Energy Syst
DOI 10.1007/s12667-017-0236-x

ORIGINAL PAPER

The economics of commercial demand response for spinning reserve

Michael Fisher[1] · **Jay Apt**[1,2] · **Fallaw Sowell**[2]

Received: 5 January 2017 / Accepted: 27 February 2017

Abstract Demand response (DR) for spinning reserve may be appropriate for customers whose operational constraints preclude participation in energy and capacity DR programs. We investigate the private business case of an aggregator providing spinning reserve in California across customer end uses and business segments. Revenues are calculated using end use level hourly load profiles. With average annual revenue of ∼\$35/kW, steady end uses (e.g., lighting) are more than twice as profitable as seasonal end uses (e.g., cooling) because spinning reserve is needed year-round. Business segments with longer operating hours, such as groceries or lodging, have more revenue potential. Total costs for participation would need to be under \$250/kW for many end uses and business segments to have payback periods less than 5 years, which is plausible given equipment cost data from California's Automated Demand Response programs. Avoided carbon emission damages from using DR instead of fossil fuel generation for spinning reserve could justify incentives for DR resources.

Keywords Demand response · Spinning reserve · Carbon reduction

✉ Jay Apt
apt@cmu.edu

Michael Fisher
mjfisher@cmu.edu

Fallaw Sowell
fs0v@andrew.cmu.edu

1 Department of Engineering and Public Policy, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

2 Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Published online: 17 March 2017

Abbreviations

AB32	Legislative act creating California's carbon cap-and-trade system
ACZ	Ancillary service zone
ARMA	Autoregressive moving average
ARRA	American recovery and reinvestment act
AutoDR	Automated demand response
BEMS	Building energy management system
BIC	Bayesian information criteria
CAISO	California independent system operator
CEUS	California Commercial End Use Survey
CO_2	Carbon dioxide
DR	Demand response
FCZ	Forecasting climate zone
kW	Kilowatt
MW	Megawatt
NGCC	Natural gas combined-cycle power plant
NGCT	Natural gas combustion turbine power plant
NP26/SP26	North/south of transmission line path 26
PG&E	Pacific Gas and Electric
SCE	Southern California Edison
SCC	Social cost of carbon
WECC	Western electricity coordinating council

1 Introduction

Load that can respond to price or reliability signals, referred to as "demand response" (DR), lowers energy demand during periods of high prices or the need for generation capacity during periods of high load [1]; grid operators are now exploring the use of DR for ancillary services [2–5].

One ancillary service is spinning reserve. This type of reserve is also referred to as synchronous reserve and is often considered under the umbrella of contingency reserves, which include spinning and non-spinning reserve. Spinning reserves have traditionally been generators running at idle power and synchronized to the phase of the 50 or 60 Hz grid; they are able to provide rapid increases in power in response to an unexpected contingency event (e.g., loss of a transmission line or generating facility) [6]. The operational requirements vary across jurisdictions, but generally require the ability to increase generation in a short time, typically 10 min [7], and to maintain that response for a minimum amount of time (typically 30 to 60 min) [8].

The intrinsic characteristics of DR are a natural match to the requirements of spinning reserve resources. Load resources can provide higher ramp rates [3,9] at lower costs [10] than traditional generation. Furthermore, a large number of loads that are individually less reliable than a generator may provide aggregate reliability in excess of that provided by a few large generators [2,3]. The timescale on which spinning reserve operates is well served by DR because the average event lasts only 10–20 min

[3]. Moreover, this short period is attractive to DR participants because it avoids customer fatigue and business operations changes required by the 1 to 8 hour interruptions [11] seen in energy or capacity events.

Wholesale markets in the US Mid-Atlantic, New York, Texas, and the Mid-West all allow DR to participate in spinning reserve markets. However, current Western Electricity Coordinating Council's (WECC) rules implicitly prevent the California Independent System Operator (CAISO) from allowing DR in the spinning reserve market, but this is a purely regulatory barrier. WECC rules require immediate and automatic response to system frequency to participate in spinning reserve [12] while DR typically requires a signal from an outside operator to initiate response. For our analysis we assume this barrier is removed and regulators permit DR in the wholesale environment.

The open question is whether market prices are sufficient to attract participation given time-varying resource availability and the magnitude of implementation costs. Previous studies have examined the use of DR for ancillary services and the economics of participation. Kirby [2] and Mathieu et al. [13] consider residential air-conditioning loads in New York and California, respectively. They characterized resource size and calculated potential revenue assuming time-invariant resource availability. MacDonald et al. [14] reviewed market clearing prices and participation requirements across the U.S., though they do not discuss potential resource revenue and assume the demand resource is time-invariant. MacDonald et al. [15] examined commercial building HVAC and lighting loads but did not discuss implementation costs or match time-varying resource availability with market clearing prices. Ma et al. [16] and Hummon et al. [17] examined the market dynamics of the western interconnection using unit commitment and economic dispatch models with increased flexible demand resources for energy and ancillary services. They did not consider the costs to enable DR for these services.

To our knowledge, no previous research has compared the costs and potential revenues of using DR for ancillary services while capturing the time-varying nature of resource availability across many end uses and customer segments. The 2009 PG&E Participating Load Pilot [18] implemented DR for non-spinning reserve, and thus faced the true operational costs and potential revenues, but included only 3 participants in the study. We take a more comprehensive view using data from over 2700 buildings in California. We examine the economics across geographic regions, building segments, and end uses within California using econometric models. California is used as an example because its varied load types and competitive market operations provide an ideal environment in which to examine the business case for DR and because the results may influence DR policy in WECC. We examine DR for commercial load, which represents approximately 50% of California's load [19]. This work adds to the existing literature by determining which commercial demand response applications are both profitable and significant to the grid, and making a first-order estimate of their environmental consequences.

We focus on the case where DR participates solely in spinning reserve (not in energy or capacity). Customers may want to participate only in spinning reserve because of the low frequency and short duration of events. Indeed, customers accounting for approximately 50% of the MW signed up through the California Automated Demand

Response program participate in a voluntary energy reduction program [20]. This suggests that these customers do not find the mandatory energy curtailment required by capacity events attractive. This work does not discuss frequency regulation (another ancillary service) because this application of commercial DR remains largely in its infancy [9] and the installation costs are highly scenario specific.

There is a growing body of literature on the optimal control of demand-side resources in market and microgrid environments [21]. Here we assume the control algorithms and equipment are sufficient to achieve the load reductions determined by our models and instead focus our analysis on the resource and the economics.

We find that steady end uses (e.g., lighting) are better able to make a profit than are seasonal end uses (e.g., cooling) because, unlike a capacity resource, spinning reserve is needed throughout the year. Payback periods of 5 years or less are plausible in certain niche applications given data on equipment costs, but longer paybacks for many resources may discourage widespread participation. Therefore, we investigate if the damages from carbon emissions avoided by procuring DR in spinning reserve are sufficient to justify monetary incentives to encourage greater DR participation.

