

DOI: 10.5281/zenodo.12426110

# REGIONAL DISPARITIES IN SOLAR PHOTOVOLTAIC INSTALLATION COSTS: A MULTI-STATE ANALYSIS OF PRICING MECHANISMS AND SCALE ECONOMIES

Kshitiz Raj<sup>1\*</sup>, Michael Walton<sup>2</sup>

<sup>1</sup>*Energy Finance Specialist, Redaptive Sustainability Services, Denver, Colorado, United States.*

<sup>2</sup>*Partner, Precursor, Chattanooga, Tennessee, United States.*

<sup>2</sup>*Former Member, White House Environmental Justice Advisory Council (2024–2025).*

Received: 01/12/2025

Accepted: 02/01/2026

Corresponding author: Kshitiz Raj

([kshitiz.raj@alumni.duke.edu](mailto:kshitiz.raj@alumni.duke.edu))

## ABSTRACT

This study examined whether economies of scale and regional variation explain observed differences in reported cost per watt in public-facing commercial solar PV records, and to what extent reporting-boundary inconsistency distorts benchmarking. A dataset of 164 project records was assembled from publicly available or publicly reference materials across California (n = 96), Illinois (n = 60), and New Jersey (n = 8). For each record, reported cost per watt was computed mechanically as total reported cost divided by reported system size in watts, and patterns were evaluated using descriptive statistics and state-level and pooled scatter plots of cost per watt versus system size with linear trend diagnostics. Mean reported cost per watt differed materially by state: California averaged 3.88 \$/W, Illinois 2.44 \$/W, and New Jersey 3.49 \$/W. In pooled analysis, cost per watt declined directionally with system size, but the relationship was weak in a simple linear specification ( $R^2 = 0.0585$ ), indicating that system size alone explained little of the observed variance across mixed project types and reporting conventions. Dispersion was highest at small-to-mid system sizes and persisted even for multi-megawatt projects, with California exhibiting the widest spread and most high-cost outliers. Configuration labels aligned with separable cost bands, with canopy-associated records more frequently occupying higher cost per watt regions than flat-roof and many ground-mounted systems. Repeated-value banding and portfolio entries suggested standardized pricing or blended reporting practices that can compress or inflate apparent costs. These findings indicate that public-record cost per watt is an informative market signal but an unstable benchmark unless stratified by configuration and reporting boundary. The study supports minimum scope-disclosure standards for public cost reporting and segmentation-based benchmarking for developers, financiers, and policymakers.

---

**KEYWORDS:** Cost per watt; Solar PV benchmarking; economies of scale; Regional cost variation; Reporting boundary inconsistency; Commercial and industrial solar.

---

## 1. INTRODUCTION

### 1.1. Background

Solar is now a mainstream infrastructure asset, but cost remains the gating variable for how fast, how widely, and where it gets built. On the demand side, corporate buyers, municipalities, and school districts are still highly sensitive to delivered price because electricity savings must clear internal hurdles that compete with core business capital needs. On the supply side, developers and EPCs live in a world where schedule risk, interconnection uncertainty, permitting friction, and labor availability can quietly dominate the final installed cost, even when module and inverter pricing looks favorable. This is why Cost per Watt persists as a blunt but durable industry language. It compresses a complex set of technical and commercial choices into a number that lets stakeholders compare projects, negotiate contracts, and size financing, even though everyone in the room knows the number is never fully comparable across deals.

The irony is that the more solar matures, the more the comparability problem matters. Large benchmark efforts have made major progress in explaining cost structure through consistent definitions and bottom-up modeling. The NREL cost benchmark framework, for example, separates hardware and soft costs and provides national-average bottom-up estimates with transparent assumptions (Ramasamy et al. 2023). Berkeley Lab's work complements that approach by analyzing empirical, project-level datasets at scale, including utility-scale plants and distributed PV pricing trends, helping the industry distinguish cost fundamentals from short-run market noise (Bolinger et al. 2024; Barbose et al. 2024). These references define the center of gravity for how the sector talks about installed cost.

But practitioners repeatedly run into a practical constraint: real transaction data that circulates in pipelines, teasers, acquisition packages, public

finance documents, and partner-provided summaries often uses accounting boundaries that do not match benchmark boundaries. Some datasets are effectively EPC price only. Others include EPC plus structural upgrades, electrical upgrades, or interconnection work. Others fold in developer overhead, margin, financing fees, and gain-on-sale economics in ways that are rarely labeled cleanly. The same label, project cost, can therefore describe materially different underlying content. That is not a minor technicality. It drives how policymakers interpret reported infrastructure costs, how financiers set covenants and contingency expectations, and how developers benchmark their own competitiveness.

This paper starts from that messy but real world condition and asks a simple industry-facing question: when I take publicly sourced project records that resemble the way many market participants first encounter a deal, what do apparent Cost per Watt patterns look like, how much signal is there around system size and geography, and what does the noise itself reveal about reporting practices.

### 1.2. Data Context for This Study

This study analyzes 164 solar project records obtained from public or partner-accessible materials across California, Illinois, and New Jersey. Each record includes at minimum system size in kW and a total cost value used to compute dollars per watt. Many records also include system type (flat roof, canopy, ground fixed, ground single-axis tracker, pitched roof), counterparty quality indicators (credit rating categories), partner identifiers, sales identifiers, and a transaction or pricing date. The dataset is weighted toward commercial and industrial and public-sector style deployments rather than utility-scale plants, and it includes a mix of single-site projects and portfolio entries.

