature of congestion pricing in a simplified manner.

when this relieves a congested line. Because the shadow value on a constraint is always positive, the negative entry in the shift factor matrix is the source of negative congestion prices. Negative congestion prices signal that withdrawals at the negative-priced node help to relieve a constraint, allowing additional (less expensive) power to flow to other nodes, and, thereby reducing the cost of serving load (i.e. the objective function).

In a more complex network with many nodes and many transmission lines, each with varying susceptances, the calculation of the shadow value is not as straightforward, but the constrained optimization analog still applies. Because transmission line constraints are inequalities, the optimization is a Karush-Kuhn-Tucker (KKT) system of equations. It remains that the congestion cost (i.e. the cost differential between two nodes) is the weighted sum of the shadow values of the transmission constraints from the KKT conditions for grid optimization. The weight for each constraint is determined by the net flow over each transmission line derived from Kirchoff’s Law. Because different nodes in a network will share transmission line paths with each other, the weights (net flows) for each transmission line are integral to understanding how local nodal prices are determined. A formal model of the KKT system is presented in Appendix [A](#).

This relationship between price differences on the grid and transmission constraints is critical to the empirical strategy of this paper. As storage discharges, changes in net load at a node will affect prices at other nodes based on shift factors and transmission constraints. Estimation of price effects requires a specification that (1) accounts for the unobserved, underlying dependence on the shift matrix, and (2) admits the inequality constraints on transmission which frequently equal zero during uncongested times. In brief, the data generating process is a weighted sum of constraints where the constraints are non-linear and frequently zero, and the weights are unobserved. This estimation problem is, at its heart, one of variable selection. When  $\neq 0$  for some edge, storage at one node will affect *some* other nodes, but not all. A natural tool, then, is the LASSO estimator. This estimation is discussed further in Section [3.1](#).

### 3 Data and Methods

Price effects of a marginal unit of storage capacity are separately estimated for the node at which the battery unit is installed and for all other nodes at which prices are affected by the battery unit. This approach is motivated by the unobserved nature of the grid. Cross-node effects depend upon unobserved network links; own node effects are defined by an assumed

link between storage and the most proximal node. The own node effect is estimated in a panel regression that pools all storage units in order to maximize statistical power. Cross-node effects are estimated separately for each node at which storage is installed, sacrificing power for flexible modeling of the network as defined by the LASSO estimator. It would be computationally unfeasible to instead estimate cross-node effects in a pooled regression with each storage node and its interactions entering as independent variables.

### 3.1 Own-node price effects

The per-MW price effect of energy storage is estimated at the node level within the CAISO network using hourly day-ahead prices. These prices are regressed on node-specific, time-varying energy storage capacity, yielding an estimate of the associated effect of storage capacity on prices. Unique coefficients are estimated by hour of the day and by season, admitting distinct charge and discharge behavior. Because electricity consumers do not face prices that instantaneously vary with marginal cost, demand is assumed to be exogenous.

Specifically price effects at the local node are estimated by:

$$\log(\lambda_{nt}) = \log(\lambda_t^{LAP}) + \sum_{s=1}^4 \sum_{h=1}^{24} \beta_{hs} ES_{nt} \times HOUR_h \times SEASON_s + \sum_{n=1}^N \sum_{y=2009}^{2017} \sum_{s=1}^4 \sum_{h=1}^{24} \gamma_{nhsy} \times NODE_n \times YEAR_y \times SEASON_s \times HOUR_h + \epsilon_{nt} \quad (6)$$

where  $\lambda_{nt}$  is hourly (total) marginal price at node  $n$  and time  $t$ ;  $ES_{nt}$  is contemporaneous MW of installed storage capacity at the node;  $NODE_n$ ,  $HOUR_h$ ,  $SEASON_s$ , and  $YEAR_y$  are indicators for each node  $n$ , hour of the day  $h$ , quarter (season) of the year  $s$ , and year of the sample  $y$ , respectively; and  $\lambda_t^{LAP}$  is a weighted average of prices at nodes other than  $n$  within the corresponding utility territory known as a Load Aggregation Point (LAP).<sup>5</sup> Interest centers on  $\beta_{hs}$ , the vector of season and hour-specific coefficients on energy storage capacity that reflects the change in node price due to a one MW change in storage capacity.

Due to the saturation of fixed effects, identifying variation comes from node-level variation in storage capacity across three-month node-hour-season-year intervals. The identifying assumption of equation 6 is that deviations from the node-hour-season-year average price, conditional on LAP prices, are uncorrelated with any omitted variables that correlate with

---

<sup>5</sup>I follow the energy literature and refer to  $\lambda_n$  as the nodal price.



the quantity of energy storage located at the node. Even node-specific trends, for instance, only confound if they introduce dynamics across three-month seasons during which storage is introduced. LAP prices,  $L_t^{AP}$ , are included specifically to control, admittedly in a parametric fashion, for any dynamics that may occur during the season in which storage is commissioned.

The timing and location of energy storage investments are endogenous. There is little information about the siting decisions of storage operators, precluding estimation of the location-choice stage of the adoption decision. Profit maximization suggests arbitrageurs site storage at nodes with large and growing differences in peak and nadir prices. If so, then the price effects estimated in this paper are attenuated. The rich set of fixed effects controls for cross-sectional heterogeneity and secular trends at fairly fine temporal levels, allowing that only node-specific and highly dynamic trends can confound. In contrast to the method of Bushnell and Novan (2018) who control for a variety of potential confounders that affect the grid price of interest and employ only month-hour fixed effects, I use the load aggregation point (LAP) price associated with each node to control for unobserved fluctuations in grid demand. Because the LAP price is a weighted average of all nodal prices (the dependent variable) in an IOU territory, I re-calculate each LAP price excluding the focal node's price. While no single node comprises more than a tiny fraction of the LAP price, this assures that the specification is not mechanically endogenous.

Using this rich set of fixed effects limits identifying variation to a very fine window within the node-hour-season-year. With few exceptions, storage generally increases only once at each node. Therefore, coefficients that vary by season will only capture the effect within the hour-season-year of each increase in storage, while the fixed effects will absorb all variation in subsequent node-hour-season-years. While this improves internal validity, it comes at the cost of understanding the longer run effects of storage.