Section 2 describes our methods and data used to characterize the implementation costs and calculate potential revenue. Section 3 presents and discusses the results of our analysis. Section 4 estimates avoided carbon emissions damages by using DR for spinning reserve and Sect. 5 presents our conclusions.

2 Methods and data

We consider a DR aggregator who contracts with individual facilities to procure DR. These facilities receive compensation for agreeing to reduce load when called upon. In turn, the aggregator sells the cumulative DR capability to a utility or grid operator. We take the perspective of an aggregator, not an individual facility owner, because aggregators are more likely than individual facilities to have the resources necessary for sophisticated forecasting models and the complex administrative requirements necessary to participate in these markets.

Aggregators are most likely to target large commercial participants. Overhead costs are lower for these customers as administrative and marketing costs often scale per customer rather than per kW. Large customers are also more likely to participate in DR programs [22] and have the internal building controls required for automated response.

Aggregators earn revenue based on the market clearing price and magnitude of load response, and incur costs to enable spinning reserve in participant facilities. Revenue is calculated by matching hourly DR resource availability with market clearing prices across geographic zones, building segments, and end uses. Detailed cost data are not available at the end-use or business segment level; we therefore treat costs parametrically to determine the level at which acceptable payback periods are achieved. We compare these cost levels to general cost estimates from the literature and from a cost database for an automated DR program in California.

We do not model the effects of a call for spinning reserve on energy cost. This eliminates the uncertainty inherent in modelling events with probabilistic frequency and duration. A first-order analysis shows that we are ignoring less than $5/kW-year in potential revenue gains from energy reductions, which would not affect the conclusions of our work. Consider the case of an end use with no energy rebound after a spinning reserve event (e.g. lighting). End uses with energy rebound (e.g. cooling) will have less change in their total energy consumption. Assuming a fairly large number of events (30), long-duration events (1 h), large energy reductions during all events (normalized value of 1 kW), and an average energy cost of $0.15/kWh, we can calculate that in this "worst-case" scenario we would be ignoring $4.50/kW-year of decreased energy costs.

2.1 Potential revenue across end uses, business segments, and geographic location

To calculate potential revenue, we gathered hourly commercial load data that have been standardized to typical weather conditions and disaggregated by geographic zone, business segment, and end use. Using models of these profiles, we created new profiles specifically for the period 2011–2013. Normalized hourly profiles were then matched with hourly market clearing prices to calculate potential revenue.

By using normalized load profiles to represent DR resource availability, we assume that DR resource availability for reserves is proportional to the load of that particular end use at that particular time. For energy or capacity events, which can last from 1 to 8 h in California [11], this may not be an appropriate assumption. Commercial customers may not want a portion of their electrical service interrupted for that period of time due to operational constraints. However, spinning reserve events typically last for only 10–20 min, and thus customers can shed larger percentages of their load without suffering major interruptions to business operations. Data from PJM, the only region to publish hourly market clearing resource amounts for DR in spinning reserve, support this assumption (see Appendix A for discussion of this topic).

2.1.1 Load disaggregation

One of the only large scale studies to quantify end use level demand across a broad geographic area is the 2006 California Commercial End Use Survey (CEUS) [23]. The CEUS collected metered data from a stratified sample of approximately 2700 buildings in order to create hourly end use level load profiles. The sample was stratified across 12 geographic zones and 12 building segments (Table 1). For each building in the survey, a simulation model that disaggregates whole-facility load into 13 end uses (Table 1) was built in a DOE-2.2 energy simulation environment. Simulation results were calibrated to actual consumption and weather data to ensure the model was accurate. Once calibrated, the building model was run on a new standardized weather set meant to represent a typical weather year in California. Buildings within each sample strata were aggregated to produce weighted average hourly profiles. 1872 unique hourly profiles were created across all geographic zones, building segments, and end uses.

Certain end uses from CEUS were removed from our consideration because they are not appropriate for spinning reserve. For example, exterior lighting is not a good

Table 1 Building segments and end uses in CEUS

Building segments	
College	School
Grocery	Restaurant
Health	Small office
Lodging	Large office
Miscellaneous	Refrigerated warehouse
Retail	Un-refrigerated warehouse
End-uses	
Heating	Interior lighting
Cooling	Exterior lighting
Ventilation	Miscellaneous
Refrigeration	Office equipment
Hot water	Motors
Cooking	Process
	Air compressor

Table 2 End uses/segments removed in this study

End-uses removed	Reason for removal
Exterior lighting	Code issues
Process	Business process constraints
Cooking	Business process constraints
Office equipment	Business process constraints
Miscellaneous	Unknown resource type
Segment removed	Reason for removal
Small office	Does not match cost data

candidate for spinning reserve because reducing exterior lighting at night may violate building codes. This left 981 profiles. The list of removed end uses and business segments (along with a reason for removal) is contained in Table 2.

2.1.2 Load modelling

To convert the standardized profiles from CEUS to 2011–2013 profiles, we first separated end uses into weather and non-weather dependent categories. Non weather-dependent end uses were converted using a day-matching method. Consumption values for each hour of the day in each month were averaged, treating weekdays and weekends separately. While heating would normally be considered a weather-dependent end use, regression modelling was not successful in capturing the variation of heating profiles. Therefore, the day-matching method was used for all heating profiles.

Regression models with ARMA errors were used for weather-dependent end uses (cooling and ventilation). Via 10-fold cross-validation, we explored over 20 model specifications. The final model (Eq. 1) showed the lowest average out-of-sample error across all cooling and ventilation profiles (See Table 3 for variable descriptions). We also investigated using lagged weather variables. Due to thermal mass, buildings often show a lagged response to outdoor temperature and humidity conditions. However, current weather conditions showed better out-of-sample prediction error than lagged weather conditions for the standardized CEUS load profiles. We believe this is an artifact of the modeling process used in the CEUS project and does not reflect what one would find if raw metered data was used.

All models exhibited significant autocorrelation in the residuals. To facilitate more accurate prediction, we chose to model the error using time-series (ARMA) parameters. A necessary condition for parameter estimation using time-series models is homoscedasticity. However, a plot of the residuals for most load profiles revealed two distinct periods during the year for which residual variance was uneven (summer vs. winter). We thus split the annual standardized models into three periods: the first winter period (Jan–Apr), summer (May–October), and last winter period (Nov–Dec). The selection of periods for boundary months (e.g. April) was performed by examining how closely the residual variance of the month compares to other months when it was included in the winter or summer model.

