Scaling Dynamic Pricing: What Can We Squeeze Out of Today's Pilots?

Jingjing Liu, Mary Ann Piette, Lawrence Berkeley National Laboratory
Jahon Amirebrahimi, Albert Chiu, Pacific Gas and Electric Company
Mark Martinez, Eva Molnar, Reginald Wilkins, Southern California Edison

ABSTRACT

California has established a policy goal of achieving 7 gigawatts (GW) of load shift by 2030, about 2 GW of which would come from dynamic-pricing-based load shift growth. The California Load Flexibility Research and Deployment Hub (CalFlexHub) has demonstrated 18 different flexible-load technologies to be dynamically price-responsive at about 95 residential/commercial field sites. Concurrently, California utilities are piloting dynamic pricing to gather evidence of viable solutions and customer experience. To meet the state's policy goal, it will be crucial to quickly scale dynamic pricing from the current pilot stage to large-scale utility programs that can promote promising solutions/models with proper incentives.

This paper presents lessons from demonstrating early commercial solutions in CalFlexHub and the utility dynamic pricing pilots, which are critical for informing the development of transformational load shift programs and incentives by utilities and community choice aggregators. Building on these insights, the paper provides a set of 10 evaluation criteria that focus on the long-term scale-up implications for utility pilot evaluation. These criteria effectively assess the following aspects of each pilot solution in three categories: (1) the load-flexibility functions (controlled load, predictability, service impact, open ecosystem); (2) accessibility (affordability, simplicity, and customer choice); and (3) adverse scenarios (connectivity resilience, customer troubleshooting support, and vendor lock-in risk). The paper uses early commercial solutions in CalFlexHub as examples to illustrate how using these evaluation criteria can identify tradeoffs, promising long-term solutions, and investment opportunities. The framework will help utilities and community choice aggregators optimize takeaways from current pilots and design targeted, market transformational programs to scale these solutions.

1. Introduction

1.1 California's Policy Background for Dynamic Pricing Pilots

The state of California has identified using dynamic retail electricity prices to motivate flexible loads as a key strategy in response to multiple grid challenges (Liu et al. 2024). These challenges include the need to regularly shift load away from the 4–9 p.m. system peak period, to reduce load under system or locational grid stress conditions, and to manage load growth. During a regional extreme heat event in mid-August 2020, the California Independent System Operator (CAISO) was forced to issue rotating blackouts, cutting power to hundreds of thousands of customers and triggering the first statewide outages due to inadequate supply in nearly 20 years

(CAISO, CPUC, and CEC 2021). As part of its comprehensive approach to bolstering summer reliability after this blackout, the California Public Utility Commission (CPUC) authorized two dynamic pricing pilots: the Valley Clean Energy Agricultural Flexible Irrigation Technology (AgFIT) pilot in Pacific Gas and Electric Company's (PG&E) service territory and Southern California Edison Company's (SCE) Dynamic Rate Pilot to incentivize load shifting by commercial and residential customers through electric vehicles (EVs), onsite energy storage, and similar flexible technologies (CPUC 2021). Shortly thereafter, the CPUC proposed and formalized the California Flexible Unified Signal for Energy (CalFUSE), a framework for continuous, price-signal-driven demand flexibility (CPUC 2022a; CPUC 2022b).

In fall 2021, the CEC launched a major research development, demonstration, and deployment program, the California Load Flexibility and Deployment Hub (CalFlexHub), through Electric Program Investment Charge (GFO-19-309). Awarded to a project team led by Lawrence Berkeley National Laboratory (LBNL), CalFlexHub has aimed to develop and demonstrate a pipeline of automated price-responsive technologies that can enable customer-flexible loads in buildings and EVs (Piette et al. 2022).

Two additional policy drivers added to this momentum. First, the California Energy Commission's (CEC) 2022 amendments to the Load Management Standards (effective April 1, 2023) require large electric utilities and community choice aggregators (CCAs) to develop and offer customers at least one retail electricity rate that changes at least hourly, creating a direct regulatory mandate for dynamic pricing deployment (CEC 2022). Second, as directed by Senate Bill 846, the CEC in May 2023 established a statewide load shift goal of 7 GW of net peak demand reduction by 2030, of which 2 GW is expected to come primarily from dynamic-pricing-based load shift (CEC 2023). Meeting this target will require scaling load shift from dynamic pricing from approximately 30 megawatts (MW) in 2022[1] to 2 GW by 2030 — a roughly 67-fold increase that makes the transition from pilots to scalable programs an urgent priority.

1.2 Status of California Utility Dynamic Pricing Pilots

In January 2024, CPUC's Decision 24-01-032 directed the state's two investor-owned utilities (IOUs) — PG&E and SCE — to expand the two dynamic pricing pilots to serve all customer classes and provide system reliability benefits between June 2024, and December 2027 (CPUC 2024). It established combined enrollment targets of 100 MW for PG&E and 50 MW for SCE. Table 1 compares the two IOU expanded dynamic rate pilots in detail.