## 3.2 Cross-Node Price Effects

Estimating the effect of energy storage on congestion pricing at other nodes requires knowledge about the characteristics of the network of generators and transmission lines. Network topology is not published by regulators, utilities, or grid operators because of grid security concerns. Hence, the network is unobservable to the econometrician. In order to recover market-relevant characteristics, I borrow an identification strategy from the Research and Development (R&D) literature that leverages a multi-step Double Pooled LASSO estimator to uncover the structures of unobserved networks.

The initial application of the Double Pooled LASSO estimator in Manresa (2016) was

to identify R&D linkages across technology firms in order to estimate spillover effects. In Manresa (2016), data on firm R&D spending and a LASSO estimator are used to reveal the binary (linked/unlinked) structure of R&D spillovers, and to estimate the magnitude of the spillovers. The LASSO estimator is an intuitive choice in this situation. Since its inception, the LASSO has been used for variable selection and regularization (Tibshirani, 1996), and the binary nature of the R&D linkages are conceptually akin to a variable selection problem.

LASSO is a penalized OLS estimator, usually categorized as a machine learning algorithm.<sup>6</sup> It is commonly used to choose variables and controls when theory has no prediction, or when there are more controls than observations because it is a data-driven procedure that avoids ad hoc variable selection such as stepwise regression. The penalty term on the LASSO is equal to a tuning parameter,  $\lambda$ , multiplied by the sum of the absolute value of all coefficients. The penalty term reduces the magnitude of coefficients and selects variables by reducing the magnitudes of some coefficients to zero<sup>7</sup>.

If one takes any two nodes on the electric grid, then storage at one node will affect price at another if and only if (1) the two nodes share at least one non-zero shift matrix entry, and (2) the transmission constraint on that shared shift matrix entry has a non-zero shadow value (i.e. the constraint is binding) during some hour and season. Because the shift factor matrix and the transmission constraints of the electric grid are unobserved, and because the shadow values of transmission constraints (  $\mu$  in Section 2) are either zero or large, the LASSO estimator is used to uncover the relevant non-zero cross-node effects.

In an electricity grid setting, the unobserved linkages between two nodes are the shared *non-zero* transmission constraints. Each node’s shift matrix column,  $\mathbf{s}_n$  represents the physical path of electricity resulting from one unit of net increased (decreased) flow at that node. The price at that node is the sum of the system-wide energy cost plus the sum of the transmission constraint shadow values (KKT multipliers) (  $\mu$  ) weighted by the shift matrix. Because constraints are zero when they are not binding, and non-zero when they are, any estimation must be able to select which condition holds.

Cross-node effects are estimated by:

---

<sup>6</sup>A distinctly different machine learning approach has also been used in a similar setting by Mercadal (2015) to uncover relevant characteristics of the grid. In this application, machine learning is used to define discrete markets formed by congestion in order to study generator bidding behavior.

<sup>7</sup>The “zeroing” property of the LASSO estimator flows from the absolute value check function.

$$\log(p_{nt}) = \log(p_t^{LAP}) + \sum_{s=1}^4 \sum_{h=1}^{24} \beta_{hs} ES_{nt} + \beta_{HR_t} HR_t + \sum_{s=1}^4 \beta_{SEASON_s} SEASON_s + \beta_n w_{nt} + \epsilon_{nt}; \quad (7)$$

where  $p_{nt}$  is the nodal price,  $p_t^{LAP}$  is the utility-level price as before, and  $w_{nt}$  is the marginal price effect of  $w_{nt}$ , which is comprised of weather covariates, node-level deviations from LAP mean weather covariates, LAP-level load, hydrological potential (the 12-month rolling average of hydropower watershed rainfall), local (within node) solar generation, and a dummy indicator for the Aliso Canyon blowout<sup>8</sup>. To capture a longer window of potential variation attributable to energy storage, I specify only node-hour-season fixed effects plus a year-node fixed effect. This extends the window of identifying variation to the year, rather than than node-hour-season-of-sample.

Variables in  $w_{nt}$  are potential confounders that are included to avoid omitted variables bias. Specifically, if storage is located in an area characterized by more frequent load or temperature spikes, then omitting these variables will attribute their effect to storage, biasing the coefficient of interest. The LASSO selection in Equation 7 presents a unique source of potential omitted variables bias best characterized as model selection omitted variables bias. The source of this bias is rooted in the machine learning nature of the parameter selection. This bias is generated as follows: If two independent variables are correlated with each other, and each has an associated effect on the dependent variable, a LASSO algorithm will select only one, zeroing out the coefficient of the other, despite their correlation. This is because the LASSO objective is parsimonious prediction and not inference. If the dependent variable can be explained with one variable instead of two, then the penalty term selects the more parsimonious specification. In this case, an OLS specification would partition out the explained variance, keeping both variables, provided the two are not perfectly correlated.

By selecting only one of two explanatory variables, LASSO introduces the possibility of an omitted variables bias Belloni et al. (2012, 2014a,b). The solution requires either including any variables in the variable selection stage that also explain the variable of interest (e.g. the treatment variable) derived from a first-stage LASSO procedure, or orthogonalizing the

<sup>8</sup>The Southern California Gas-operated Aliso Canyon natural gas storage facility failed in Fall of 2015, limiting gas storage for Southern California. This supply chain interruption affects the marginal cost of all gas plants in the region, which in turn affects electricity prices in the region.

variables of interest and potential covariates (Belloni et al., 2014a). Following Manresa (2016), I do the latter. Results are robust to either method, and are little changed when following a "naive" single-stage approach.

The Double Pooled LASSO is introduced in Manresa (2016). It performs both selection and estimation of the network effects of energy storage at each node  $ES_{nt}$ , and addresses the model selection omitted variables bias. The objective of the procedure is to estimate an unbiased sparse matrix of cross-node effects of energy storage for every hour-season combination,  $w_{ijhs}$ , where each entry  $w_{ijhs}$  is the price effect at node  $j$  of 1 MW of energy storage at node  $i$  in hour  $h$  and seasons  $s$ .

The Double Pooled LASSO is estimated over three stages. The first two stages generate values of  $w_{nt}$  and  $ES_{nt}$  in Equation 7 that are orthogonal to  $w_{nt}$ , resolving the potential for model selection omitted variables bias. The third stage is the unbiased LASSO estimation.