The same ARMA model specification for the error term was used across all load profiles because it was successful in removing most of the autocorrelation in residuals across load profiles. We attempted to include parameters at other lags but often found that they did not reduce the Bayesian Information Criteria (BIC) and/or the coefficient estimates were not statistically significant.

$$\ln(\mathrm{kW}_t) = \alpha + \sum_{d=Wkdy}^{Wknd} \sum_{h=1}^{24} \begin{bmatrix} \beta_{Temp}^{d,h} \mathrm{I}^d \mathrm{I}^h Temp_t \\ +\beta_{Temp^2}^{d,h} \mathrm{I}^d \mathrm{I}^h Temp_t^2 \\ +\beta_{T.RH}^{d,h} \mathrm{I}^d \mathrm{I}^h Temp_t RelH_t \end{bmatrix} + \eta_t \qquad (1)$$

where:

$$(1 - \phi_1 \mathrm{B}^1 - \phi_{24} \mathrm{B}^{24})\eta_t = (1 + \theta_1 \mathrm{B}^1)\epsilon_t$$

Once the model coefficients were estimated for each load profile, predicting the 2011–2013 hypothetical load profiles was a 2-step process. Coefficients for exogenous weather variables were multiplied against actual hourly 2011–2013 weather data to form the base of the prediction. Next, 5000 separate ARMA simulations were conducted using the time-series coefficients from each of the three period models (the length of the simulations was tailored for the period of the year). The simulated error at each time-step was independent and identically distributed (i.i.d.) and randomly drawn from a normal distribution with variance equal to the residual variance of the model. All simulations used a burn-in period of 50,000 iterations. The average path of the 5000 simulations was added to the predictions from the exogenous variables to form the overall predicted load profile.

In using load data captured in 2002 to infer load profiles for 2011–2013, we assume the shape of the end use load profiles has not changed over time. Load shapes could

Table 3 Equation 1 variable descriptions

Variable	Description
kW_t	Average kilowatt consumption in hour of the year t
I^d, I^h	Indicator variables for day type d (weekday/weekend) and hour of day h
$Temp_t$	Temperature (°F) in hour t
$RelH_t$	Relative humidity in hour t
$\boldsymbol{\beta_x^{d,h}}$	Regression coefficient for day type d and hour of day h for weather variable x (Temp, Temp^2, or Temp*RelH)
η_t	Error in hour t unexplained by exogenous weather variables
B	Backshift operator
$\boldsymbol{\theta}_{\mathrm{i}}$	Coefficient for moving average term of lag i
$\boldsymbol{\phi}_{\mathrm{i}}$	Coefficient for autoregressive term of lag i
$\boldsymbol{\varepsilon}_t$	Unexplained error in time t

change due to shifts in equipment stock (e.g. higher saturations of more efficient equipment) and equipment use patterns. However, commercial load has not grown in California since 2005 [19]. Load growth is not a perfect measure of changes in end use load profiles, but the authors believe it is reflective of a load environment that is in steady-state.

2.1.3 Normalization and revenue potential

Normalization of the load profiles was necessary to express our results in a standardized measure of size (per kW). Profiles were normalized to the average load during the top 50 h in each year by temperature, which closely mirrors the method used to calculate peak kW for incentive payments in California's AutoDR program. Equation 2 displays the normalization calculation for each hour t in the profile.

$$\mathrm{kW}_{norm,t} = \frac{\mathrm{kW}_t}{average\left(\sum \mathrm{kW}_{50\ \mathrm{h},2011}, \sum \mathrm{kW}_{50\ \mathrm{h},2012}, \sum \mathrm{kW}_{50\ \mathrm{h},2013}\right)} \tag{2}$$

The calculation of revenue was completed by matching the hourly normalized resource availability with the day-ahead market clearing price in that hour. Market clearing prices for spinning reserve are not the same across the entire CAISO region. CAISO has established separate procurement requirements for operating reserves in areas "north of path 26" (NP26) and "south of path 26" (SP26) to ensure that contingencies can be mitigated even in the case of congestion on the Path 26 transmission line. The Path 26 transmission line in central California roughly delineates the boundary between SCE and PG&E. This area is a bottleneck for power trying to flow between northern and southern California. Variations in generation mix and transmission network topology among the two regions lead to price differences. Prices for NP26 and SP26 were matched with the different forecasting zones from the CEUS. Table 4 details how the load forecasting zones (FCZ) were mapped to ancillary service zones (ACZ).

Table 4 Forecasting zone mapping to ancillary service zone partitions

FCZ in service zone "CAISO"	FCZ in service zone "SP26"
FCZ 1	FCZ 7
2	8
3	9
4	10
5	13
6	

In making this calculation we assume perfect forecasting of resource availability, which would tend to increase our revenue numbers. However, this did not affect the final conclusions of the study. We also assume that load resources are price-takers that do not affect the market clearing price. While ancillary service participants are worried that markets will saturate quickly and prices will collapse [24], as long as some traditional generation remains in the spinning reserve market prices may not decrease significantly due to the payment of lost opportunity costs of energy production [5].

2.2 Costs

An aggregator would incur a number of costs in setting up a spinning reserve portfolio, including equipment installation for controls and automated response, telemetry for monitoring loads, equipment maintenance, participant incentives, program administration, forecasting, and CAISO administrative fees. The communications architecture of such a system in described in Fig. 1. These expenditures would also allow participation in capacity/energy DR programs, or frequency regulation markets as control devices advance in sophistication, though in our analysis we assume end users face business constraints that prevent them from participating in energy/capacity programs.

Unfortunately, detailed cost information for this type of system across many types of loads/businesses does not exist. The closest program for which information is publicly available is PG&E's Participating Load Pilot [18]. It had only 3 participants and much of the cost for the program was spent on one-time startup costs. We were able to obtain generalized cost information for DR equipment installation (described in the next section) but cannot tie the data to specific end-uses or business types. Therefore, we treat the costs an aggregator would incur as a parametric variable in our results, reporting ranges that would provide a sufficient payback on invested capital. We use the generalized cost information on equipment installation to provide context for the reported cost ranges.

2.2.1 Equipment cost for event communication and automated response

In order for DR to provide spinning reserve within the required 10 min, automated response is necessary. Personal notifications (email or phone) and manual changes to equipment operating parameters cannot guarantee 10-min response. Automated

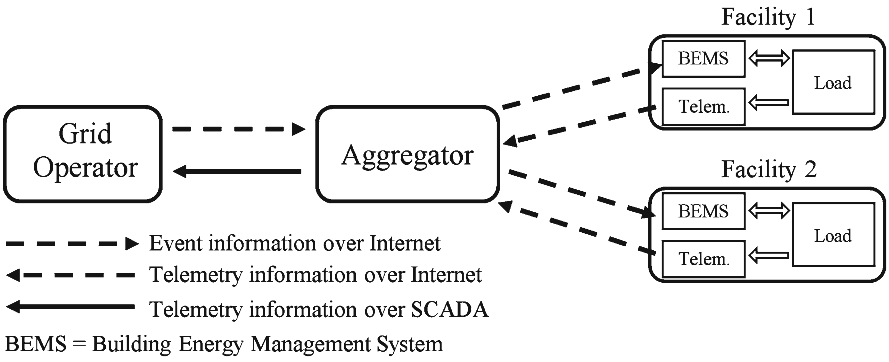

Fig. 1 System communication architecture for loads participating in spinning reserve. Communication from the grid operator to the aggregator, and to and from the aggregator and the facility can take place over secure internet connections. Telemetry reporting from the aggregator to the grid operator must take place via a more demanding Supervisory Control and Data Acquisition (SCADA) protocol. Communication architecture design based on the OpenADR 2.0 standard (OpenADR Alliance 2014). Telemetry architecture from [25]

response can be enabled by pre-programming DR strategies into control equipment so the response is implemented without human intervention.