*Table 1: Summary statistics of solar installation costs by state*

State	Number of Sites	Average Cost per Watt (\$/W)
CA	96	3.88
IL	60	2.44
NJ	8	3.49
Total	164	— <sup>a</sup>

<sup>a</sup> Weighted average not computed due to heterogeneity across states.

Two immediate implications follow. First, the sample is not balanced across states, and New Jersey is a small-n segment where any single outlier can meaningfully swing averages. Second, the spread between CA and IL is large enough that the paper

cannot stop at reporting a mean. I have to explain what the mean likely contains, what it likely omits, and why two states can look structurally different even when they are buying many of the same physical components.

### 1.3. Existing Evidence and What It Helps Us Do

At the macro level, global cost declines for solar PV are well documented and remain a central reason solar deployment has accelerated worldwide. IRENA reports continuing competitiveness improvements and provides global cost statistics for utility-scale renewables, including solar PV (IRENA 2024). The IEA similarly documents rapid growth trajectories and identifies barriers that prevent even faster buildout, including non-hardware constraints that increasingly dominate timelines and risk (IEA 2024). Global market tracking by IEA PVPS provides additional context on deployment patterns, supply chain dynamics, and regional policy drivers (Masson et al. 2024). For the US specifically, industry outlooks continue to show large build rates and a pipeline that is sensitive to policy stability, grid constraints, and financing conditions (SEIA and Wood Mackenzie 2025).

At the benchmark and project level, NREL's annual cost benchmarking work offers a disciplined framework that decomposes PV and storage cost into categories and makes assumptions explicit, which is essential for any serious discussion of Cost per Watt (Ramasamy et al. 2023). Berkeley Lab's empirical work provides two important complements: a view of distributed PV pricing trends at scale, and plant-level patterns for utility-scale PV and PV plus storage (Barbose et al. 2024; Bolinger et al. 2024). Taken together, these references form a strong baseline for what installed cost means under consistent definitions and large samples.

The literature on technological learning and experience curves also explains why, over long horizons, larger scale and accumulated deployment correlate with cost declines, even if short-run datasets often look noisy. The foundational learning curve concept dates back decades (Wright 1936). More recent work shows that cost trajectories reflect not only scale but also RD, efficiency gains, supply chain maturation, and process improvements (Kavlak et al. 2018). Research on predictability of technological progress provides useful framing for what kind of signal one should expect in observational cost datasets (Farmer and Lafond 2016). The key point for this paper is not to re-prove learning curves, but to avoid the native mistake of treating cross-sectional scatter in a small multi-state sample as evidence for or against technological learning. Cross-sectional variation can be dominated by scope differences, site constraints, and accounting boundaries even when the long-run

learning story is true.

Finally, policy and regulatory context matters because it shapes both cost and what is recorded as cost. In the US, interconnection and transmission constraints increasingly affect project viability and can add both direct costs and indirect schedule costs, prompting federal efforts to reform generator interconnection processes (Federal Energy Regulatory Commission 2023; Federal Energy Regulatory Commission 2024). Broader system planning assessments also emphasize grid bottlenecks as a practical limiter on deployment (U.S. Department of Energy 2024). On the tax incentive side, the Inflation Reduction Act expanded clean energy credits but also introduced pre-vailing wage and apprenticeship rules that can materially shift labor cost structure for many projects, while providing exemptions for certain small facilities (Internal Revenue Service 2025). These policy mechanics do not just change true cost. They change reported cost, depending on whether labor premiums, compliance overhead, and sponsor-side costs are captured in the data source used.

What this existing evidence does well is establish the centerline: how solar costs are expected to behave under standardized definitions, and what structural forces are pushing deployment. What it does less well, by design, is explain how messy, public deal-level cost numbers behave when accounting boundaries vary and when project types are mixed. That is the gap my dataset can address in a way that is immediately recognizable to practitioners.

### 1.4. Research Gap

There is a persistent mismatch between how cost is defined in benchmark research and how cost is surfaced in many real-world deal records. Benchmarks are careful about categorization and assume a consistent scope. Public deal records often are not. Yet those public numbers still influence decision-making. They show up in early-stage screening, partner conversations, board materials, public finance documentation, and policy discussions about whether solar is cheap or expensive in a given region.

This mismatch creates two blind spots that matter for the industry.

- **First** is interpretive. When experts see a scatterplot of Cost per Watt versus system size and the regression fit is weak, they know not to overread it. But non-expert stakeholders, including some policymakers and generalist investors, often do. They may interpret noise as inefficiency, or they

may interpret a low price as evidence of superior execution, when it may simply reflect a narrower accounting boundary. Without a structured discussion of reporting practices, the same dataset can support contradictory narratives.

- **Second** blind spot is operational. Developers and financiers care about variance as

much as the mean because variance drives contingency, reserves, DSCR resilience, and pricing buffers. If reported costs understate certain categories in some regions or for some project types, teams can miscalibrate risk. That miscalibration shows up later as renegotiation, schedule slips, or underperformance against underwriting assumptions.

My dataset, while smaller than national benchmark datasets, is precisely the kind of dataset that can illuminate these blind spots because it resembles the way trans-action cost information is actually encountered. The small-n nature is not just a limitation. It forces the paper to be honest about signal, noise, and why the noise exists.

### 1.5. Central Research Problem and Research Question

The central research problem is that apparent Cost per Watt values drawn from public deal records are not consistently comparable, which can distort perceived economies of scale and regional cost differences.

The primary research question is: How do apparent solar Cost per Watt values vary across California, Illinois, and New Jersey in a 164-project public-record dataset, and to what extent do system size and project geography explain observed cost dispersion once reporting limitations and scope inconsistencies are explicitly considered?