For PG&E, D.24-01-032 expanded the AgFIT Pilot to (a) bundled and unbundled agricultural customers and (b) bundled and unbundled[2] residential, commercial, and industrial customers (some of which have load-shifting technologies) across its service territory. PG&E's current expanded dynamic pricing pilot — marketed as Hourly Flex Pricing (HFP) — is using a shadow billing approach that bypassed earlier billing system barriers while providing customer

[1] Although 2026 enrolled MW for PG&E and SCE's dynamic rate pilots are available in Table 1, those represent controlled load rather than measured load shift MW. Therefore, 2022 available load shift from Senate Bill 846 is referenced.

[2] A bundled customer in CA purchases all of their electricity generation and delivery services directly from their local IOU. An unbundled customer receives delivery from the IOU but buys their power generation from an alternative supplier, typically a CCA.

bill protection[3]. Separately, CPUC required the Phase 2 of PG&E's vehicle grid integration (VGI)[4] pilots to offer real-time pricing rates[5] to the participants, all VGI pilot customers migrated into HFP, and the measurement and evaluation (M&E) of HFP and VGI were subsequently combined, with CPUC's approval. Since late 2024, HFP has been offering hourly day-ahead prices with a seven-day forecast via public application programming interface (API), with locational distribution pricing. The 50 MW agricultural and 50 MW residential/commercial enrollment targets were reached by June 2025, with more than 2,300 sites and 18 automation service providers (ASPs) as of April 2026 — exceeding initial expectations.

SCE's initial Dynamic Rate Pilot initiated in 2022 enrolled 40 bundled customers and generated important design lessons around multiple challenges: the original concept of transactive hourly pricing proved administratively burdensome and was rarely used by customers, subscription development as a hedging product was complicated when customers had large year-over-year load changes, and partnering with a vendor lacking interval data access limited scalability (Hansen et al. 2025). These lessons directly shaped SCE's expanded dynamic rate pilot design, which eliminated transactive pricing and selected a scalable billing platform provider. In addition, the expanded pilot allowed CCA customers to participate and has also expanded from HVAC only to participation from more end uses. The expanded pilot began delivering hourly day-ahead prices to a first set of customers in July 2025 and has developed a self-service enrollment portal. It retains a subscription-based shadow billing structure and introduces distribution site-specific locational pricing. As of March 2026, approximately 700 customer service accounts totaling about 150 MW are enrolled, primarily commercial EV chargers and large customer accounts.

Table 1. Summary of California utility expanded dynamic pricing pilots

	PG&E Hourly Flex Pricing (HFP)	SCE expanded Flexible Pricing Rate pilot
Pilot launch	December 2024	July 2025
Eligible customers	All classes: Residential, C&I, agricultural, EVs (VGI), CCAs	All classes: Residential, C&I, CCAs
Target technologies	V1G/V2G EV charging, ag. pumps, behind-the-meter batteries, heat pumps, commercial HVAC	EV charging (home and commercial), HVAC, and water pumping (ag./municipal)
Enrollment target	50 MW Ag + 50 MW Res/Com	50 MW
Actual enrollment	~338 MW / 2,300+ sites / 9 CCAs / 18 ASPs (as of April 2026)	~700 service accounts and ~150 MW (as of March 2026)

[3] PG&E provides annual bill credits to customers whose electricity bill is lower compared to their bill under the otherwise applicable tariff (OAT); therefore, no customers will be charged for more than their OAT bill.
[4] Refers to any method of altering the time, power level, or location at which EVs charge or discharge in a manner that optimizes the interaction with the electric grid and provides net benefits to ratepayers.
[5] A dynamic electricity rate where prices are adjusted frequently, typically on an hourly or sub-hourly basis, to reflect real-time system conditions, wholesale market pricing, or the utility's marginal cost of production.

Rate structure	Hourly day-ahead + 7-day forecast via API; locational (circuit-level); subscription-based; shadow billing; annual true-up credits	Hourly day-ahead via API; locational by 220-kV substation; subscription-based; shadow billing; annual true-up credits
Customer incentives	V2G: up to $3,300 (interconnection) Ag: $160/kW (automation) Res/Com: annual bill credit only	Shadow bill credit for OAT outperformers; no upfront incentive by SCE
Third-party incentives	CCA: $20/kW enrolled load/yr ASP: $42/kW managed load/yr	CCA: $20/kW enrolled load/yr ASP: $4.29/kW load/mo. $250K total for ASP enablement

Sources: Tiell (2025); Molnar (2025); CPUC (2024). (Notes: ASP = Automation Service Provider; C&I = Commercial & Industrial; CCA = Community Choice Aggregator; HVAC = heating, ventilation, and air-conditioning; OAT = Otherwise Applicable Tariff.)

1.3 Learning from Pilot Evaluation Is Critical for Scaling Dynamic Pricing

As mentioned earlier, the CEC load shift goal implies that dynamic-pricing-based flexible load needs to grow from the 2022 level of 30 MW to roughly 2 GW by 2030. It is worth noting that the 311 MW and 150 MW enrolled in PG&E's and SCE's expanded pilots, respectively, can be described as a less-defined "controllable load" rather than measured load shift. Therefore, the current progress toward the policy goal is lower compared to these enrollment numbers. Moving from pilots to full programs will be required to close this gap, and extracting and implementing lessons from the pilots will be critical for programs to succeed.