California investor-owned utilities provide incentives for the installation and programming of such equipment through the Automated Demand Response (AutoDR) program. Salient features of the California AutoDR program are:

1. Designed for commercial/industrial customers with peak load >200 kW.
2. Requires participation in utility energy or capacity DR programs.
3. Incentives are capped at the minimum of 100% of total project cost or $300/kW of load response. These are one-time payments (not annual).
4. The amount of load response must be proven through a test event or actual performance history from energy or capacity events.

Incentive data were collected from Pacific Gas and Electric (PG&E) and Southern California Edison (SCE). Project-level incentive information from San Diego Gas and Electric and Sacramento Municipal Utility District was not available. See Appendix B for more information on the treatment of incentive data. Figure 2 displays the combined SCE and PG&E incentive information. The mean cost is approximately $180/kW.

We assume participating commercial buildings have a building energy management system (BEMS) that can communicate with end use level equipment. The market share of BEMS in California commercial buildings is approximately 60% for buildings with an average demand of 200 kW [22].

An aggregator would also have to install telemetry at a participating building because it is required for participation in spinning reserve markets. Telemetry allows the grid operator to obtain real-time information on load characteristics, such as real and reactive power. Energy and capacity DR programs do not rely on telemetry for measurement and verification of load reductions—they use interval meter data that are already captured for billing purposes.

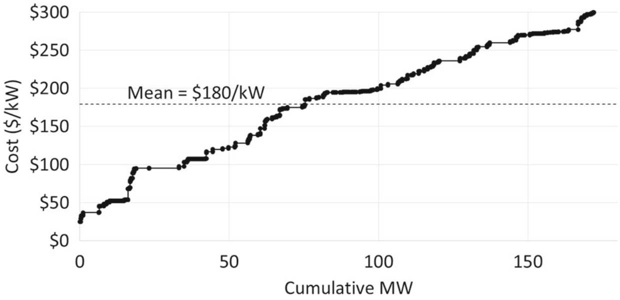

Fig. 2 Incentives provided to install communications equipment, program and commission DR strategies. *Note* Incentive data includes commercial and industrial customers. Industrial customers could not be removed because the project database lacked identifying information. We do not believe that removing industrial customers would significantly affect the cost distribution as large projects were evenly spread across higher and lower $/kW values

For small distributed resources like DR, the cost of telemetry is a significant obstacle to participation in ancillary service markets. Estimates of the cost of telemetry for a large commercial building are approximately $50,000–$80,000 [25]. Given the average load response in the AutoDR program, this cost would translate to over $200/kW. However, new designs have the potential to provide telemetry at much lower cost. Early tests show large commercial buildings could be outfitted with telemetry at an approximate cost of $50/kW of controlled load [25] or 1/4 of the current cost estimate. We use the $50/kW estimate to provide context for our results.

3 Results

We find end uses with relatively constant load profiles throughout the year, such as lighting or refrigeration, are better suited for spinning reserve than seasonal end uses like cooling and heating. This is counter to the intuition behind traditional capacity-based DR programs that focus on seasonal end uses because they are highly correlated with the system peak demand. Spinning reserve, however, is needed at all times and is therefore best served by resources which are available at all times. Figure 3 shows the results by end use and building segment combinations across all of the forecasting zones.

While cooling is the largest end use by peak load in California, it nevertheless has very low revenue potential because of its seasonal nature. Interior lighting is a large end use and is well suited for spinning reserve, especially in building segments that operate on continuous schedules such as lodging. The school and college segments which have lower seasonal loads during capacity strained periods do especially poorly.

These revenue figures are next used to determine the maximum allowable cost at which an aggregator would find the simple payback of their investment to be 5 years or less. Simple payback can be calculated as the ratio of costs to annual revenue. The 5 year simple payback threshold is important because many companies use simple payback

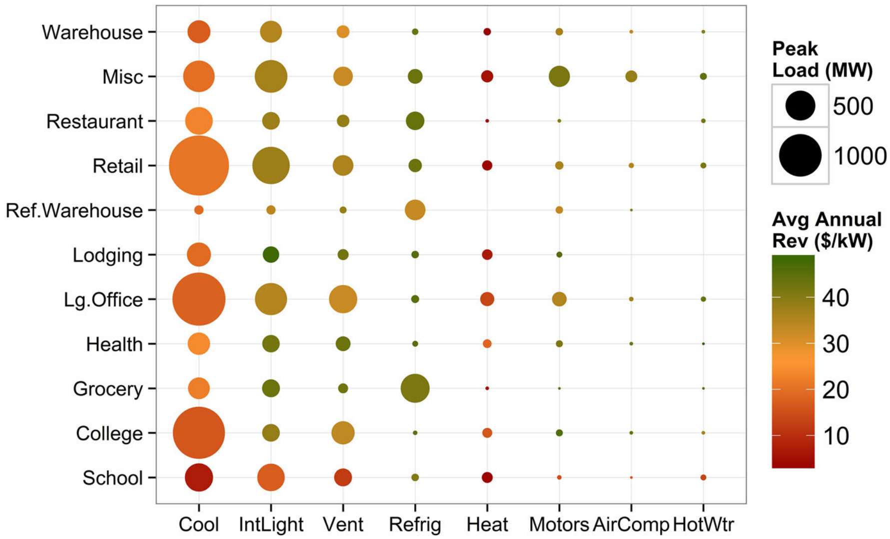

Fig. 3 Average annual revenue for end use/building segment combinations. The *area of the dot* represents the total peak load for that combination across all forecasting zones. The *shading of the dot* corresponds to the average annual revenue potential. Average annual revenue is calculated as a weighted average across all zones, weighted by peak load

as a metric for energy decisions and most of these companies use a threshold of 5 years or less [26]. Figure 4 shows the distribution of these maximum costs across (a) end uses and (b) business segments. When viewing the figure, if the reader imagines that the true cost to an aggregator was $200/kW, any point on a distribution below $200/kW would have a payback greater than 5 years. In general, higher maximum allowable costs represent those end uses/business segments that have higher revenue. The horizontal lines spanning the graphic show the low and high end of the cost distribution for communication and control equipment discussed in Sect. 2.2.

For the majority of potential participants, total costs incurred by the aggregator would need to be below $250/kW to achieve a payback of 5 years or less, though the highest cost for any end use to achieve the 5 year threshold is $340/kW. The median cost for a 5 year payback across all end uses excluding cooling and heating is $173/kW. These are plausible maximum cost values given the distribution of equipment installation costs, though we should remind the reader that this does not include many other costs an aggregator would face (e.g., participant incentives). Thus we find that the business case probably exists to provide spinning reserve from pooled DR resources, though the aggregator would need to be selective in targeting participants.

We do not find important differences in revenue potential across geographic zones. The largest driver of difference across zones is the market price for spinning reserve; southern California (below the Path 26 transmission line) often has higher prices than northern California. Figure 5 shows price duration curves for northern and southern California.