The first stage regresses  $w_{nt}$  and  $ES_{nt}$  on  $w_{nt}$ :

$$\begin{aligned} w_{nt} &= \beta_n ES_{nt} + \epsilon_{nt} \\ \epsilon_{nt} &= w_{nt} - \beta_n ES_{nt} \end{aligned}$$

and generates  $w_{nt}$  and  $\tilde{w}_{nt}$ , orthogonal to  $ES_{nt}$ :

$$\begin{aligned} w_{nt} &= w_{nt} - \beta_n ES_{nt} \\ \tilde{w}_{nt} &= w_{nt} - \beta_n ES_{nt} \end{aligned}$$

In the second stage, a consistent estimate of  $\beta_n$  is generated using  $w_{nt}$  and  $\tilde{w}_{nt}$ :

$$\tilde{w}_{nt} = \beta_n w_{nt} + \epsilon_{nt}$$

and  $\hat{\beta}_{nt}$ , orthogonal to  $w_{nt}$  is generated:

$$\hat{\beta}_{nt} = \beta_n w_{nt}$$

$\hat{v}_{nt}$  is a consistent estimate of prices that is orthogonal to the portion  $\alpha v_{nt}$  that is uncorrelated with  $ES_{nt}$ . Thus, the effect of  $w_{nt}$  is accounted for, but because the effect of  $w_{nt}$  that is removed is orthogonal to  $ES_{nt}$ , LASSO is not subject to model selection omitted variables bias | the covariance between  $ES_{nt}$  and  $v_{nt}$  remains intact.

Finally, the LASSO estimator is used to estimate  $\beta_{hs}$ , the  $n \times i$  matrix of cross-node effects on price at node  $n$  from storage at node  $i$  for hour  $h$  and seasons. The LASSO procedure estimates each row  $n$ , of the matrix separately:

$$\arg \min_{\beta_{hs}} \sum_{t=1}^T \sum_{i \in n} \beta_{ihs}^2 + \sum_{i \in n} \lambda_{ij} |\beta_{ihs}|;$$

where  $\lambda$  is the LASSO penalty term,  $\lambda_{ij}$  is the penalty loading, and  $\lambda_{ihs}$  is the hour-season specific effect of node  $i$  on node  $n$ , which is assumed to be sparse.

Penalty loadings,  $\lambda_{ij}$  are selected through an iterated estimation process described in Belloni et al. (2012). The penalty weights allow for sharp properties of the LASSO estimator, even in the presence of dependence in errors in a time-series setting (Manresa, 2016). The LASSO is estimated by adapting Christian Hansen's "shooting algorithm"<sup>9</sup>. The penalty term, is calculated via a simulation process.

To remove the shrinkage bias inherent in LASSO feature selection (owing to the penalty term), a final OLS estimation is done for each  $n$  using nodes  $i$  identified as having  $\lambda_{ihs} > 0$ , setting aside all storage nodes designated by LASSO as being zero. Let this set be denoted  $L_{LASSO}$ . This Double Pooled LASSO estimator eliminates shrinkage bias (see Manresa (2016); Belloni et al. (2014b)):

$$v_{nt} = \sum_{s=1}^4 \sum_{h=1}^4 \beta_{hs} ES_{nt} + HR_t + SEASON_t + w_{nt} + \sum_{i \in n} \beta_{ihs} ES_{nt} + HR_t + SEASON_t + \epsilon_{nt} \quad (8)$$

Proofs and methods for estimation are provided by Manresa (2016).

<sup>9</sup><http://faculty.chicagobooth.edu/christian.hansen/research/#Code>

### 3.3 Data

Estimating the effects of energy storage on nodal prices requires three main datasets. First, the nodal prices themselves. Second, the location and commissioning date of energy storage, and finally, data to control for potentially confounding factors that may affect nodal electricity prices, such as temperature, precipitation, or the price of natural gas. This section describes the data sources in detail.

Commercial energy storage capacity in the CAISO area between 2009 and 2017 is observed in the U.S. Department of Energy's (DoE) Global Energy Storage (GS) Database (U.S. Department of Energy, 2016). During the study period, a total of 372MW of energy storage was brought online, equivalent to approximately 1% of the average peak daily load. The DoE GS database contains the commissioning date, size (MWh), capacity (MW), owner, purpose, and geographic coordinates for every unit. I remove any unit whose purpose is related to emergency backup or other non-arbitrage applications, as these are unlikely to engage in regular discharges. This yields 77 unique storage capacity additions to the CAISO service area. Storage installations are concentrated in urban areas, likely to exploit local congestion-related price spikes (see Figure 8). Capacity also exists in rural areas, and some capacity is located proximal to utility-scale renewables.

Figure 5: Energy Storage Capacity Added During Study Period by Type/Source. Totals do not include storage introduced before September 2009. Does not include 120MW of Pumped Hydro taken online in 2012 due to unexpected re. SGIP data on household-level "behind the meter" storage is not used in this study, but is presented for reference.

Figure 6: Energy storage locations in CAISO territory

I obtain Locational Marginal Prices (LMPs),  $\lambda_{nt}$ , for the day ahead market using CAISOs OASIS data portal for the period September 1, 2009 to December 31, 2017<sup>10</sup>As shown in Figure 7, seasonal average LMPs exhibit expected patterns|daily peaks occur during the hours of 5-9pm, and some seasons also exhibit a smaller morning peak. These peaks show the potential for arbitrage using energy storage. I also obtain each of the three Load Aggregation Point (LAP) prices, one for each major investor-owned LSE in California. Figure 8 shows spatial variation in nodal prices for select hours and select seasons. Note that these prices

---

<sup>10</sup><http://oasis.caiso.com>



different from CAISO system price, as nodal prices include congestion whereas the CAISO system price does not.

Figure 7: Hourly Nodal Price (\$), by Season; 95% CI

In some applications, storage capacity is weighted by nodal, hourly energy withdrawals to allow price effects to vary by the magnitude of storage capacity relative to node demand.

Figure 8: Example of spatio-temporal variation in hourly nodal price (\$/MW),  $p_n$ . Note the "hot-spot" of congestion in the Northwest portion of the state which is greatest during the lower-right panel, 6pm on July 20, 2016. Red outlined polygons indicate storage nodes.

Hourly energy withdrawals at each node are imputed as the product of hourly total demand for each utility and respective, static Load Distribution Factors (LDFs) that report the share of utility demand withdrawn at each node. Utility demand and LDFs are reported by CAISO.

I derive the physical location of each CAISO node from OASIS maps using a data scraping algorithm. The scraped location is not exact. Department of Homeland Security standards for information disclosure of "critical infrastructure" generally forbid disseminating information on electricity grid infrastructure. While reported node locations are not precise, they nevertheless appear to correspond to locations at which Google satellite imagery depicts large electricity substations. I assign each energy storage unit in the GS database to the most proximal node. Error in the assignment may introduce bias through measurement error, biasing estimates towards zero. The magnitude of this attenuation is likely to be small, however, because proximal nodal prices are highly correlated. Nodes are clustered by name and proximity to form aggregate clusters of nodes and energy storage. These clusters are called "nodes" throughout this paper. Polygons are generated around each node containing all spatial points that are closest in proximity to that node.