The IOUs have already been working on extracting lessons from the current pilots under CPUC's guidelines. The key evaluation elements outlined in Decision 24-01-032 include: (1) the response of customer loads to prices; (2) monthly bill impacts compared to the otherwise applicable tariff; (3) recovery of generation and resource adequacy costs for pilot customers; (4) recovery of delivery costs for pilot customers; (5) the number of participating customers, shiftable load enrolled, and shadow-bill credits for specific population; and (6) the impact on emissions and lessons learned regarding dynamic-rate designs that provide benefits specific population. D.24-01-032 also requires that the IOUs use an independent third-party evaluator for the pilot M&E; the CPUC be consulted in developing the M&E plan; and the final evaluation of the expanded pilot be completed by March 1, 2028.

Under these guidelines, the two IOUs are at different stages of developing detailed midterm M&E plans for the expanded pilots, which are not published yet. However, the Advice Letter (AL) 6850 titled M&E Study Proposal for PG&E's Real-Time Pricing (RTP) Pilots (PG&E, 2023) provides a proxy for PG&E's midterm M&E plan for the current HFP pilot, which was developed subsequently referencing it and another Advice Letter 6587-E (PG&E, 2022). AL 6850-E is considered a reasonable representation of the state-of-the-practice of dynamic pricing pilot evaluation framework, which involved a public workshop as part of its development process. It not only includes all objectives from D.24-01-032 required evaluation elements but also a dedicated objective around customer research addressing key research questions below using marketing/outreach metrics, customer enrollment and survey data.

- What are the factors (drivers/barriers, risks/benefits) that influenced customers' decision to enroll or not enroll in the pilot?
- What appliances or technologies did customers use to shift loads?
- What optimizing strategies did customers use to respond to real-time prices?

According to PG&E staff, additional topics are also part of the midterm M&E discussion, including: tracking interoperability standards (OpenADR, IEEE 2030.5, ISO 15118) as part of data collection; analyzing business model maturity and effectiveness for technology use cases; benchmarking program costs against demand response program cost; evaluating technologies' ease of use and access, automation strategies, revealed drivers/barriers, and satisfaction; and technology readiness and compatibility of platforms used to implement the pilot.

2. Gap in Current Pilot Evaluation Path and New Evaluation Lens

The above additional topics move the key question from "Can it work?" to "How does it work?" but still not "Will it scale and endure?" Dynamic pricing is arguably more complex than any existing rates for residential and small business customers, and its long-term success will depend on structural market conditions — not just near-term pilot performance. Several studies have identified lasting barriers to adoption and scaling, such as transaction costs, performance uncertainties, lock-in risks, and limited competition (York et al. 2022); onboarding friction, multivendor troubleshooting, and ecosystem fragmentation (Pereira & Regnier 2025, Vols. II & III); and device incompatibility and uncoordinated single-device architectures (Daly et al. 2026; Culler et al. 2026). Therefore, customer acceptance and satisfaction must be understood at a structural level — assessing price communication and control system architecture for long-term customer experience rather than measuring satisfaction percentage for a specific implementation. Table 2 summarizes the key gaps and evaluation criteria proposed in this paper to address them.

Table 2. Evaluation gaps and proposed criteria for assessing long-term scale-up potential

Gap in current evaluation	Why it matters for scaling	Criterion that addresses it	Reference
Load response measured but not site-level controllable load depth, end-use service impact, or predictability at solution level	Site-holistic load control approaches deliver larger total impact than fragmented approaches over the long term; enrolled MW ≠ reliable flexible load; service impact is key to customer retention; architecture influences reliable signal communication and therefore load response predictability	Controlled load; predictability; service impact	Satchwell et al. (2021, 146); Culler et al. (2026); Varsano et al. (2025); Haupt et al. (2025); Faruqui and Sergici (2010)
Interoperability tracked as adoption metric, not as structural market incentive	Open ecosystem is a prerequisite for scaling; it promotes competition that drives affordability, more product options, and better product designs	Open ecosystem; affordability; customer choice	O'Fallon and Gopstein (2019, 2022); Daly et al. (2026); York et al. (2022)

Customer satisfaction assessed at program level, not at architecture level	Setup complexity determines whether adoption scales beyond early adopters; Troubleshooting accountability across all essential vendors and lock-in risk determine long-term customer retention	Simplicity; customer troubleshooting support; vendor lock-in risk	Faruqui and Sergici (2010, 207-210); Pereira and Regnier (2025, Vol. II, §3–4); Hansen et al. (2025, 15-18)
Connectivity resilience not evaluated	Price-response failure during connection outages is a known operational risk that varies by architecture due to inherent vulnerability differences; negative impact on customer bills can hurt customer retention	Connectivity resilience	Hansen et al. (2025, 22-25); Daly et al. (2026, 11-13); Pereira and Regnier (2025d, Vol. IV, §3); Culler et al. (2026)

To address the gaps identified in Table 2, this paper proposes evaluating dynamic rate pilots at the architectural level — i.e., the pathway through which hourly price signals and responsive control commands reach load-flexible devices in buildings and homes. Four distinct architecture types have been identified through CalFlexHub (see Section 3). This paper proposes (a) test-applying the criteria in Table 2 to show the pros and cons of each type and (b) proposing measurable metrics for each criterion to be considered in utility dynamic rate pilot evaluation.