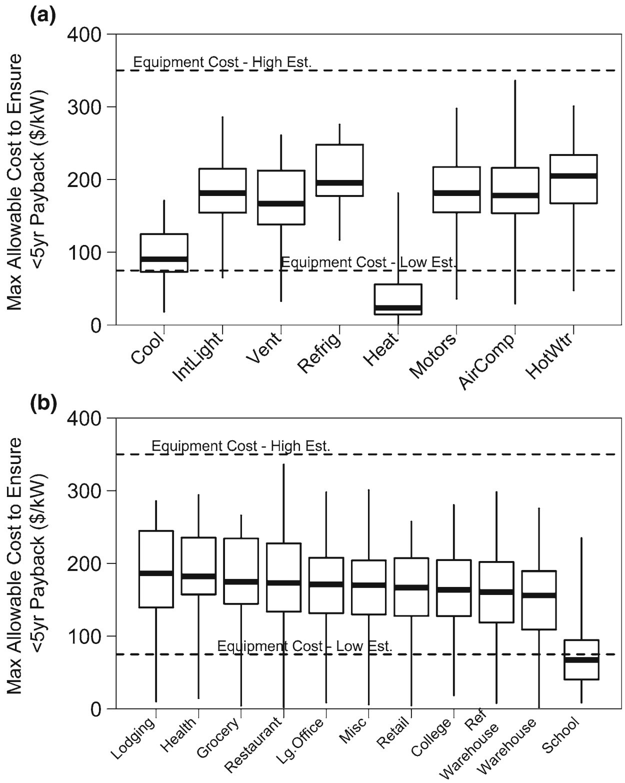

Fig. 4 Distribution of maximum allowable costs ($/kW) incurred by an aggregator to keep payback periods under 5 years across **a** end uses and **b** business segments. Higher maximum allowable costs represent end uses/segments that have higher revenue. *Horizontal lines* spanning the figure represent the low and high estimate of equipment installation costs, including control equipment and telemetry (ignores other types of costs like participant incentives). Each combination of geographic zone, business segment and end use represents a single point within each distribution. The *heavy horizontal line in the middle of each box* marks the median. The *range of the box* represents the interquartile range. The *whiskers* extend to the extremes of the distribution

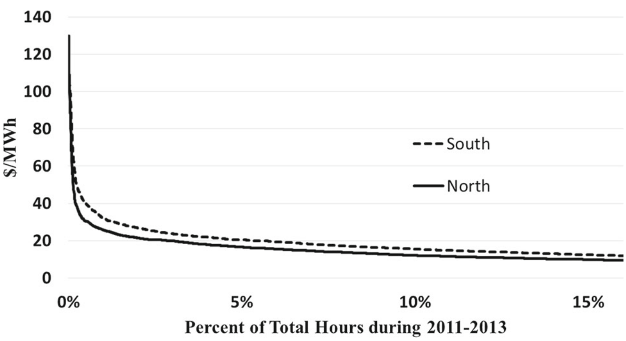

Fig. 5 Partial price duration curves for California spinning reserve prices from 2011–2013. Prices are often higher in the southern California zone (below the Path 26 transmission line). *Horizontal axis* abbreviated for clarity

4 Policy implications: avoided carbon emissions

We have shown that aggregators will have to be selective in targeting potential DR participants, possibly leaving a large amount of DR on the sidelines of the market. However, providing incentives to DR would improve economics and encourage participation. We now consider if such an incentive is justified by a market failure not currently captured in spinning reserve clearing prices: the damages associated with carbon dioxide (CO_2) emissions from fossil fuel power generation. California already considers the social cost of carbon in their cost effectiveness tests for utility energy efficiency and DR programs [27].

To our knowledge, there has been no detailed study of the emissions avoided from DR participation in electricity markets for either energy or ancillary services. Studies of avoided emissions in reserve markets have mostly focused on renewable energy [28] or pumped hydroelectric power [29]. The most rigorous approach to this problem would make use of a dispatch model of the California grid to understand the quantity and type of fossil fuel power plants offset from DR and the duration of offset. Here we instead make a first-order estimate.

The procurement of spinning reserve is fundamentally an *option* to produce power, not an actual call for power. Marginal changes in the fuel mix of reserves that do not change the overall energy dispatch will not displace emissions, as nothing has physically changed on the grid. However, if enough DR is procured to offset the reserve provided by an entire plant, that plant can shut down. This assumes that the marginal plant used for reserves is online only because of the need to provide reserve. We adjust for this assumption in our calculations. The emissions saved would be the difference between the reduction from turning off the partly-loaded reserve plant and the increase of the base load plant that is now making up for the energy generation of the reserve plant.

To calculate emissions savings, it is thus important to understand the fuel types that typically provide spinning reserve and base load. The 2013 CAISO Annual Report on Market Issues and Performance [30] reports that hydro supplies approximately half of the spinning reserve in a typical year. Natural gas and imports supply approximately a quarter of this reserve each. Droughts and changing climate patterns, however, may reduce the potential for high-elevation hydropower production in California in the future [31]. Reduced hydropower energy production is typically offset by natural gas in California [32]. We assume that reduced spinning reserve from hydropower is also offset by natural gas.

Natural gas plants represent the majority of the available dispatchable generation in CAISO, hence the energy production from plants providing reserve that are offset by DR is likely assumed by other natural gas generation. We assume that all natural gas generation is performed by combined-cycle (NGCC) plants. In reality, some spinning reserve is provided by natural gas combustion turbines (NGCTs). NGCTs have higher heat rates than NGCC plants. Thus, ignoring NGCTs likely underestimates carbon savings. In this analysis, we focus just on the emissions and associated damages from CO_2 and not from criteria pollutants (e.g., sulfur dioxide, nitrogen oxide, particulate matter). This first-order analysis does not consider the emissions savings during actual spinning reserve events, only the savings from a different economic dispatch of generation resources. However, criteria pollutant emissions savings during spinning reserve events may be significant. Nitrogen oxide ramping emissions from simple-cycle natural gas combustion turbines can be significantly higher than steady state emissions [34]. Thus during a spinning reserve event, demand response can offset much higher emissions from ramping natural gas plants than it does under normal dispatch conditions.

Social damages from CO_2 are orders of magnitude larger than damages from criteria pollutants for natural gas plants. Assuming damages of \$37 per tonne of CO_2 [35] and emissions of 0.375 tonne of CO_2/MWh [36] for natural gas plants, we calculate damages of ~\$14/MWh. From [37], we find damages from criteria pollutants emitted from natural gas plants on the order of \$0.05/MWh.