Two secondary questions support the main question. First, do larger systems in this dataset exhibit lower dollars per watt, consistent with economies of scale, or is that relationship weak once mixed project types and reporting boundaries are present. Second, are observed regional differences in mean Cost per Watt better explained by underlying cost structure differences, by project mix and site constraints, or by systematic differences in what costs are included in public reporting.

### 1.6. Objectives

The first objective is to quantify basic descriptive statistics for Cost per Watt across the full dataset and by state, including mean values and dispersion. The purpose is not to claim a definitive state-level benchmark, but to document what these public-record costs look like and how wide the spread is.

The second objective is to test for economies of scale using system size versus dollars per watt scatterplots and simple regression fits, both for the combined dataset and state subsets. In your combined-state scatterplot, the fitted line slopes downward but the explained variance is low, with R-squared around 0.0585, which is a strong early indicator that size alone explains little of the cross-sectional variance in this dataset. That is not a failure. It is a result that needs interpretation, and it aligns with the idea that scope differences and site-specific constraints can dominate (Ramasamy et al. 2023). It also invites a more nuanced view of scale: scale economies may show up within narrower categories, such as within one system type, one permitting regime, or one contractor set, rather than across a mixed sample.

The third objective is to evaluate regional variation by comparing state-level patterns and discussing plausible structural drivers without overstating causality. California's higher average in your dataset likely reflects a combination of project mix, including canopy-heavy entries that are structurally more expensive, and differences in labor and construction complexity that may or may not be captured consistently. Illinois' lower mean likely reflects a different mix and possibly different cost boundaries reported in the sources used. New Jersey sits between them but is not statistically stable due to sample size. The paper's job is to make these statements carefully and to separate what the data shows from what the paper hypothesizes.

The fourth objective is to make reporting practices a first-class analytic object. Instead of treating inconsistent bundling as a nuisance footnote, the paper treats it as a meaningful finding with direct implications for benchmarking, procurement, and financing. Development costs may be underreported or inconsistently bundled across records, including EPC margin, sponsor overhead, interconnection upgrades, legal and advisory fees, and gain-on-sale economics. These omissions can make projects appear cheaper than they are on an all-in basis, which can mislead readers and distort downstream decisions.

The fifth objective is to translate findings into implications for three stakeholder groups. For policymakers, the key question is how to interpret cost claims in public discourse and how reporting standards might improve transparency. For developers and EPCs, the question is how to benchmark competitiveness without fooling themselves by comparing inconsistent scopes. For financiers, the question is how to set contingency, reserves, and covenants in the face of cost variance

that is not explained by size alone, and how to identify when a low Cost per Watt is a real advantage versus an accounting artifact.

### 1.7. Scope and Limitations

The scope of this study is intentionally narrow in geography and broad in real-world deal heterogeneity. It covers three states with different market characteristics and regulatory environments, and it uses project records that reflect public or semi-public deal-level reporting rather than a standardized survey instrument.

Several limitations are material and must be addressed explicitly. The first is accounting boundary ambiguity. Many records do not specify whether cost includes full EPC scope, interconnection upgrades, structural reinforcement, owner-side costs, financing fees, or developer margin. This is not a trivial limitation, and the paper will not pretend it can be fully corrected without additional primary data. Instead, the study treats this limitation as central evidence of why public cost benchmarking is difficult and why caution is warranted when comparing across projects.

The second limitation is project mix. Canopies, flat roof systems, and ground-mount systems often have meaningfully different structural and civil scopes, which affects dollars per watt. If one state subset contains a higher share of canopy projects, state mean differences can reflect mix more than region. This limitation can be partly mitigated by segmenting the analysis by system type, but any segmentation reduces sample size and statistical stability.

The third limitation is sample imbalance and small-n issues. New Jersey has only eight observations in this dataset, and any inference about New Jersey should be framed as preliminary. Even Illinois and California, while larger, may be clustered around certain partners or channels and therefore not representative of the entire state market.

The fourth limitation is time heterogeneity. The records span multiple dates across 2023 to 2025 in the snippet provided, during which supply chain conditions, interest rates, and labor availability changed. Without a controlled time-series design, the analysis is cross-sectional and descriptive rather than causal.

The fifth limitation is that a simple linear regression of dollars per watt on system size may be structurally mis-specified for this problem. The relationship is likely non-linear, and the variance is likely driven by categorical variables and scope differences that are not fully captured. The low R-

squared in your combined scatterplot is consistent with that reality and should be interpreted as evidence that size alone is not a sufficient explanatory variable in this dataset, not as evidence that scale economies do not exist in general (Kavlak et al. 2018; Farmer and Lafond 2016).

Within these constraints, the paper remains valuable because it does something many papers avoid. It uses a dataset that looks like real market information, shows what patterns do and do not emerge, and uses the limitations as insight rather than as a way to dismiss the exercise. For an expert audience, the contribution is not that the paper discovers that larger systems are cheaper. They already know that as a long-run tendency. The contribution is that the paper demonstrates how easily that tendency can be obscured in public deal data, and it maps the practical implications of that obscuring for procurement, financing, and policy evaluation.

## 2. METHODS AND MATERIALS

### 2.1. Study design and analytic approach

This study used an observational, cross-sectional design to characterize reported cost per watt patterns in commercial-scale solar PV projects across three U.S. states: California (CA), Illinois (IL), and New Jersey (NJ). The objective was not to estimate a causal treatment effect, but to quantify and interpret the empirical structure of publicly reported project cost data as it is commonly encountered by market participants. The analysis focused on two primary relationships: (i) economies of scale, assessed through the association between system size and reported cost per watt, and (ii) regional variation, assessed through state-level differences in distributional statistics and scatterplot structure. A third analytic dimension, treated as a first-order measurement issue, was reporting boundary inconsistency, evaluated through distribution shape, outlier structure, and repeated-value banding in the cost per watt metric.