3. Uncovering the Tradeoffs Between Architectures in CalFlexHub

Four representative system architectures have been selected from CalFlexHub technology demonstration projects and named based on the control ownership and partnership between ASPs and the Original Equipment Manufacturers (OEMs) of load-flexible appliance:

1. **OEM cloud-to-device**: price signal is received by the OEM’s cloud, which optimizes load control and communicates with each OEM device through network;
2. **Cross-cloud open integration**: price signal is received by a third-party ASP’s cloud, which optimizes load control and passes commands onto the OEM’s cloud through API; then OEM further communicates with each device in the network;
3. **OEM full integration**: price signal is received locally by the OEM (e.g., on a local controller), who also fulfills the ASP role by optimizing load control;
4. **Localized orchestration**: price signal is received locally by a third-party ASP’s gateway, which optimizes load control and sends commands to devices.

These four architectures and associated business models are evaluated against a set of 10 criteria refined from Liu et al. (2024), grouped into three categories: load flexibility functions, accessibility, and adverse scenarios. Each architecture/model is rated as “high,” “medium,” and “low” based on one or two technology examples in CalFlexHub. The results are shown in a scorecard-style analytical tool in Table 3. In the remaining part of this section, we synthesize the most important aspects of the evaluation results for each model. Areas colored green in Table 3 highlight criteria scored high or medium for the technology examples. Criteria colored blue or light orange are either positively or inversely correlated, which are explained later in this section.

Table 3. Evaluating four system architectures from CalFlexHub against 10 criteria

System architecture/business model			**OEM full integration**	**OEM cloud-to-device**	**Cross-cloud open integration**	**Localized orchestration**
Example Technology			Integrated heat pump*	EV	Thermostat; HPWH	Smart panel / HEMS
Cat.	**Name of criteria**	**Criteria description**	**Rating**	**Rating**	**Rating**	**Rating**
LF functions	**Controlled load**	Controlled total load per household	**Medium** (1-2 major end uses)	**Medium** (1-2 major end uses)	**Med-High** (API-compatible end uses)	**High** (potentially all end uses)
	Predictability	LF performance is less dependent on user behavior	**High** (set it and forget it)	**High** (set it and forget it)	**Variable** (Low for HPWH; high for thermostat)	**Med-High** (pre-set priority)
	Service impact	Ability to modulate vs. only on/off to minimize service interruption	**High** (modulate with thermal storage or battery)	**High** (modulate with thermal storage or battery)	**Med-High** (modulate by setpoint change)	**Low-Med** (on/off towards modulation)
	Open ecosystem	Inherent need to promote interoperability	**Low** (no need to integrate with other vendors)	**Low** (no need to integrate with other vendors)	**High** (requires integration with other vendors)	**High** (requires integration with other vendors)
Accessibility	**Affordability**	Affordable to customers based on installed and operational costs	**Low** (high installed cost)	**High** (cloud enabling cost is low)	**Medium** (API & device fees)	**Low** (first cost of installation is very high)
	Simplicity	Not requiring high skill level or long time in setup	**Low** (by qualified installer)	**High** (customer can set up)	**Med-High** (customer can set up)	**Low** (by qualified installer)
	Customer choice	The number of available vendors for similar products	**Low** (specialized and unique product)	**Medium** (other car OEMs)	**Med-High** (options for both ASPs & OEMs)	**Med-High** (options for both ASPs & OEMs)
Adverse scenarios	**Connectivity resilience**	Ability to function normally during loss of price signal connection	**Med-High** (can operate on either home WIFI or cellular)	**Low** (dependent on cellular signal; poor garage reception)	**Low** (dependent on multiple clouds and home WIFI)	**Med-High** (positioned for backup using on-device compute and local storage)
	Customer troubleshooting support	Requiring fewer parties involved in troubleshooting	**High** (only OEM involved)	**High** (only OEM involved)	**Medium** (potentially require ASP & OEM)	**Medium** (potentially require ASP & OEM)
	Vendor lock-in risk	Customer's ability to replace service or product without significant loss	**Low** (specialized product, cannot be replaced)	**Low** (specialized product, cannot be replaced)	**High** (ASP & OEM can both be replaced)	**Medium** (OEM can be replaced, but ASP cannot without big loss)

Note: Areas colored green highlight criteria scored high or medium for the technology examples. Criteria colored blue or light orange are either positively or inversely correlated. LF = load flexibility; HPWH = heat pump water heater; HEMS = home energy management system.

OEM full integration. A technology example used for rating this architecture is an integrated heat pump system for space and water heating with an integrated hot water thermal storage tank. Because a single vendor provides an entire solution for homes, it has a major advantage of not relying on a separate ASP to deliver connectivity resilience and streamlined customer troubleshooting support. However, the flip side of the same advantage is that it does not encourage many customer choices or the development of an open ecosystem, and it presents a higher risk of vendor lock-in because of the specialized, fully integrated solution. Other advantages include consistent performance and minimal service impact associated with the integrated thermal energy storage and controls, which, on the other hand, also lowers affordability due to upfront cost and requires specialized professional installation.