The relationship between CO_2 output and power generation is nearly linear for a NGCC plant [33], thus marginally unloading one plant and reloading another of the same type saves no CO_2. But if one plant is able to be fully shut down, the CO_2 saved is equal to the no-load emissions of that plant. To make a first-order estimate of the annual CO_2 saved from procuring DR for spinning reserve we use Eq. 3. The input assumptions are presented in Table 5. Total reserve was divided by the idle generating capacity of an average plant in order to calculate the number of plants shut down by procuring DR. We assume that there is enough DR to offset the reserve of natural gas plants that provide half the average annual spinning reserve requirement. This corresponds to a future scenario where the proportion of reserves provided by natural gas has increased due to falling hydro reserves.

$$Annual\ CO_2\ Savings = \frac{Total\ Reserve}{Idle} * Carbon_{no_{load}} * \%Reserve * 8760 \qquad (3)$$

We estimate annual carbon savings at approximately one million metric tons ($0.2 \times 10^6 - 2.8 \times 10^6$ tonnes for the low and high scenarios, respectively). Avoided damages associated with carbon emission savings were calculated using two different values of carbon: (1) the social cost of carbon (SCC) computed by the United States government for emissions year 2010 under the average 3% discount rate scenario (\$37 in 2014 dollars) [35], and (2) the average 2014 market price for carbon under California's cap and trade system set up under AB32 (\$12) [39]. The annual results are shown in Fig. 6 relative to the up-front capital required to install telemetry on DR resources.

Figure 6 demonstrates that meaningful incentives for DR might be justified by avoided damages from carbon emissions. The value of avoided damages under AB32 produce far less compelling results than under the SCC, but still reflect a payback of the up-front telemetry capital costs in approximately 2 years.

5 Discussion and conclusion

To allow DR to participate in spinning reserve in California, WECC must modify the definitions that govern eligible resources by removing the requirement to be immediately and automatically responsive to system frequency; thereby bringing its policy

Table 5 Variable descriptions and assumed values for Eq. 3

Variable (units)	Description	Assumed value	Range	
			Low savings scenario	High savings scenario
Total reserve (MW)	The total MW of spinning reserve in CAISO offset by DR	500 MW (approx. half of average spin requirement)	250 MW[a]	750 MW[b]
Idle (MW)	The amount of spinning reserve provided by each natural gas plant (idle generating capacity)	50 MW[c] (approx. 10-min ramp capability for 200 MW combined-cycle turbine)	100 MW[d]	40 MW[e]
$Carbon_{no\_load}$ (Tonnes CO_2/h)	CO_2 emissions at no load	17.5 tonnes [33]	14 tonnes[f]	21 tonnes[g]
% Reserve (Unitless)	Percent of annual hours that system dispatch is reserve-constrained[h]	77%[i]	74%[j]	80%[k]
8760 (h)	Number of hours in a year			

[a] Scenario where DR displaces current reserves from natural gas (∼25% of requirement)
[b] Scenario where DR displaces current reserves from natural gas and hydro (∼75% of requirement)
[c] Ramp rate of 2.5%/min [28]
[d] Ramp rate of 5%/min [38]
[e] Ramp rate of 2%/min (lower end of ramp rates shown in [34, Fig. 4-13])
[f] 5% quantile of the true intercept of [33, Figure S4]
[g] 95% quantile of the true intercept of [33, Figure S4]
[h] Reserve-constrained means that the system dispatch was different from a hypothetical scenario where reserves are not required. Alternatively, a dispatch is *not* reserve-constrained if the removal of the reserve constraints from the system optimization problem does not change the overall dispatch. Reserve-constrained periods are those in which increased DR procurement would cause marginal reserve plants to shut down
[i] Calculated from the average number of hours that spinning reserve prices are above the minimum value from 2011–2013. A reserve price at the minimum value reflects a system which is not reserve-constrained
[j] Calculated from the low annual number of hours that spinning reserve prices are above the minimum value from 2011–2013
[k] Calculated from the high annual number of hours that spinning reserve prices are above the minimum value from 2011–2013

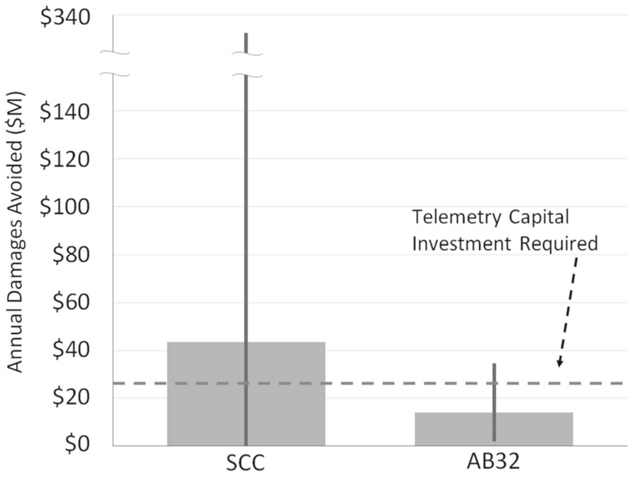

Fig. 6 Damages avoided from carbon emission savings due to DR procurement in spinning reserve market. Uncertainty bars reflect 90% confidence interval for uncertainty in the value of damages per metric ton and the uncertainty in the estimated magnitude of carbon savings (low to high savings scenarios of Table 5). No correlation was assumed between the value of damages per metric ton and the savings scenario. Uncertainty in the value of damages per metric ton for SSC were derived from the distribution of carbon value per ton for the 3% discount rate for emission year 2010 [35]. Uncertainty in the value of damages per metric ton for AB32 were derived from the variance of carbon allowance futures prices during 2014. Capital investment for telemetry calculated at $50/kW

into alignment with most other US wholesale markets. Diversifying the resources providing ancillary services will allow the grid to be more resilient and less operationally expensive.

With an average revenue of ~$35/kW-year, steady end uses (e.g., lighting) have more than twice the revenue than seasonal end uses (e.g., cooling) because spinning reserve is needed year-round. Similarly, business segments with longer operating hours, such as groceries or lodging, have more revenue potential. We find that niche applications of DR could present an attractive business opportunity: certain business segments in southern California can achieve nearly $60/kW-year in revenue from interior lighting. However, this will depend on the total cost to attract spinning reserve resources. To achieve a simple payback of 5 years or less, the median DR resource in California would need to have a total enablement cost of $173/kW or less. Refrigeration resources with more constant profiles could be profitable with median enablement costs of $200/kW, while cooling loads would require costs below $90/kW to be profitable. This is plausible given data on equipment installation costs for automatic DR in California, but the large range of cost data suggests an aggregator would need to be careful in targeting participants.

Enablement costs for DR are likely to decrease in the future as technologies find a common standard and production volumes increase. NIST is working on smart grid interoperability standards [40] and California recently required new control systems for lighting, heating and air conditioning be able to receive automated DR signals [41]. Our analysis included a cost reduction for telemetry of a factor of 4 under current cost estimates. This will help make DR for spinning reserve more economically attractive.

Avoided carbon emissions from using DR instead of fossil fuel generation for spinning reserve could justify the provision of incentives for the cost of installing telemetry (∼\$50/kW) for DR resources. If 500 MW of DR replaced fossil generation in the spinning reserve market, we estimate an annual carbon savings of approximately one million metric tons. Avoided emissions may be larger in other regions with higher proportions of coal-fired resources.