Over each polygon, I process the daily maximum temperature and precipitation data from the PRISM weather dataset (PRISM Climate Group, 2017). Assuming energy consumed in a given area is drawn primarily from the nearest node, a safe assumption given the cost of building transmission lines, maximum local temperature gives a localized measure of nodal demand due to air conditioner use and heating needs. To focus on variation in local nodal net demand that may drive nodal price differences, I calculate a node's daily maximum temperature deviations from the LAP area's average daily maximum temperature, excluding the focal node, and supplement this measure with half-hour temperature measures for the centroid of each polygon as reported by the National Renewable Energy Laboratory's National Solar Radiation Database (NREL NSRDB) (National Renewable Energy Laboratory, 2018). The half-hour temperature measures are recalculated as relative to the LAP area's average as well.

I obtain half-hourly Global Horizontal Irradiance (GHI) measurements for the centroid of each polygon from the National Renewable Energy Laboratory's (NREL) National Solar Radiation Database (NSRDB) (National Renewable Energy Laboratory, 2018). This measure is used to calculate hourly GHI, total daily GHI (summing over each day), and a rolling measure of variation in GHI over a 2-hour window for each polygon. GHI is an input into the estimation of solar generation. Combined with a physical solar model such as NREL's

System Advisory Model (SAM) and data on a system's tilt, azimuth, and inverter, it is possible to convert GHI to electricity generated as in Craig et al. (2018). I use solar insolation (combined with solar capacity installed) only as a proxy for generation. This eliminates the need for perfectly modeling hourly electricity generation at the expense of interpretation of the coefficient on solar insolation.

To create a within-node measure of relative local solar generation, daily solar capacity within each polygon is combined with each polygon's GHI measure and standardized. Locations and installation dates of household solar are extracted from the California Solar Initiative (CSI) database (California Solar Initiative, 2018) and interpolated across each zip code as exact coordinates are not available in the CSI. An hourly time series of total installed solar capacity is generated for each polygon and multiplied by hourly GHI and standardized. The resulting standardized measure represents within-node solar generation.

Utility-scale solar generation dominates distributed generation in California by more than 2:1. Because the focus of this paper is on nodal price deviations from the grid average, and not on the aggregate, system-wide price of electricity, it is important to distinguish between distributed and utility-scale solar. When examining effects on nodal prices, which vary due to congestion, a MW of electricity generated within a node will have a very different effect on nodal prices than a MW of electricity generated at a distant utility-scale solar facility. To distinguish the two, I use the EPA eGrid database to find the location, capacity, and commission date of each utility-scale solar facility in the CAMX eGrid sub-region that contains the CAISO territory. I then merge the half-hourly time series of solar insolation (GHI) from NSRDB for each of the facilities, and generate a measure of utility-scale solar generation in a manner identical to the within-node solar generation measure already described. This allows me to control for the localized effects of distributed generation separately from the system-wide effect of utility-scale generation.

## 4 Effects of Energy Storage on Nodal Prices

### 4.1 Own-node Effects

Point estimates of the  $\beta_{hs}$  in Equation (6) and their 95% confidence intervals are plotted in Figure 9. Errors are clustered at the node. Results for Spring (April-June) are noisy, but point estimates are largely negative during peak (afternoon) periods. When afternoon demand is highest on the CAISO grid, storage effects are statistically significant and equal to approximately -2% with a maximum price effect of -2.2% at 6pm. In Summer (July-

September), when theory posits the price effect will be most pronounced because of high demand and the convexity of the supply curve, price effects are smaller, with a maximum price decline of 0.31% for the hour ending at 8pm. Although the magnitude is smaller during these months relative to Spring, the absolute (dollar) value of the effect is closer in magnitude to Spring due to higher Summer afternoon peak prices.

Coefficients for hours usually considered "off-peak" are statistically insignificant, consistent with markets clearing at a flat portion of supply curves. Point estimates for the early morning hours are small and insignificant but positive in three of four seasons, in line with the theory presented in Section 3.

Figure 9: Coefficient Estimates - Eq 6 with 95% CI, by Season and Hour

Note y-axis re-scales between seasons

Lower hourly nodal prices benefit LSEs and, ultimately, ratepayers by reducing the cost of serving load. To estimate the magnitude of these total savings from a MW storage capacity, denoted  $T_S$ , I multiply the point estimates in Figure 9 with the observed price at each node in each hour of sample, the total load withdrawn at that node, and the total amount of energy storage located at the node:

$$TS = \sum_{d=1}^D \sum_{n=1}^N \sum_{h=1}^H \Delta p_{hs} ES_{dn} \text{LOAD}_{hdn}$$

where  $\Delta p_{hs}$  is the log-price (percent) change estimated in Equation 6, and  $d$  is the day-of-sample. Total load withdrawn at each node is not reported, though total grid load is reported. To calculate node-level load, I multiply total grid load by load distribution factors (LDFs) | the share of total load served, on average, by each node | to calculate hourly load at each node,  $\text{LOAD}_{hdn}$ .

The total savings for the entire study period (2009-2017) is \$50,346,178. These effects are depicted in Figure 10 for 2017. The effects of storage on load-serving costs are greatest in the second and third quarters of the year | not only because point estimates are largest, but also because load is greatest during these hours and seasons. Because each energy storage installation is introduced on a different day of the sample (see Figure 5), I calculate an average total savings per hour, per megawatt of storage ("capacity-hour"). I divide the total benefits of energy storage (\$50,346,178) by the total number of MW of capacity-hours on the grid to generate a per-hour public benefit of \$7.13 per MW of capacity-hour on grid.

Over one year (8,760 hours), the annual own-node benefit of 1MW of energy storage is  $\$7.13 \times 8,760 = \$62,466$ . For perspective, a Tesla Powerwall, the most widely-known brand of behind-the-meter energy storage, has a capacity of 2kW, or .002MW, yielding cost savings from own-node changes of \$124.93 per year.