#### 2.1.1. Data sources and record assembly

The dataset comprised 164 project records derived from public-facing sources and public referencable deal materials used in commercial solar procurement and transactions. These sources included procurement award documentation, project summaries, portfolio acquisition materials, and partner-provided public disclosures. Each record represented either a single site or an individual site entry within a portfolio-level disclosure. Portfolio entries were retained to reflect the

way cost information is frequently disseminated in practice, but they were flagged during cleaning so that their influence could be recognized in variance and banding diagnostics.

The unit of observation was a project record with at minimum: (1) state, (2) system size in kW (DC nameplate as provided), and (3) total reported cost in U.S. dollars. Additional descriptive fields were captured when available, including system configuration labels (flat roof, pitched roof, canopy, ground fixed, ground single-axis tracker), transaction/pricing date, partner label, sales label, and an indicative counterparty credit category. Missingness in descriptive fields did not disqualify a record if the minimum data required to compute reported cost per watt were present.

## 2.2. Inclusion and exclusion criteria

Records were included if they met all of the following criteria:

1. Project location was identifiable as CA, IL, or NJ.
2. System size was reported as a numeric value in kW.
3. Total cost was reported as a numeric value in USD sufficient to compute cost per watt.
4. The record corresponded to a PV installation (standalone PV or PV paired with storage where PV cost could not be separated were retained only if the source represented the reported cost as solar cost; see reporting boundary notes below).

Records were excluded if any of the minimum variables were missing or non-interpretable (e.g., “TBD” cost, size reported without units and not resolvable). Duplicate records were removed when the project identifier and attributes were substantively identical. When the same project appeared multiple times with differing sizes or costs due to scope revisions, the most complete record with explicit size and cost was retained, and earlier partial entries were excluded.

## 2.3. Variable definitions and cost per watt computation

The primary dependent metric was reported cost per watt, calculated for each record as:

$$\text{Cost per Watt (\$/W)} = \frac{\text{Reported Total Cost (\$)}}{\text{Reported System Size (kW)} \times 1000} \quad (1)$$

This calculation converted system size from kW to W to align with standard \$/W reporting conventions. System size was treated as the independent scale variable and remained in kW for

plotting and interpretability.

Because the dataset was derived from heterogeneous reporting sources, the total cost field was interpreted strictly as “reported total cost,” not as a standardized all-in installed cost. No attempt was made to retroactively normalize scope categories (EPC, owner’s costs, interconnection upgrades, roof work, structural reinforcement, developer fee, financing fees, acquisition premiums, or gain-on-sale economics) because these categories were not consistently disclosed. Instead, reporting boundary inconsistency was treated as an empirical characteristic of the dataset and assessed through diagnostic patterns (dispersion, outliers, and banding).

## 2.4. Data cleaning and standardization procedures

All records were standardized into a consistent schema with the following steps:

1. **Unit harmonization:** System sizes were converted to numeric kW values. Where size was presented with separators (commas) or embedded text, values were parsed into numeric form.
2. **Currency normalization:** Cost values were standardized to numeric USD. Currency symbols, commas, and whitespace were removed.
3. **Derived metric creation:** Cost per watt was computed using the equation above.
4. **Sanity checks:** Records were screened for nonphysical or implausible values. This included checking for negative costs, zero sizes, and cost per watt values outside a broad plausibility envelope for commercial PV. Values were not automatically removed based on magnitude alone because scope inclusions (e.g., canopy civil work, roof work, interconnection mitigation, portfolio allocation) can legitimately produce high cost per watt in public reporting. Instead, extreme values were flagged for discussion as outliers.
5. **Portfolio flagging:** Records explicitly labeled as portfolio acquisitions or multi-site portfolios were flagged as portfolio-derived. These records remained in the main analysis because they reflect real-world dissemination of cost information; however, their presence informed interpretation of horizontal banding and repeated \$/W values.
6. **System type normalization:** System configuration labels were mapped into a

standardized set: Flat Roof, Pitched Roof, Canopy, Ground Fixed, Ground SAT (single-axis tracker), and Other/Unknown. The mapping used the label provided in the record without inferring missing configurations.

**2.5. Statistical analyses and visualization methods**

Analyses were executed in a descriptive statistics framework supplemented by simple bivariate regression diagnostics. The goal was to report patterns transparently rather than to overfit models to heterogeneous data.

**2.5.1. Descriptive statistics**

Table 2:

State	Records (n)	Mean reported cost per watt (\$/W)
CA	96	3.88
IL	60	2.44
NJ	8	3.49
Total	164	Pooled distribution analyzed

**2.5.2. Economies of scale assessment**

Economies of scale were evaluated through scatterplots of reported cost per watt against system size (kW). Four scatterplots were generated: one pooled plot (CA+IL+NJ) and one plot per state. Each plot included a fitted linear trendline. For the pooled plot, the coefficient of determination ( $R^2$ ) displayed on the chart was reported as a descriptive indicator of how much variance in reported cost per watt was explained by system size in a linear specification.