Acknowledgements This work was supported by grants from the Richard King Mellon Foundation and the Electric Power Research Institute. Neither funding source had any role in study design, collection, analysis and interpretation of data, writing of the paper, or in the decision to submit the article for publication. We thank Ines Azevedo, Alex Davis, Roger Lueken, Stephen Rose and Kyle Siler-Evans of Carnegie Mellon University for helpful comments and suggestions. We also thank Girish Ghatikar of Lawrence Berkeley National Laboratory for providing background and incentive information related to the AutoDR program, as well as Daniel Olsen of Lawrence Berkeley National Laboratory and Mark Ciminelli of the California Energy Commission for their help in deciphering the CEUS data and results.

Appendix A: DR availability proportional to load

Figure 7 displays DR clearing MW in each hour of the day across 4 seasons for spinning reserve. The Pearson correlation coefficient between the median DR MW cleared in a given hour across all days of 2012–2013 and the median load for that hour of the day was 0.92. DR clearing amounts in the summer appears quite low—this may be due to other more lucrative DR opportunities (such as capacity) during those times.

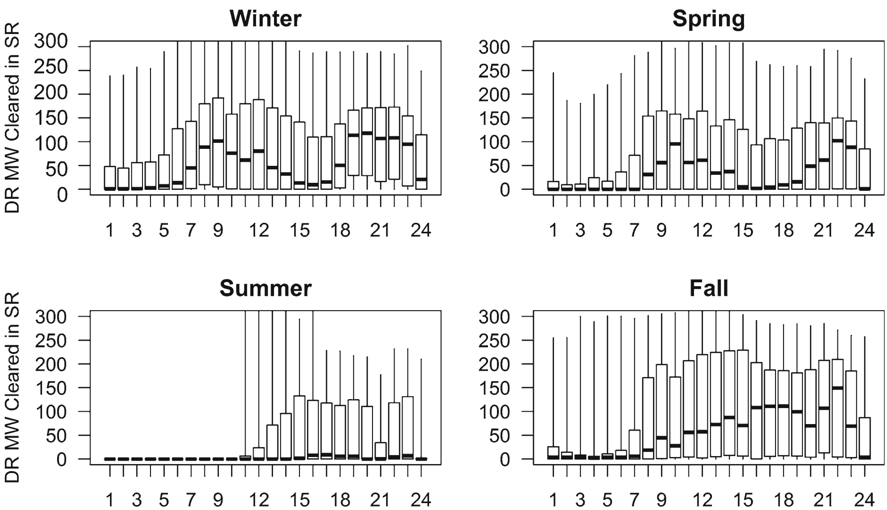

Fig. 7 Demand response MW cleared in spinning reserve market for each hour of the day in PJM during the period 2012–2013. The pattern of cleared demand response mimics the typical overall load pattern seen in each season. The heavy *horizontal line* in the middle of each box marks the median. The *range of the box* represents the interquartile range. The *whiskers* extend to the extremes of the distribution

Appendix B: Reasoning for removal of projects from cost information in SCE

Figure 8 displays the incentive information from PG&E and SCE.

The authors conducted an investigation into the AutoDR program costs and found that nearly all of the projects which had incentives of \$300/kW in the SCE territory were likely from one contractor that received money from the American Recovery and Reinvestment Act (ARRA) grant funds. We surmise that the use of ARRA funds may have led to different recruitment practices and cost reporting. Thus, we do not believe that the incentive information reported for these projects is representative of the rest of the project population. The list below provides details on why the authors believe that these projects were from one contractor.

- An AutoDR program report stated that "the U.S Department of Energy's \$11.4 million American Recovery and Reinvestment Act grant influenced a larger load shed and enablement cost in the SCE territory" [19].
- ARRA records show a total AutoDR project cost of \$22.8M in SCE [42] attributable to one company. The 50% cost sharing required by ARRA leads to a grant of \$11.4 million.
- There are 348 facilities in the project incentive database from SCE that had project incentives of \$300/kW. These projects have a total load response of 67MW. The total rebate amount given to these participants was just over \$20M, which closely matches the ARRA project cost report.

We believe that most, if not all of the projects with incentive values at \$300/kW were not representative of the true costs to install, program, and commission this equipment. This is especially apparent when you compare the incentive distribution from SCE with that of PG&E. There may be other projects in the database with incentive costs of less than \$300/kW that were implemented by this DR contractor. However, we have no way of differentiating those projects.

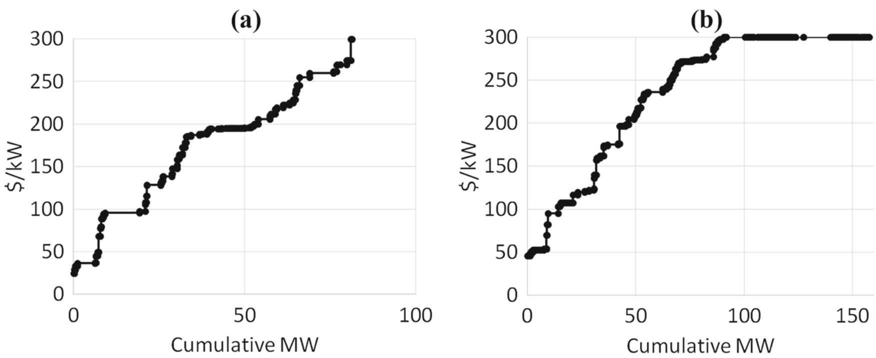

Fig. 8 **a** Incentives Provided by PG&E for AutoDR. **b** Incentives Provided by SCE for AutoDR