OEM cloud-to-device. A technology example used for this is an EV OEM using its cloud to optimize dynamic price and control each car's charging via telematics. This can largely work as “set it and forget it” for the customer with minimal service impact. As an OEM-owned, end-to-end solution, it shares some similar advantages and shortcomings as the OEM full integration model above, streamlining customer troubleshooting support but exposing customers to risks such as fewer choices, closed ecosystem, and vendor lock-in. Compared to the model above, though, the cloud-based approach requires no hardware installation and makes initial setup simple and highly cost-effective.

Cross-cloud open integration. Two technology examples are used: heat pump water heater (HPWH) and connected thermostat. The promises of this model are that (1) it enables integration of potentially all end-use devices in a building that are API-compatible, although they may end up being controlled by multiple rather than the same ASPs; and (2) the API-based integration fundamentally requires interoperability, structurally promoting an open ecosystem where alternative products from many device OEMs and equivalent services from ASPs will compete. The cloud-to-cloud integration approach makes the enabling cost low; however, currently prevailing OEMs tend to charge significant API and device fees to the ASPs, making the solution costly to utilities. In the future, the growing competition should bring such costs down, in theory. There is also motivation for ASPs and OEMs to make the setup simple for the customer. However, these promises are not realities yet, considering the challenges of inconsistent communication protocols, data gaps (e.g., real-time device status), and fragmented coverage of ASP-device integration noted by utilities, all of which makes the current implementation burden high. Furthermore, in this model, the connectivity is inherently dependent on the uptime of multiple vendors’ cloud services and home Wi-Fi; multiple vendors’ support and cooperation may be needed in troubleshooting. These are significant risks that should be considered in customer experience. Other load flexibility function aspects such as performance predictability and service impact will depend on the technology type — for example, load flexibility from HPWHs is typically much less subject to user cooperation than thermostats due to less impact on service.

Localized orchestration. The smart panel and home energy management system (HEMS) are often offered as a combined technology package. The solution demonstrated in CalFlexHub is not a production system, nor does an existing production system have incorporated dynamic price response feature. Therefore, the ratings for this model are partially based on projected architecture and business model. This orchestration model potentially enables most end-use devices in a home by either wire or API-integration, therefore inherently promoting an open ecosystem, which makes it very promising. For this reason, it brings similar advantages to the cross-cloud open integration model above in terms of promoting innovation and market competition, leading to more customer choices and lower vendor lock-in risk on the device side. However, the significant upfront cost of the smart panel/HEMS still presents significant lock-in risk. Given the complexity of connecting different types of devices, the setup requires significant expertise and effort; troubleshooting may require cooperation from multiple vendors. Given the local computing and storage, it is positioned to offer robust connectivity backup. The holistic orchestration may allow delivering a load reduction target based on a user-predefined load shed priority order. A downside of the earlier smart panel technology is that it relies on on/off control of loads, which creates bigger service impact compared to the API-based setpoint adjustment approach. However, more recent leading smart panel technologies are already capable of the latter approach, although those devices connected with the panel/HEMS via API are not considered reliably controlled as the panel-wired loads are, for the same reasons discussed in the cross-cloud open integration model.

Through the above rating exercise based on evidence from CalFlexHub research, it was evident that no single system architecture or associated business model has advantages across all aspects. In fact, being rated high for some criteria (colored blue in Table 3) almost certainly leads to low ratings on some other criteria (colored light orange); in other words, there are inherent tension among some of the criteria. Other criteria not colored are considered less inherent to the architecture but more dependent on the subject technology. In summary, the fundamental tradeoffs surfaced from the above evaluation results are:

1. The two vertically integrated, closed-system solutions provided by an OEM that also serves as its own ASP streamlines setup and troubleshooting support but leaves customers fewer choices and a higher lock-in risk;
2. The two open integration solutions both promote a “many-to-many” ASP-to-OEM mapping possibility, inherently increasing controlled load flexibility while propelling innovation and open market competition in favor of greater customer choices, lower cost, and reduced lock-in risks. However, both are subject to complexity and implementation burden from integrating various types of devices, fragmented customer support responsibilities, and likely less connectivity resilience.
3. Further comparing with the cross-cloud open integration model, localized orchestration offers more holistic coordination of more types of devices via both wiring and API, but it also requires a much higher upfront investment as a tradeoff.

Among the four models, cross-cloud open integration is currently the furthest developed and likely will have more data points to be collected in the utility pilots. On the other hand, localized orchestration is the most nascent model and will likely need to be evaluated based on lab results and evidence-backed projections, to some extent.

4. Informing Pilot Evaluation Metrics

Once the structural pros and cons of each architecture/model and the underlining tension among the 10 criteria is explained, it would be useful to develop metrics to guide data collection from the utility pilots to further understand how the above conceptual analysis transpires in real pilot customers' experience. Below are proposed program-level metrics under the three categories introduced in Section 4 for a utility dynamic rate pilot evaluation team's consideration and adjustments. All metrics should be measured at the customer site level and aggregated to the architecture level for comparison, unless otherwise stated.

4.1 Load Flexibility Functions

Controlled load. Average kilowatt of controllable load enrolled per site, disaggregated by architecture type and end-use device category; this allows direct comparison of load-shift depth across architecture types in the pilot.

Predictability. Coefficient of variation[6] of load shift across daily high-price period of interest (requires time series of site load response against baseline[7]); rate of customer-initiated overrides per customer per year.