The linear trendline was estimated as:

$$\$/W_i = a + \beta \cdot kW_i + \varepsilon_i \quad (2)$$

where  $\$/W_i$  is reported cost per watt for project  $i$ ,  $kW_i$  is reported system size, and  $\varepsilon_i$  captures residual variance, including unobserved scope and reporting differences. The regression was not treated as causal;  $\beta$  was interpreted directionally as a scale association under mixed-scope reporting.

Because the scatterplots indicated heteroskedasticity (wider dispersion at smaller sizes) and potential nonlinearity, the linear fit was used as a minimal, interpretable diagnostic rather than as a definitive model. The analysis therefore emphasized visual structure, density shifts across size, and the stability of the association across states.

**2.5.3. Regional variation assessment**

Regional variation was assessed via (i) differences in state-level mean reported cost per watt, (ii)

For each state and for the full sample, the following descriptive statistics were computed for cost per watt: count (n), mean, and distributional shape assessment based on scatterplot density and outlier presence. Although median and quantiles can provide additional robustness, the primary summary values reported in the study emphasized mean values to align with common industry benchmarking practices, while explicitly interpreting those means as reported-cost indicators rather than normalized installed-cost benchmarks.

A compact summary table was produced to report the number of sites and mean cost per watt by state:

differences in dispersion and outlier structure across state scatterplots, and (iii) differences in apparent clustering patterns. The state plots were interpreted as state-specific distributions of public-report cost signals rather than as controlled samples of installed-cost reality. New Jersey results were treated as illustrative because the state sample was small ( $n = 8$ ), which increases sensitivity to individual observations.

**2.5.4. Reporting boundary diagnostics**

Reporting boundary inconsistency was evaluated using three observable diagnostics:

1. **Dispersion disproportion:** large vertical spread in cost per watt at similar sizes, particularly at small-to-mid system sizes.
2. **Outlier structure:** the presence of a high-cost tail that may reflect scope-intensive projects or broader cost inclusion.
3. **Horizontal banding:** repeated or near-repeated cost per watt values across multiple records, consistent with standardized pricing templates, blended portfolio allocation, or repeated disclosure conventions.

These diagnostics were used to interpret why simple size-based models showed limited explanatory power and why cross-state comparisons based solely on means can be misleading without scope stratification.

**2.6. Treatment of bias, replicability, and integrity controls**

To minimize analyst-driven bias, the study

followed a rule-based transformation pipeline: standardize units, compute cost per watt mechanically, and present results using a consistent plotting template across pooled and state-level views. No project records were removed solely because they were high or low cost per watt; instead, such records were retained to preserve the empirical reality of public reporting and were flagged for interpretation as potential scope-intensive or boundary-inclusive observations.

Replicability was supported through explicit variable definitions, a deterministic cost per watt formula, and consistent inclusion criteria. A researcher replicating the study would require access to the same public sources or equivalent record extracts containing size and cost fields. Because public sources can evolve, the study treated each record as a snapshot of what was available and interpretable at the time of compilation, with the understanding that public reporting practices can change.

### 2.7. Ethical consideration

The study used project-level records obtained from public-facing materials and did not involve human subjects, private personal data, or non-public proprietary disclosures. As such, formal institutional review board approval was not required. The analysis was presented at an aggregate and project-record level, focusing on cost metrics and not on sensitive operational details. Where partner or counterparty labels were included in the dataset schema, the purpose was to support internal segmentation and interpret banding patterns rather than to evaluate any specific entity's performance.

The methodological boundary of the study was intentionally aligned to the research question: to characterize how reported cost per watt behaves in public deal records and what that implies for benchmarking. The study did not attempt to infer true installed cost for each project, nor did it normalize costs for inflation, financing conditions, or supply chain indices. These extensions are feasible and valuable, but they require additional structured metadata and scope disclosure that were not consistently available in the source materials used for this dataset.

## 3. RESULTS

### 3.1. Sample composition and summary statistics

The final dataset comprised 164 commercial-scale solar PV project records drawn from publicly available or publicly referencable deal materials across California, Illinois, and New Jersey. California contributed 96 records, Illinois contributed 60 records,

and New Jersey contributed 8 records. Each record contained a reported system size (kW) and a total reported project cost, which were used to calculate a reported cost per watt (\$/W) for cross-project comparability. Where available, records also contained system configuration labels (flat roof, pitched roof, ground fixed, ground single-axis tracking, canopy), transaction or pricing dates, and counterparty or partner metadata.

State-level averages computed from the dataset indicated mean reported \$/W of 3.88 for California, 2.44 for Illinois, and 3.49 for New Jersey, with the interpretation of these means conditioned on mixed system types and mixed reporting boundaries within and across states. The state means were accompanied by visibly different dispersion patterns in the scatterplots, with California exhibiting the broadest range, Illinois the tightest cluster, and New Jersey the smallest sample with a steep apparent size gradient.

### 3.2. Distributional characteristics of reported cost per watt

Across the full sample, reported \$/W values spanned a wide interval, and the dispersion did not appear symmetric around the mean. Visual inspection of the combined scatterplot indicated a dense concentration of observations between approximately \$2.0/W and \$3.2/W at smaller and mid-sized systems, alongside a meaningful tail extending above \$4.5/W and reaching into the \$6 to \$7/W range. The upper tail was disproportionately represented by California records and appeared to include a higher share of canopy-labeled projects and other scope-intensive records. Illinois exhibited fewer high-cost points, with most observations residing in a narrower band in the low-to-mid \$2/W range and relatively few observations exceeding \$3.0/W. New Jersey displayed a moderate spread relative to its small sample size, with points ranging from the mid-\$2/W range to the mid-\$4/W range.