References

1. US DOE: Benefits of demand response in electricity markets and recommendations for achieving them. A Report to Congress in fulfillment of Sec. 1252(e) of the Energy Policy Act of 2005 (2006)
2. Kirby, B.: Spinning Reserve From Responsive Loads. Oak Ridge National Laboratory. ORNL/TM-2003/19 (2003)
3. Eto, J.H., Nelson-Hoffman, J., Parker, E., Bernier, C., Young, P., Sheehan, D., et al.: Demand Response Spinning Reserve Demonstration. Lawrence Berkeley National Laboratory. LBNL-62761 (2007)
4. Callaway, D.S.: Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy. Energy Convers. Manag. **50**(5), 1389–1400 (2009)
5. Cappers, P., MacDonald, J., Goldman, C.: Market and Policy Barriers for Demand Response Providing Ancillary Services in U.S. Electricity Markets. Lawrence Berkeley National Laboratory. LBNL-6155E (2013)
6. NERC: Balancing and frequency control: a technical document prepared by the NERC resources subcommittee. Available at: http://www.nerc.com/docs/oc/rs/NERC%20Balancing%20and%20Frequency%20Control%20040520111.pdf. Accessed April 2014 (2011)
7. CAISO: Business practice manual for market operations. Version 40, last revised 7/9/2014. Available at: http://www.caiso.com/rules/Pages/BusinessPracticeManuals/Default.aspx. Accessed October 2014 (2014)
8. Ellison, J.F., Tesfatsion, L.S., Loose, V.W., Byrne, R.H., 2012. Project report: a survey of operating reserve markets in U.S. ISO/RTO-managed Electric Energy Regions. Sandia National Laboratories. SAND2012-1000
9. Callaway, D.S., Hiskens, I.: Achieving controllability of electric loads. Proc. IEEE **99**(1), 184–199 (2011)
10. Watson, D.S., Matson, N., Page, J., Kiliccote, S., Piette, M.A., Corfee, K., et al.: Fast Automated Demand Response to Enable the Integration of Renewable Resources. Lawrence Berkeley National Laboratory. LBNL-5555E (2012)
11. Southern California Edison: Demand response event history. Available at: https://www.sce.openadr.com/dr.website/scepr-event-history.jsf. Accessed Nov 2014 (2014)
12. 145 FERC 61,141: Regional reliability standard BAL-002-WECC-2—contingency reserve, Order No. 789 (2013)
13. Mathieu, J.L., Dyson, M., Callaway, D.S.: Using residential electric loads for fast demand response: the potential resource and revenues, the costs, and policy recommendations. In: ACEEE Summer Study on Energy Efficiency in Buildings, pp. 189–203 (2012)
14. MacDonald, J., Cappers, P., Callaway, D., Kiliccote, S.: Demand response providing ancillary services. Presented at Grid-Interop Forum 2012, Irving, TX. LBNL-5958E. Available at: http://drrc.lbl.gov/sites/all/files/LBNL-5958E.pdf. Accessed March 2014 (2012)
15. MacDonald, J., Kiliccote, S., Berkeley, L.: Commercial building loads providing ancillary services in PJM. In: 2014 ACEEE Summer Study on Energy Efficiency in Buildings, pp. 192–206 (2014)
16. Ma, O., Alkadi, N., Cappers, P., Denholm, P., Dudley, J., Goli, S., et al.: Demand response for ancillary services. IEEE Trans. Smart Grid **4**(4), 1988–1995 (2013)
17. Hummon, M., Palchak, D., Denholm, P., Jorgenson, J., Olsen, D., Kiliccote, S., et al.: Grid Integration of Aggregated Demand Response, Part 2: Modeling Demand Response in a Production Cost Model. National Renewable Energy Laboratory. NREL/TP-6A20-58492 (2013)
18. PG&E: 2009 Pacific Gas and Electric Company participating load pilot evaluation. Available at: http://www.pge.com/includes/docs/pdfs/mybusiness/energysavingsrebates/demandresponse/cs/2009_pacific_gas_and_electric_company_large_commercial_industrial_participating_load_pilot.pdf. Accessed Sept 2014 (2009)
19. U.S. Energy Information Administration (EIA): Electric Power Annual (2014)
20. Ghatikar, G., Riess, D., Piette, M.A.: Analysis of Open Automated Demand Response Deployments in California and Guidelines to Transition to Industry Standards. Lawrence Berkeley National Laboratory. LBNL-6560E (2014)
21. Siano, P.: Demand response and smart grids—a survey. Renew. Sustain. Energy Rev. **30**, 461–478 (2014)

22. Itron: California commercial saturation survey. Prepared for the California Energy Commission. Available at: http://calmac.org/publications/California_Commercial_Saturation_Study_Report_Finalv2.pdf. Accessed Oct 2014 (2014)
23. Itron: California commercial end use survey. Prepared for the California Energy Commission. CEC-400-2006-005 (2006)
24. US DOE: Load participation in ancillary services: workshop report. Available at: https://www1.eere.energy.gov/analysis/pdfs/load_participation_in_ancillary_services_workshop_report.pdf. Accessed Feb 2014 (2011)
25. Kiliccote, S., Lanzisera, S., Liao, A., Schetrit, O., Piette, M.A.: Fast DR: controlling small loads over the internet. In: 2014 ACEEE Summer Study on Energy Efficiency in Buildings, pp. 196–208 (2014)
26. Prindle, W., de Fontaine, A.: A survey of corporate energy efficiency strategies. In: ACEEE Summer Study on Energy Efficiency in Buildings, pp. 77–89 (2009)
27. California Public Utilities Commission: Demand Response Cost Effectiveness Protocols. R.07-01-041 (2010)
28. Fripp, M.: Greenhouse gas emissions from operating reserves used to backup large-scale wind power. Environ. Sci. Technol. **45**(21), 9405–9412 (2011)
29. Koritarov, V., Veselka, T., Gasper, J., Bethke, B., Botterud, A., Wang, J., et al.: Modeling and Analysis of Value of Advanced Pumped Storage Hydropower in the United States. Argonne National Laboratory. ANL/DIS-14/7 (2014)
30. CAISO: 2013 Annual report on market issues and performance. Department of Market Monitoring. Available at: http://www.caiso.com/Documents/2013AnnualReport-MarketIssue-Performance.pdf. Accessed Sept 2014 (2013)
31. Phinney, S., McCann, R.: Potential changes in hydropower production from global climate change in California and the Western United States. Prepared for the California Energy Commission. CEC-700-2005-010 (2005)
32. U.S. Energy Information Administration (EIA): California drought leads to less hydropower, increased natural gas generation. Available at: http://www.eia.gov/todayinenergy/detail.cfm?id=18271#tabs_SpotPriceSlider-1. Accessed Nov 2014 (2014)
33. Katzenstein, W., Apt, J.: Response to comment on "Air emissions due to wind and solar power". Environ. Sci. Technol. **43**(15), 6108–6109 (2009)
34. Katzenstein, W.: Wind power variability, its cost, and effect on power plant emissions. Dissertation submitted for Doctor of Philosophy. July 2010. Carnegie Mellon University (2010)
35. Interagency Working Group on Social Cost of Carbon: Technical support document: technical update of the social cost of carbon for regulatory impact analysis under executive order 12866 (US Government, Washington, DC). Revised November 2013 (2013)
36. U.S. Energy Information Administration (EIA): Assumptions to the Annual Energy Outlook 2014. U.S. Department of Energy, Washington (2014)
37. National Research Council: Hidden Costs of Energy: Unpriced Consequences of Energy Production and Use. National Academy, Washington (2010)
38. Black & Veatch: Cost report: cost and performance data for power generation technologies. Available at: http://bv.com/docs/reports-studies/nrel-cost-report.pdf%E2%80%8E. Accessed Nov 2014 (2012)
39. Climate Policy Initiative: California carbon dashboard. Available at: http://calcarbondash.org/. Accessed 24 Nov 2014 (2014)
40. National Institute of Standards and Technology: NIST framework and roadmap for smart grid interoperability standards, Release 3.0. Special Publication 1108r3. http://dx.doi.org/10.6028/NIST.SP.1108r3 (2014)
41. California Energy Commission: 2013 Building energy efficiency standards: title 24. CEC-400-2012-004-CMF-REV2. Available at: http://www.energy.ca.gov/2012publications/CEC-400-2012-004/CEC-400-2012-004-CMF-REV2.pdf. Accessed Nov 2015 (2013)
42. US DOE: Recovery act selections for smart grid investment grant awards—by category. Available at: http://www.energy.gov/sites/prod/files/SGIG%20Awards%20by%20Category%202011%2011%2015.pdf. Accessed 15 Sept 2014 (2011)