Service impact. Rate of customer-reported service impact per year; share of load response control delivered via modulation (e.g., setpoint change) vs. on/off.

Open ecosystem. Number of distinct OEM device brands and ASPs known to have the ability to participate in the pilot in addition to those already participating; number of open-standard protocols (OpenADR, IEEE 2030.5, CTA-2045) actively in use.

4.2 Accessibility

Affordability. Full incremental installed cost relative to a reference alternative solution[8] combined with ASP/OEM annual service fee (counted for the lifetime of the solution or a certain number of years, at the utility's discretion), normalized by either kilowatt of controllable load or customer bill savings.

Simplicity. Hours of self-reported customer time required for setup (no higher education should be required); share of enrollments requiring professional installation; enrollment drop-off rate during onboarding, reflecting interested customers who gave up.

---

[6] Coefficient of variation = (Standard deviation of load response) / (Mean load response) × 100%
[7] Creating accurate counterfactual baseline methods is a known challenge for dynamic pricing, not expanded here.
[8] Examples: for integrated heat pump systems, this is premium over a standard heat pump installation; for smart panel, this is premium over a standard panel upgrade if unavoidable

Customer choice. Number of alternative ASP vendors on the market (as opposed to participating in the pilot) offering equivalent essential product/service to which price response feature can be added quickly at a non-substantial cost.[9] This metric allows some degree of projection of growth to reduce bias against emerging architectures.

4.3 Adverse Scenarios

Connectivity resilience. Multiple key factors play into this criterion, making it complex to directly measure: number of separate cloud services required; communication media redundancy (e.g., home vs. dedicated Wi-Fi, cellular); reliable fallback operation, including executing prestored price schedules or default load control logic, during loss of price signal connectivity; and vulnerability in local network and reception. Utilities should evaluate the overall ability to continue price response in multiple common failure modes (up to a few hours) and come up with an overall rating.

Customer troubleshooting support. Average number of vendor parties required to resolve customer reported price response problems; average days of customer waiting time before reported issues are resolved.

Vendor lock-in risk. Number of alternative ASP vendors on the market that can step in and replace the price-response service without leaving the customer significant stranded equipment (e.g., heat pump, EV, electric panel). This metric will be in favor of the cross-cloud open integration model and likely present challenges for the other three architectures.

5. Discussion

5.1 A Challenging Interoperability Reality

Utilities and state policymakers recognize that creating an open technological ecosystem with widespread interoperability is essential to scaling load flexibility because it reduces implementation friction, promotes innovation and market competition, and favors greater customer choices, affordability, and reduced lock-in risks (O'Fallon and Gopstein 2019 & 2022). However, the path to true, functional openness has been a challenging one as experienced by the IOUs, with challenges including:

- **Data gaps and format inconsistencies.** OEMs and aggregators are not always able to provide data in the format or level of granularity needed for program functionality needs. For example, certain aggregators can only provide thermostat set point data and are unable to supply the actual energy consumption or response data needed for measurement and verification purposes.
- **Stated vs. actual capability mismatches.** There is often a gap between the resources that OEMs and aggregators represent during program development or contracting and what they

[9] For example, for products that can already handle demand response events, it is reasonable to assume that price response feature can be enabled by the vendor without a major investment if market incentive signal is clear enough.

are able to deliver when dispatched. These mismatches introduce program risk and reduces impacts.

- **Lack of universal protocol adoption.** Communication standards such as OpenADR and IEEE 2030.5 continue to evolve, but neither have been universally adopted across the vendor ecosystem between OEMs and ASPs. This inconsistency complicates integration efforts and increases the technical burden of onboarding new partners.
- **Fragmented device and vendor coverage.** Vendor capabilities are often siloed within specific device categories (e.g., EVs, thermostats, HPWHs) or brands, resulting in incomplete coverage across a customer's full suite of assets (Varsano 2025). This fragmentation limits the potential program scale and flexible design.
- **Limited API integration maturity.** The relatively limited depth and standardization of API integrations across third parties and OEMs makes it difficult to achieve seamless, scalable onboarding. Gaps in integration capability slow time-to-market, increase manual intervention requirements, and create inconsistent customer onboarding experiences.
- **Limited marketing value from third-party ASPs.** Third-party ASPs that lack an established OEM-based customer relationship generally have limited ability to drive enrollment. OEMs are “stickier” with the customer due to the original underlying service.

Given these challenges, utilities should not aim to find one "best" architecture but need to actively prioritize and manage a strategic growth portfolio, making strategic bets and investments to mitigate specific risks while improving readiness and scalability.

5.2 Interpreting the Scorecard for Program Design

The scorecard exercise presented in Section 4 using four representative architectures from CalFlexHub is not meant to either capture all system architectures that may be worth considering or deliver refined/final ratings on those four architectures. The intended ways for utilities to use this paper in their pilot evaluation are:

1. Add measurable metrics from Section 5 (refine as needed) to Table 3;
2. Update ratings in Table 3 using measurable metrics based on pilot-collected data;
3. Add other models or technology examples from the pilot;
4. Decide how the utility would like to weigh the 10 criteria — in other words, how the utility views the importance of multiple inherent tradeoffs discussed in Section 4;
5. Place the criteria in parallel to other regulatory or otherwise required pilot evaluation criteria as the utility sees fit in comparing different pathways for scale-up.