Figure 10: Aggregated own-node effects of all storage over study period

Another way to compare the ratepayer benefits is to state them as a fraction of the private benefits that can be gained from daily arbitrage of prices. This is of particular importance in understanding the magnitude of ratepayer benefits not captured by private operators. A high amount of unappropriated (public) benefits justifies mandates or subsidies, as are used in California. Arbitrage opportunities decline in storage capacity though ratepayer benefits of storage do not. Hence, internalizing the externality may become increasingly important.

An exact determination of private revenues from arbitrage requires information about daily battery charge and discharge, which is not publicly available data. However, private revenues for one pilot storage project in California are reported in Penna et al. (2016), which estimates private revenues between \$49,500 and \$208,500 per MW of storage per year. Hence, the storage benefits accruing to ratepayers (and not captured by the operator) of \$62,466 amount to 29-126% of these private revenues.

## 4.2 Cross-node Effects

Results from estimating Equation 8 are plotted in Figure 11. Each point represents one non-zero coefficient from  $\beta_{hs}$ , a cross-node price response to a MW of storage at another node linked through the network identified by the LASSO estimator. Confidence intervals are not included for visual clarity, however, nearly all coefficients are significant, a result

not surprising given the LASSO method used to select non-zero effects. In the final stage of estimation, each node is estimated separately. This sacrifices some precision that would be gained from pooling across nodes, but gains flexibility in estimating  $\beta_{hs}$ . Appendix X contains plots with standard errors clustered at the node-week level.

Figure 11: Cross-node effects. Each point represents an entry in  $\beta_{hs}$ , and is the coefficient representing the change in one node's log price at hour  $h$  seasons, resulting from a 1MW increase in storage at another node. Points are plotted with 20% transparency to show masses.

The summer season (Jul-Sep), hypothesized to be the period when congestion costs would be highest and thus storage would exhibit the greatest effect, contains the most non-zero coefficients. A mass of negative coefficients is observed in the early afternoon hours (3pm to 5pm) during Jul-Sep. Furthermore, during the off-peak times (e.g. 10-12pm; 2-5am), many coefficients are positive, though this represents a smaller positive change (on a lower quantity) than the afternoon on-peak effects. This is consistent with the effects hypothesized in Section 2. Some negative effects are observed during evening in October-December, and during the morning peak hours during Jan-Mar. Note that Figure 7 shows a second daily peak occurring during the morning in January-March. This morning price peak occurs as households rise to prepare for their day during the colder, darker Winter mornings.



Nodes designated by the LASSO estimation as linked during some hour and season ( $\beta_{inhs} \neq 0$ ) may not necessarily be spatially contiguous as transmission and distribution infrastructure is not a perfect lattice. However, linked nodes should be in proximity; it is unlikely that a node in San Diego influences at the northern border of California. Selected linked nodes are depicted in Figures. As these figures show, linked nodes are spatially contiguous or proximal, lending validity to the LASSO selection procedure.

Figure 12: Cross-node effects for storage located near Oroville, CA (yellow polygon)

Figure 13: Cross-node effects for storage located near Escondido, CA (yellow polygon)

These cross-node results indicate that price effects of energy storage extend beyond the capacity node to other nodes in the network. The sign of these effects is predominantly negative. When aggregated in a manner similar to the own-node effects in Figure 10, cross-node effects generate total savings (e.g. decreases in cost of serving load) of approximately \$1,170,026, or around 2.5% of the total benefits realized at the own-node. This is approximately 1% to 3% of private revenues. Cross node effects are heterogeneous across the 77 storage installations. In some areas, effects appear to be widespread across a region as in Figure 13. In others, effects are concentrated in the area around the location of storage as in Figure 12. This heterogeneity highlights the importance of considering the local network topology and congestion patterns in calculating the benefits of energy storage to ratepayers.

## 5 Discussion

Ratepayer benefits from own-node price responses are estimated to be \$62,466 per year, per MW of capacity, and cross-node benefits are 2.5% of that amount, \$1,561, totalling \$64,028 per MW of capacity, per year. Per hour on the grid, the average ratepayer benefit is \$7.13. While the price of battery installations is falling rapidly, estimates from 2016 for the full installation and connection cost in California are around \$6.5M per installed MW of storage capacity. Arbitraging energy prices in CAISO energy or participating in the ancillary services market generates private revenues of \$45,500 to \$208,500 per year, per MW (Penna et al., 2016).

Standing on their own or in tandem, the total public and private benefits from energy storage only partially justify the estimated cost of storage in 2016. Mandates such as AB 2514 can still be justified by learning spillovers and first-mover advantage, and California has historically been a first-mover in renewable and low-carbon energy policy. This paper estimates only the ratepayer benefits, draws assumptions on private benefits and storage cost from published reports, and does not seek to quantify other tangible economic benefits. This is largely because costs of batteries are declining rapidly, with some reports of large-scale installations dipping below \$1,000,000 per MW, less than one-sixth of the reported 2016 price (Lambert, 2018).

### 5.1 Ratepayer Benefits Relative to Private Benefits

It is important to consider the economic incentives present in a low-cost storage world. If storage acts as any other generation asset, then investment in storage will be driven by the private marginal benefits. Private investors will not consider the ratepayer benefits that cannot be captured. In fact, higher ratepayer benefits will decrease private revenues as the magnitude of the diurnal arbitrage opportunities (peak prices) are reduced. Therefore, a public goods problem arises, and a policy intervention is necessary to ensure that the ratepayer benefits are accounted for in the investment decision.

To this end, I calculate the ratepayer benefits as a fraction of the total private benefits from arbitrage. If ratepayer benefits are large relative to private benefits, then policy intervention in the form of subsidies is justified. If the benefits are small relative to private benefits, then the unsubsidized outcome will be close to the efficient outcome, and no policy is necessary.

Results show that public ratepayer benefits are a significant portion of benefits from

storage. Own-node effects total 29-126% of the total private benefits from arbitrage, and cross-node effects total another 1-4%, summing to 30%-130%. Therefore, ratepayer benefits are a considerable portion of total overall benefits from storage operation, possibly even exceeding private benefits. These benefits, not captured by private storage operators, justify policy if broad utility regulation does not achieve these investments. However, as storage is added to the grid, the price effects will reduce the private benefits of each subsequent unit as arbitrage opportunities decline. An anticipated decline in such opportunities may dissuade potential entrants from making capacity investments, exacerbating the under-provision of storage capacity. In markets fully saturated with storage capacity, hourly prices would be invariant across hours of the day as all deviations from an average are arbitrated away. However, this also removes any incentive for adding storage. Thus, policy is likely to play an important role in the storage market. The dynamic aspects of storage investment are not examined here, but are left for subsequent research. Clearly, there exists a per-MW cost of storage where private benefits alone do not justify investment, but the sum of private and public ratepayer benefits do justify investment.