A recurring pattern in the dataset was the presence of repeated or clustered \$/W values across multiple records, particularly within certain portfolios or partner groupings. This clustering was observable as horizontal bands in the scatterplots, suggesting that subsets of records used standardized pricing templates, blended portfolio costs, or uniform allocation methods across multiple sites. These banding features were most visible in the combined plot and the California plot, and they remained observable after visually segmenting by system size.

### 3.3. Economies of scale assessment using cost per

### watt versus system size

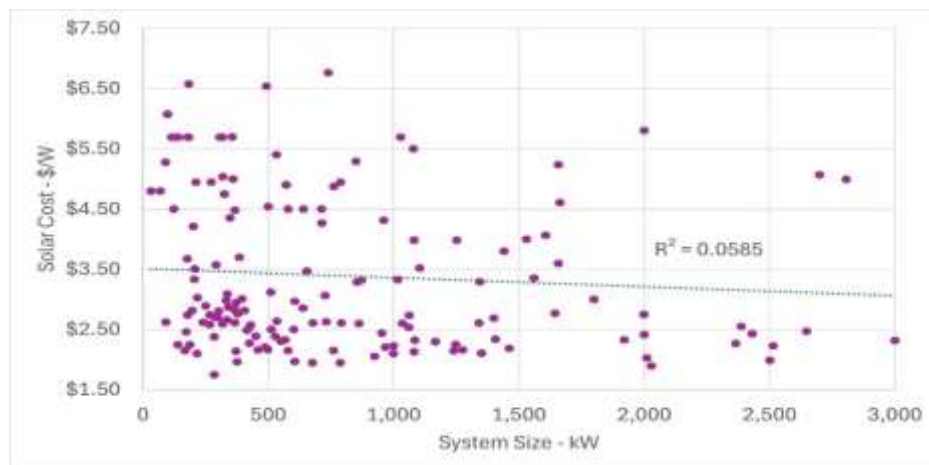
Economies of scale were evaluated using scatterplots of reported \$/W against system size (kW) for the combined dataset and for each state subset. In the combined scatterplot for California, Illinois, and New Jersey, the fitted linear trendline sloped downward, indicating lower reported \$/W at larger system sizes on average. However, the fit remained weak, with the displayed coefficient of determination indicating low explanatory power ( $R^2 = 0.0585$ ). The combined plot therefore indicated that system size alone accounted for a small share of cross-project variance in reported \$/W within this mixed-scope, mixed-configuration dataset.

Despite the weak global fit, the data showed a

visually identifiable shift in the central tendency of points as system size increased. Observations below roughly 500 kW contained a dense mix of projects in the \$2 to \$3/W band plus a sizable number of higher-cost projects, particularly in California. Observations above roughly 1,000 kW displayed a larger share of points clustered in the low-to-mid \$2/W range, especially in Illinois, with fewer observations in the highest cost bands. At the largest sizes shown on the plots, a subset of observations remained near or above \$3/W, indicating that cost dispersion persisted even as system size increased.

### 3.4. Interpretation of the combined scatterplot

The combined scatterplot displayed three outcomes that were consistently observable across the plotted range.



**Figure 1: Relationship between system size and cost per watt across California, Illinois, and New Jersey. The scatter plot shows considerable price variation with a weak negative correlation ( $R^2 = 0.0585$ ), indicating that system size explains only 5.85% of cost variation. Data includes 164 installations: 96 in CA, 60 in IL, and 8 in NJ.**

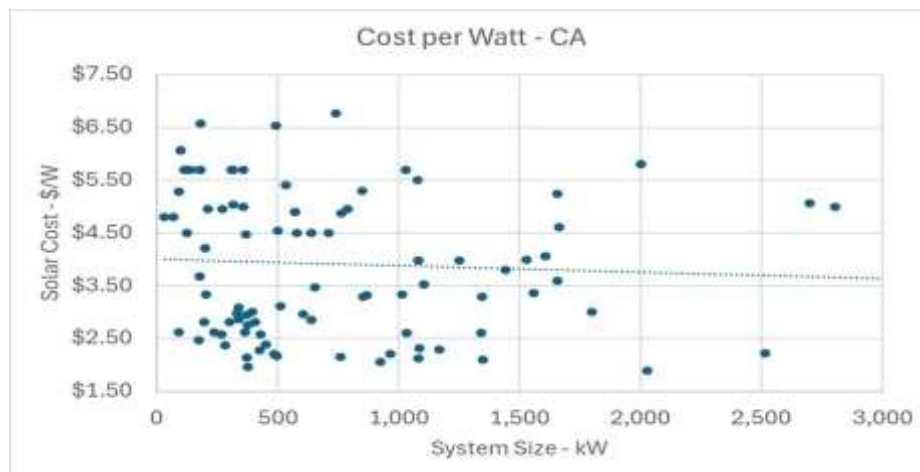
First, the relationship between reported \$/W and system size was directionally negative but statistically weak in the simple linear specification shown, as reflected by the low  $R^2$  value printed on the plot. The plotted trendline indicated that the average reported \$/W declined as system size increased, but the cloud of points around the trendline remained broad at nearly all sizes.

Second, the variance structure was non-uniform across the x-axis. The smallest systems contained the widest vertical dispersion, with reported \$/W values spanning from the low \$2/W range to above \$6/W. As systems increased in size into the mid-commercial band, the majority of points converged toward a narrower region centered in the low-to-mid \$2/W range, while a smaller but persistent group remained above \$3.5/W.