There are a few things to be mindful about in evaluating results and weighing the tradeoffs. First, special attention should be given to the adverse scenarios category, because these can potentially create significant setbacks in dynamic pricing adoption if systematic negative customer experience becomes an issue (e.g., stranded assets when vendors of closed system solutions exit the market, prolonged waiting for troubleshooting service). Second, low ratings in any criteria should be carefully reviewed to understand if high ratings in other areas are enough to justify the risk on the low-rated criteria. Those four uncolored criteria in Table 3 are more indictive of the underlying end-use technology, while the light-blue- and light-orange-colored criteria tend to be structurally representing "two sides of the same coin." Third, some scoring criteria (e.g., customer choice, vendor lock-in risk) are biased toward commercialized solutions and against emerging solutions. Utilities should gauge adjustments in metrics definitions and weighting factors to account for this bias. Last but not least, it is important to reiterate that no single architecture is best. A desirable outcome from incorporating these criteria into a pilot evaluation is a strategically prioritized portfolio that balances proven and emerging architectures.

6. Conclusion and Path Forward

Dynamic rates are an important part of California's overall strategy in achieving the 7 GW load shift goal by 2030. The two largest CA IOUs — PG&E and SCE – have both launched their expanded dynamic rate pilots since late 2024 and mid-2025, respectively, and they have well exceeded their enrollment targets. However, compared to the 2030 goal, the flexible load enabled by dynamic pricing will need to grow 67-fold from the 2022 level. This requires moving from pilots to full programs and making strategic decisions to scale up fast. This paper closely reviews regulatory documents of pilot M&E plans and suggests adding a new lens in evaluation to include the perspective of long-term implications on business model, customer experience, and trajectory of the ecosystem.

A set of 10 criteria in three categories — load flexibility functions, accessibility, and adverse scenarios — are proposed to evaluate the pros and cons of each common system architecture and associated business model. These criteria are test-applied to four architecture-and-technology examples in Table 3, which serves as a starting point scorecard. A next step is to add technology examples in each architecture category to gain a fuller and more accurate picture of each architecture for its strengths and weaknesses. The scorecard results uncovered some fundamental tradeoffs: OEM-integrated architectures deliver simpler customer experience but constrain market competition and elevate lock-in risk; open integration architectures structurally promote interoperability and customer choice but carry higher implementation and support complexity and connectivity vulnerability. No architecture dominates all criteria.

More data-informed insights about these underlying tensions will be gained from the IOU dynamic rate pilot evaluations and from the experience of the stakeholders. As the pilots progress, the utilities, OEMs and ASPs will gain many "lessons learned" that will address the challenges of interoperability, data gaps and lack of full resource capabilities, and ultimately develop a robust set of criteria that will facilitate a scalable model of program development and design to meet the future needs of California's demand flexibility goals.

Acknowledgements

This work was supported by the California Energy Commission's Electric Program Investment Charge. We are grateful to the CEC's Matt Fung for his support of the California Load Flexibility and Deployment Hub (CalFlexHub) program. This work was supported by the

Assistant Secretary for Critical Minerals and Energy Innovation, Building Technologies Office, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Generative AI tools used: Gemini, Claude. Used for: Gemini was used to generate the initial paper structure, revised by author. Claude was primarily used to as a thought partner in formulating metrics definitions in section 5, revised by author. Claude was also used to generate fodder text for pilot description, carefully reviewed by utility co-authors. All literature review, methodology development, technical analysis, and final manuscript content were carefully produced and reviewed by the authors. Extent: Preliminary outlining, generating metrics ideas, creating fodder text in targeted sections.

References

CAISO, CPUC, and CEC. (2021). Final Root Cause Analysis: Mid-August 2020 Extreme Heat Wave. https://www.caiso.com/Documents/Final-Root-Cause-Analysis-Mid-August-2020-Extreme-Heat-Wave.pdf

CEC (California Energy Commission). (2022). 2022 Load Management Standards Rulemaking (Docket 21-OIR-03), adopted October 12, 2022, effective April 1, 2023. Accessed March 2026. https://www.energy.ca.gov/programs-and-topics/topics/load-flexibility/load-management-standards

———. (2023). Commission Report: Senate Bill 846 Load-Shift Goal Report, Docket 21-ESR-01, May 2023. https://efiling.energy.ca.gov/GetDocument.aspx?tn=250357&DocumentContentId=85095

CPUC (California Public Utility Commission). (2021). Decision 21-12-015. https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M428/K821/428821475.PDF

———. (2022a). Advanced Strategies for Demand Flexibility Management and Customer DER Compensation – Energy Division White Paper and Staff Proposal (June 2022). https://www.cpuc.ca.gov/-/media/cpuc-website/divisions/energy-division/documents/demand-response/demand-response-workshops/advanced-der---demand-flexibility-management/ed-white-paper---advanced-strategies-for-demand-flexibility-management.pdf

———. (2022b). Order Instituting Rulemaking to Advance Demand Flexibility Through Electric Rates (R.22-07-005), filed July 14, 2022. https://www.cpuc.ca.gov/industries-and-topics/electrical-energy/electric-costs/demand-response-dr/demand-flexibility-rulemaking