The magnitude of public ratepayer benefits relative to private benefits highlight the importance of rate regulation as well. Under pass-through utility regulation, savings in the cost of serving load are passed on to ratepayers. This paper does not examine the underlying assumption, but acknowledges that the intent of a policy and the implementation of that policy may differ dramatically. In California, the California Public Utilities Commission examines all rate change requests, and receives input from ratepayer advocates in determining its approval of a rate change. When storage is necessary to avoid or delay a local distribution line upgrade (i.e. to lower costs associated with local congestion costs), the utility-borne costs may be included in a ratepayer's billed distribution charge (Bierman, 2018). Under these conditions, it is important to understand the ratepayer benefits, as the ratepayer is bearing the capital costs.

## 5.2 Relevance to Other Grids

California has historically been a first-mover in renewable and low-carbon energy policy, and it retains this position in mandating storage under AB2514. California's experience with energy storage provides insight for other states and grid operators that may consider the role of storage in grid operations. The effects measured in this paper pertain to local nodal prices relative to a systemwide average LAP; therefore, grids with little congestion may be unlikely to see similar effects, even with a proportionally equivalent amount of storage.

Historically, California has seen higher than average congestion costs, with the New York ISO (NYISO) having the highest average ratepayer cost of congestion (Lesieutre and Eto, 2003). Comparison across grids is hampered by variation in how congestion is priced and measured. However, California's congestion costs are driven in part by internal, branch group (e.g. intra-distribution level) congestion, or congestion resulting from insufficient capacity on the smaller distribution grid, rather than exclusively by constraints on larger high-voltage regional transmission lines (DOE, 2015). For example, the spatial distribution of "hot spots" in Figure 8 illustrates this issue in the Northwest corner of the state during Summer peak hours. Figure 8 also illustrates the larger transmission-level congestion constraints | the lower-right panel shows a clear gradient between the Southern and Central regions around 34 N. This transmission-level congestion has persisted for years in the CAISO. Note that, in this figure, the local "hot-spot" in the Northwest corner is of greater magnitude than the longstanding South-Central gradient, although the Northwest "hot-spot" is in far fewer nodes, and covers nodes with much lower total volume. For 2016, the CAISO-estimated total cost of congestion was \$99 million (CAISO, 2016). This stands in contrast to ERCOT, the Texas grid, where transmission constraints between wind-generation rich areas in the Northwest and load centers in the East drive congestion costs (LaRiviere and Lu, 2017).

### 5.3 Does Storage Facilitate Renewables?

This paper demonstrates the role of storage in "facilitating integration of renewables", the goal of AB2514. Results indicate that storage reduces local nodal prices at the times in which local solar generally imposes high costs, i.e., late-afternoon hours during in spring and summer. Solar generation has two observable price effects throughout the day. First, it decreases the grid-level cost of electricity during the afternoon hours as generation is high relative to demand, which has the effect of depressing prices due to low net load (demand net of zero-marginal-cost renewables). Second, during both the morning ramp-up of solar generation and, especially, the evening ramp-down as the sun sets, more expensive fossil-fuel generation is necessary to accommodate the fast ramp. Bushnell and Novan (2018) quantify these effects, and identify more expensive Gas Turbine generation as responding to the evening ramp in lieu of the more efficient Combined Cycle Gas Turbine fleet. Therefore, over the day, solar reduces total costs, but accommodating the morning and evening ramp requires more expensive generation, increasing prices during those hours.

Energy storage reduces local congestion costs during these critical hours, as shown in Figures 9 and 11. Thus, storage appears to facilitate integration of solar on the grid. By

reducing local congestion, storage also makes more quickly-dispatched plants available to follow changes in solar generation, potentially alleviating system-wide cost increases associated with solar capacity (Bushnell and Novan 2018; Bierman 2018). Effects are not limited to only those days in which solar generation is highest (and thus the evening ramp is steepest). Additional analysis indicates that storage's effect does not interact with daily solar generation | the price effect is consistently present regardless of solar generation. This follows the theory in Section 2, where price effects are timed to peak prices, regardless of whether the driver of peak prices are normal demand patterns, or the steep ramp caused by solar. Storage addresses price increases during the same hours solar generates those price increases, but storage also addresses price increases when at those times, even when solar is not the driving force.

As California moves towards a zero-carbon grid and implements its Title 24 mandate of rooftop solar on all new residential construction (California Energy Commission, 2018), the additional cost of the evening solar ramp-down will likely increase in magnitude and will represent a significant portion of the cost of serving load. I show that storage operates to reduce prices in the hours most affected by the evening solar ramp-down. Hence, storage plays an important role in "facilitating integration of renewables" to the CAISO grid.

## 6 Conclusion

This research provides the first empirical estimates of the price effect of energy storage. Panel data estimates show a significant downward effect on nodal prices during peak hours and during peak summer months, suggesting that energy storage has a measurable effect on grid operations. Summer afternoon peaks coincide with times of highest demand for expensive "peaking" generation, and are the costliest demand periods for which grid operators must plan. The effect during spring evening hours reaches as much as -2.2%, and summer months reaches up to -.31%. These price effects generate an average \$62,466 in cost savings per year, per MW to load-serving entities like utilities that are passed on to consumers through typical regulation. The 372MW of energy storage capacity in California is estimated to have reduced energy costs across the grid by \$50.35M, but this amount is not sufficient to fully justify mandates. Battery operators are not compensated for these cost savings to utilities and their rate payers. These public ratepayer benefits constitute a substantial fraction of the private benefits from energy price arbitrage and point to a public interest in storage capacity expansion in the future, when private investment is likely to be under-provided.