Third, the plot exhibited visible horizontal banding and state-driven stratification. A concentration of points in the low \$2/W range appeared heavily influenced by Illinois records, while the higher-cost bands and most of the extreme outliers appeared dominated by California records. The New Jersey records were few and did not materially influence the combined trendline, but they appeared in both mid-range and higher-cost regions.

#### 3.4.1. California results

California's scatterplot displayed the widest dispersion of reported \$/W in the study. The California dataset included observations spanning from below \$2/W to well above \$6/W across a broad system-size range. The plot showed a mild downward trendline, but dispersion remained high at small and mid-sized systems.



**Figure 2:** Cost per watt versus system size for California solar installations ( $n=96$ ). The scatter plot shows considerable price variation ranging from \$1.75/W to \$6.75/W, with a slightly declining trend line (dotted line) indicating weak economies of scale. Most installations cluster in the 200-600 kW range at \$2.25-\$3.00/W

Three outcomes were evident in the California figure. First, California contained a dense cluster of projects below 500 kW with reported \$/W concentrated around roughly \$2.2/W to \$3.2/W, alongside a substantial number of higher-cost observations in the \$4.0/W to \$6.0/W region. Second, California retained high-cost observations even at larger system sizes, including several projects above 1,000 kW with reported \$/W above \$3.5/W and some exceeding \$5/W. Third, the California plot contained the most prominent outlier tail, with multiple points approaching or exceeding \$6/W, which was not mirrored to the same degree in Illinois or New Jersey.

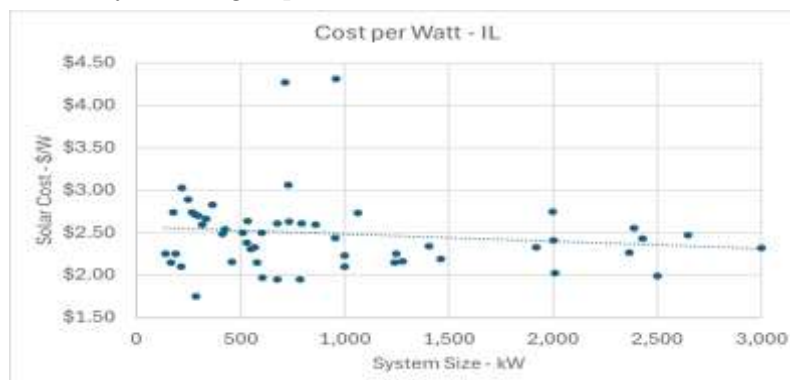
When visually segmented by system type from the record labels, canopy-labeled California observations appeared more frequently in the upper cost bands than flat-roof and ground-mounted observations. Similarly, certain grouped or

portfolio-related entries appeared as cost bands, with repeated \$/W values across multiple sizes, consistent with standardized pricing or allocation approaches within those groupings.

#### 3.4.2. Illinois results

Illinois displayed a comparatively tight distribution of reported \$/W across the plotted range. The trendline in Illinois sloped downward and the cloud of points clustered more narrowly than in California. Most observations appeared concentrated between roughly \$2.0/W and \$2.8/W, with a smaller number of observations above \$3.0/W and only a few reaching into the low \$4/W range.

**Results interpretation for Figure 3.** Three patterns emerge. First, the fitted line slopes downward, indicating that larger systems in Illinois tend to achieve lower



**Figure 3:** Installed cost per direct-current watt versus system size for Illinois projects in the study sample. Each marker is a distinct project; the dotted line is the ordinary least-squares fit. Costs are shown on a pre-ITC basis and include interconnection where applicable. The negative slope indicates modest economies of scale, while the dispersion at sub-MW scales points to site-specific soft costs, roof complexity, and interconnection scope.

delivered \$/W, but the gradient is shallow: scale helps, yet it does not erase cost variation. Second, variance is widest between roughly 200 and 1,000 kW, where outliers above \$3/W coexist with projects near \$2/W; this is consistent with heterogeneous roof conditions, prevailing-wage retrofits, change orders, and uneven interconnection requirements at these sizes. Third, beyond 1.5–2 MW, project costs cluster in a narrower band a little above \$2/W, suggesting partial convergence once projects clear minimum efficient scale and EPCs can standardize scope. These observations reinforce the paper's central result that CapEx is a first-order driver of financeability: economies of scale exist but are modest, so the most durable cost reductions come from disciplined engineering, soft-cost control, and interconnection management rather than from size alone.