———. (2024). Decision 24-01-032: Expansion of Demand Flexibility Pilots for System Reliability (R.22-07-005). https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M523/K972/523972750.PDF

Culler, M., R. Stolworthy, R. Edsall, P. Heeter, A. Tylecote, J. Kmiec, and T. Ponto. (2026). *Virtual Power Plant Architecture and Resilient Design* (January 2026). INL/RPT-26-89692. Idaho National Laboratory. https://inl.gov/content/uploads/2026/01/VPP-Architecture-and-Design-Report__Final.pdf

Daly, M., C. Briggs, E. Langham, and S. Craft. (2026). *Energy Masters: Lessons Learnt Report 2* (January 2026). Prepared for the Australian Renewable Energy Agency and RACE for 2030. SA Power Networks. https://arena.gov.au/assets/2026/02/SA-Power-Networks-Energy-Masters-Lessons-Learnt-Report-2.pdf

Faruqui, A., and S. Sergici. (2010). "Household response to dynamic pricing of electricity: a survey of 15 experiments." Journal of Regulatory Economics, 38(2), 193–225. https://doi.org/10.1007/s11149-010-9127-y

Hansen, D., M.T. Clark, and C. Goodrich. (2025). *Final Evaluation of Southern California Edison's Dynamic Rate Pilot* (February 28, 2025). https://www.dret-ca.com/wp-content/uploads/2025/03/PUBLIC-SCE-DRP-Final-Evaluation-Report.pdf

Haupt, T., J. Jungwirth, G. Hofmann, and H. Vaidya. (2025). *Home Energy Management Systems (HEMS): 2nd Market Overview for Germany* (2025). PV-Symposium Proceedings, 2. https://doi.org/10.52825/pv-symposium.v2i.2641

Liu, J., M.A. Piette, M. Pritoni, B. Nordman, P. Grant, S. Smith, R. Brown, J. Bourg, F. Godinez, M. Fung, and M. Martinez. (2024). "Assessing Customer Experience and Business Models around Price-to-Device Communication and Smart Control Pathways in CalFlexHub." 2024 ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA.

Molnar, E. (2025). "SCE's Flexible Pricing Rate Pilots." Presented at the Peak Load Management Alliance (PLMA) 2025 Fall Conference, Henderson, NV, November 3–5, 2025.

O'Fallon, C., and A. Gopstein. (2019). *Economics of Interoperability in the Context of Smart Grid Architectures*. NIST Special Publication 1900-604. Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.1900-604

———. (2022). "Reducing barriers to entry and hedging against obsolescence with smart grid interoperability." Competition Policy International. https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=935801

Piette, M.A., J. Liu, B. Nordman, S. Smith, R. Brown, M. Pritoni, B. Gerke, A. Yee, M. Hart, and M. Fung. (2022). "Accelerating Decarbonization with the California Load Flexibility Research and Deployment Hub." 2022 ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA.

PG&E (Pacific Gas and Electric Company). (2022). Advice Letter 6587-E: Dynamic Rate Pilot Program Evaluation Plan (May 6, 2022). https://www.pge.com/tariffs/assets/pdf/adviceletter/ELEC_6587-E.pdf

———. (2023). Advice Letter 6850-E: Measurement and Evaluation Study Proposal for PG&E's Real-Time Pricing (RTP) Pilots, Pursuant to D.22-08-002 (January 30, 2023). https://www.pge.com/tariffs/assets/pdf/adviceletter/ELEC_6850-E.pdf

Pereira, G. and C. Regnier. (2025). *Business Models for Scaling Demand Flexibility* (Vols. I–IV). Lawrence Berkeley National Laboratory. Vol. I: https://doi.org/10.20357/B7BG7J; Vol. II: https://doi.org/10.20357/B76P5M; Vol. III: https://doi.org/10.20357/B73031; Vol. IV: https://doi.org/10.20357/B7Z60Q

Satchwell, A., M.A. Piette, A. Khandekar, J. Granderson, N. Mims Frick, R. Hledik, A. Faruqui, L. Lam, S. Ross, J. Cohen, K. Wang, D. Urigwe, D. Delurey, M. Neukomm, and D. Nemtzow. (2021). *A National Roadmap for Grid-Interactive Efficient Buildings* (May 2021). Lawrence Berkeley National Lab, Berkeley, CA. https://doi.org/10.2172/1784302

Tiell, O. (2025). "PG&E’s Hourly Flex Pricing Pilots." Presented at the Peak Load Management Alliance (PLMA) 2025 Fall Conference, Henderson, NV, November 3–5, 2025.

Varsano, J. (2025). "Unleashing the demand-side revolution: The case for a unified VPP platform." Utility Dive, June 12, 2025. https://www.utilitydive.com/spons/unleashing-the-demand-side-revolution-the-case-for-a-unified-vpp-platform/750629/

York, D., S. Nadel, and S. Subramanian. (2022). *U.S. and International Experience with Market Transformation*. ACEEE White Paper. American Council for an Energy-Efficient Economy. June 2022. https://www.aceee.org/white-paper/2022/06/us-and-international-experience-market-transformation