## References

- A. Belloni, D. Chen, V. Chernozhukov, and C. Hansen. Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369{2429, 2012. ISSN 1468-0262.
- A. Belloni, V. Chernozhukov, and C. Hansen. Inference on treatment effects after selection among high-dimensional controls *The Review of Economic Studies*, 81(2):608{650, 2014a. ISSN 0034-6527.
- A. Belloni, V. Chernozhukov, and C. Hansen. High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29{50, 2014b. ISSN 0895-3309.
- E. M. Bierman. Direct Testimony of Evan M. Bierman On Behalf of San Diego Gas and Electric Company Before the Public Utilities Commission of the State of California, 2018.
- R. E. Bohn, M. C. Caramanis, and F. C. Schweppe. Optimal pricing in electrical networks over space and time. *The Rand Journal of Economics*, pages 360{376, 1984. ISSN 0741-6261.
- K. Bradbury, L. Pratson, and D. Patiño-Echeverri. Economic viability of energy storage systems based on price arbitrage potential in real-time U.S. electricity markets *Applied Energy*, 114:512{519, 2014. ISSN 03062619. doi: 10.1016/j.apenergy.2013.10.010. URL <http://www.sciencedirect.com/science/article/pii/S0306261913008301>.
- J. Bushnell and K. Novan. Setting with the Sun: The impacts of renewable energy on wholesale power markets. 2018.
- CAISO. Annual Report on Market Issues and Performance. 2016. URL [http://www.caiso.com/Documents/2014AnnualReport\\_MarketIssues\\_Performance.pdf](http://www.caiso.com/Documents/2014AnnualReport_MarketIssues_Performance.pdf).
- California Energy Commission. Energy Commission Adopts Standards Requiring Solar Systems for New Homes, First in Nation. Technical report, State of California, Sacramento, CA, 2018. URL [https://www.energy.ca.gov/releases/2018\\_releases/2018-05-09\\_building\\_standards\\_adopted\\_nr.html?platform=hootsuite](https://www.energy.ca.gov/releases/2018_releases/2018-05-09_building_standards_adopted_nr.html?platform=hootsuite).
- California Solar Initiative. CSI Database, 2018. URL [https://www.californiasolarstatistics.ca.gov/archived\\_working\\_data\\_files/](https://www.californiasolarstatistics.ca.gov/archived_working_data_files/).

- J. Chen, M. Macauley, and A. Marathe. Network topology and locational market power. *Computational Economics*, 34(1):21–35, 2009. ISSN 0927-7099.
- M. T. Craig, P. Jaramillo, B.-M. Hodge, N. J. Williams, and E. Severnini. A retrospective analysis of the market price response to distributed photovoltaic generation in California. *Energy Policy*, 121:394–403, 10 2018. ISSN 0301-4215. doi: 10.1016/J.ENPOL.2018.05.061. URL <https://www.sciencedirect.com/science/article/pii/S0301421518303781#>.
- U. S. DOE. National electric transmission congestion study, 2015.
- R. Fioravanti, M. Kleinber, W. Katzenstein, S. Lahiri, N. Ton, A. Abrams, J. Harrison, and C. Vartanian. Energy Storage Cost-effectiveness Methodology and Results. Technical report, DNV KEMA Energy and Sustainability, Oakland, CA, 2013.
- GTM Research. US Energy Storage Monitor: 2016 Year in Review and Q1 2017 Executive Summary. Technical report, Energy Storage Association, 2017.
- E. Hittinger, J. F. Whitacre, and J. Apt. What properties of grid energy storage are most valuable? *Journal of Power Sources*, 206:436–449, 2012. ISSN 0378-7753.
- F. Lambert. Teslas massive Powerpack battery in Australia cost \$66 million and already made up to ~\$17 million. *Electrek*, 9 2018.
- J. LaRiviere and X. Lu. Transmission Constraints, Intermittent Renewables and Welfare. 2017.
- B. C. Lesieutre and J. H. Eto. Electricity transmission congestion costs: A review of recent reports. 2003.
- R. Lueken and J. Apt. The effects of bulk electricity storage on the PJM market. *Energy Systems*, 5(4):677–704, 2014. ISSN 1868-3967.
- E. Manresa. Estimating the structure of social interactions using panel data. *Unpublished Manuscript. MIT Sloan*, 2016.
- I. Mercadal. Dynamic competition and arbitrage in electricity markets: The role of financial players. Technical report, 2015.



- National Renewable Energy Laboratory. National Solar Radiation Database (NSRDB), 2018. URL <https://nsrdb.nrel.gov/nsrdb-viewer>.
- A. Nottrott, J. Kleissl, and B. Washom. Energy dispatch schedule optimization and cost benefit analysis for grid-connected, photovoltaic-battery storage systems. *Renewable Energy*, 55:230–240, 2013. ISSN 0960-1481.
- K. Novan. Valuing the wind: renewable energy policies and air pollution avoided. *American Economic Journal: Economic Policy*, 7(3):291–326, 2015. ISSN 1945-7731.
- M. D. Penna, M. Yeung, and D. Fribush. Electric Program Investment Charge (EPIC) - Energy Storage for Market Operations. Technical report, Pacific Gas and Electric, Electric Asset Management Department, San Francisco, CA, 2016. URL [https://www.pge.com/pge\\_global/common/pdfs/about-pge/environment/what-we-are-doing/electric-program-investment-charge/PGE-EPIC-Project-1.01.pdf](https://www.pge.com/pge_global/common/pdfs/about-pge/environment/what-we-are-doing/electric-program-investment-charge/PGE-EPIC-Project-1.01.pdf).
- PRISM Climate Group. PRISM Gridded Climate Data, 2017. URL <http://prism.oregonstate.edu/>.
- R. Sioshansi, P. Denholm, T. Jenkin, and J. Weiss. Estimating the value of electricity storage in PJM: Arbitrage and some welfare effects. *Energy economics*, 31(2):269–277, 2009. ISSN 0140-9883.
- R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996. ISSN 0035-9246.
- U.S. Department of Energy. Global Energy Storage Database, 2016. URL <http://www.energystorageexchange.org/projects>.
- U.S. Department of Energy. Office of Energy Projects Energy Infrastructure Update Office of Energy Projects Energy Infrastructure Update For October 2012. Technical Report December, 2017. URL <https://www.ferc.gov/legal/staff-reports/2017/dec-energy-infrastructure.pdf>.
- R. Walawalkar, J. Apt, and R. Mancini. Economics of electric energy storage for energy arbitrage and regulation in New York. *Energy Policy*, 35(4):2558–2568, 2007. ISSN 0301-4215.

- R. Walawalkar, S. Blumsack, J. Apt, and S. Fernands. An economic welfare analysis of demand response in the PJM electricity market. *Energy Policy*, 36(10):3692–3702, 2008. ISSN 0301-4215.
- F. A. Wolak. Level versus Variability Trade-offs in Wind and Solar Generation Investments: The Case of California. Technical report, 2016.
- F. A. Wolak. The Evidence from California on the Economic Impact of Inefficient Distribution Network Pricing. Technical report, 2018.
- C.-K. Woo, I. Horowitz, J. Moore, and A. Pacheco. The impact of wind generation on the electricity spot-market price level and variance: The Texas experience. *Energy Policy*, 39(7):3939–3944, 2011. ISSN 0301-4215.
- C.-K. Woo, J. Moore, B. Schneiderman, T. Ho, A. Olson, L. Alagappan, K. Chawla, N. Toyama, and J. Zarnikau. Merit-order effects of renewable energy and price divergence in Californias day-ahead and real-time electricity markets. *Energy Policy*, 92:299–312, 2016. ISSN 0301-4215.

## Appendix A Formal model of grid pricing

Let the congestion portion of the nodal energy constraint be denoted as  $c$ , with elements  $c_n$ , and the total nodal energy constraint (the nodal price) be denoted as  $\pi_n$  with elements  $\pi_n$ . The KKT shadow value for transmission congestion is  $\lambda$ , with elements  $\lambda_e$ . For a network with  $E$  edges and  $N$  nodes,  $\lambda \in \mathbb{R}^E$  and  $c \in \mathbb{R}^N$ :

$$\begin{aligned} \pi_n &= \lambda_e + c_n \\ &= \lambda_e + \kappa_{en}; \end{aligned} \tag{9}$$

where  $\lambda_e$  is the marginal cost of energy at the reference node,  $c$  is the congestion component, and  $\pi_n$  is the  $n$ th column of the shift factor matrix derived directly from the physical resistance of each edge (transmission line) in the network. The  $\{1; \dots; E\}$  elements of  $\pi_n$  represent the energy flow across each edge  $e$  from the reference node to node  $n$ . They are not constrained to positive numbers, though the value of the elements of  $\pi$  must be greater than zero. Therefore, it is possible for  $c_n < 0$ . This can occur at a node that has low-cost energy

accessible to it, but congestion keeps that low-cost energy from flowing to the reference node. If one increases the withdrawal at a negative congestion cost node ( $c_n < 0$ ), and increases the injection at the reference node (which was unable to be served by the low-cost generator due to congestion), then (1) the total cost of serving demand has decreased as the cheaper generator is being used, and (2) the net flow on the congested line has decreased.

Congestion is priced at every node in a network but can be affected by withdrawals or injections at *any* node, especially those that are grid “neighbors”. As a result, the effect of energy storage does not follow a congestion analog to equation. 3— the discharge of energy at peak times may *increase* the nodal price, or may cause another previously uncongested transmission line to become the binding constraint, raising nodal prices elsewhere. However, decreases in demand during peak periods will, over the entire grid, strictly decrease total congestion costs. The magnitude and location of price changes due to energy storage are determined by network topology, which is not publicly available, and, therefore, unobserved to the econometrician. Changes in network congestion costs affect transfers to LSEs and ratepayers.

The vector  $c$  represents the congestion component of the cost of injecting one unit of energy at the reference node and withdrawing one additional unit of electricity at any node  $n$ . The price paid for *all* units of energy at node  $n$  includes  $c_n$ . Interest centers on the effect of a change in nodal demand on congestion costs:  $\frac{dc}{dQ_n}$ . A change in demand at node  $n$ , however, may change congestion costs at *all* nodes by changing the transmission constraint shadow value vector,  $\lambda$ . Furthermore, the change in  $c$  is itself dependent on the vector of net demand,  $Q$ :

$$\begin{aligned} \frac{dc}{dQ_n} &= \frac{dc^e}{dQ_n} + \frac{dc(Q)}{dQ_n} \\ &= \frac{dc^e}{dQ_n} + \sum_{e=1}^E \frac{dc_e(Q)}{dQ_n} \\ &= \frac{dc^e}{dQ_n} + \sum_{e=1}^E \frac{dc_e(Q)}{dQ_n} \sum_{e'=1}^E \frac{dc_{e'}(Q)}{dQ_n} \sum_{e''=1}^E \dots \sum_{e=1}^E \frac{dc_e(Q)}{dQ_n} \sum_{Ne}^{\#_0} ; \end{aligned} \quad (10)$$

where the first equality follows from the definition in 9 and that  $c^e$  is a constant. The second equality states that the change in each node’s congestion cost,  $dc_n$ , is the sum of the changes in each of the  $E$  transmission constraints,  $dc_e$ , weighted by the Kirchoff’s Law-derived flows between the reference node and  $n$ . The KKT slack conditions on  $\lambda$  provide intuition for

the difference between peak period and nadir period price effects. When a constraint is not binding on edge  $e$ , slack conditions require  $\lambda_e = 0$ . If energy storage at node  $n$  is charged during the nadir, and transmission constraints are likely to be slack at this time, then the probability distribution of price effects will have mass at zero (e.g. charging during off-peak times will not increase congestion costs).

Amid congestion, interest centers on storage effects on both *own-node* congestion prices,  $\frac{d c_n}{d Q_n}$  and *cross-node* congestion prices,  $\frac{d c}{d Q_n}$ . These determine the total congestion price effect of discharging stored energy at node  $n$  during high demand  $Q^H$  recharging storage capacity during low demand,  $Q^L$ . A change in  $Q_n$  will result in a change in generation at *some unknown node(s)*, further complicating estimation of the total effect. The total change in congestion cost per unit of energy storage at node  $n$  may be written as:

$$\text{Total Change} \approx -s \times \frac{d c(Q^H)}{d Q_n} \times Q^H + s \times \frac{d c(Q^L)}{d Q_n} \times Q^L \quad (11)$$

The first term in parenthesis is the energy storage transaction  $-s$  at node  $n$  multiplied by the price effect at each node from a change in  $Q_n$ , multiplied by the total demand at each node during the peak period. The second term is identical, but for an increase in withdrawal of  $s$  during the nadir.

In the previous sections, peak and nadir times are summarized by  $\{H; L\}$ . Operation of energy storage generally extends beyond a single hour. Storage is charged (or discharged) over 1-4-hour periods. Storage operated to shave peak demand or engage in energy arbitrage will discharge during a node's peak period. Because nodal prices are unique to each node and determined as solutions to a complex non-linear optimization problem, the peak and nadir periods may occur at different times for different nodes. Therefore, storage price effects may be measured over multiple hours that differ from node to node. Battery operations are not observed, so total daily benefits of a marginal unit of storage capacity are calculated as the sum of price changes over *all* hours of the day:

$$\text{Total Benefits} = s \times \sum_{t=1}^{24} \frac{d p_t(Q_t)}{d ES} Q_t \quad (12)$$