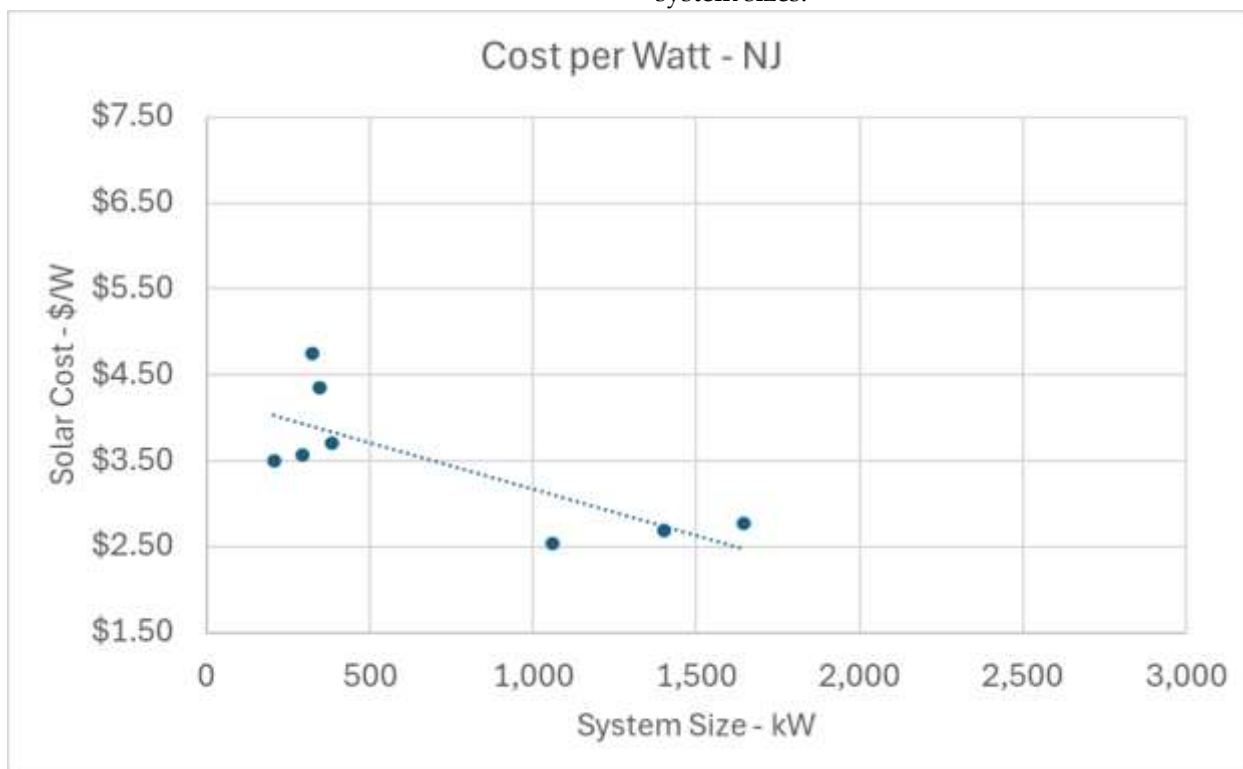
Three outcomes were evident in the Illinois figure. First, the Illinois dataset exhibited a strong central clustering in the low-to-mid \$2/W range across a wide system-size span, including systems well above 1,000 kW. Second, larger projects were

frequently observed at lower reported \$/W, including multiple large systems near or below approximately \$2.5/W, with fewer high-cost exceptions at the largest sizes than observed in California. Third, Illinois included a small set of high-cost observations relative to its main cluster, including one or more outliers above \$4/W, indicating the presence of scope-intensive or atypical records within a generally cost-consistent market segment.

The Illinois plot also exhibited banding behavior in portions of the dataset, with repeated or near-repeated \$/W values across multiple records. This pattern aligned with the presence of standardized project templates, multi-site portfolio entries, and recurring counterparties in the underlying records.

### 3.4.3. New Jersey results

New Jersey's scatterplot reflected a limited sample ( $n = 8$ ), and the trendline appeared steeper than those in California and Illinois. The plotted points spanned from the mid-\$2/W range to the mid-\$4/W range, with a downward trend across the observed system sizes.



**Figure 4:** Installed cost per direct-current watt versus system size for New Jersey projects in the study sample. Each marker is a distinct project; the dotted line is the ordinary least-squares fit. Costs are pre-ITC and include interconnection where applicable. The steeper negative slope relative to Illinois indicates stronger economies of scale in this subset, with sub-MW systems clustering between \$3.4–\$4.6/W and 1–1.7 MW systems near \$2.5–\$2.7/W.

**Results interpretation for Figure 4.** The trend shows a clear size–cost gradient: moving from a few hundred kW to 1–1.7 MW corresponds to a drop of roughly \$1/W in delivered EPC cost. Dispersion is highest below 500 kW, consistent with site complexity, roof work, and variable interconnection scope. Above 1 MW, variance tightens and costs converge toward the mid-\$2/W range, reinforcing that reaching minimum efficient scale is a practical lever for CapEx reduction in New Jersey.

Three outcomes were visible in the New Jersey figure. First, New Jersey exhibited the smallest sample and therefore the greatest sensitivity to individual records, with each observation meaningfully affecting the visual slope and implied state pattern. Second, the plotted trendline indicated lower reported \$/W at larger system sizes within this sample, with larger systems appearing in the \$2.5/W to \$3.0/W region while smaller systems included values above \$3.5/W. Third, at least one canopy-labeled record appeared in the higher-cost region, consistent with the broader observation that canopy-related entries tended to appear at higher \$/W than typical flat-roof entries in the dataset as a whole.

### 3.5. Configuration-linked patterns in reported cost per watt

Across the full dataset, system type labels aligned with distinct cost bands. Flat-roof and ground-mounted records frequently appeared within the \$2/W to \$3/W band in each state, while canopy-labeled records appeared more frequently in higher-cost regions. California exhibited the highest concentration of canopy projects and also contained the largest set of extreme \$/W observations. Illinois included fewer canopy-labeled observations and correspondingly fewer high-cost points, though at least one canopy record appeared at a high \$/W relative to the Illinois cluster. New Jersey included a canopy record at an elevated \$/W relative to other New Jersey records, consistent with the configuration-linked separation visible in the broader dataset.

The dataset also contained entries that were explicitly labeled as portfolio acquisitions or multi-site portfolios. These records commonly appeared as repeated \$/W values across multiple sites and sizes, consistent with pooled pricing, blended cost reporting, or standardized allocation across the portfolio. These records remained in the analysis to preserve the integrity of the dataset as encountered in public materials, and they contributed to the observed horizontal banding.

#### 3.5.1. Summary of principal empirical findings

Across 164 project records spanning three states, reported \$/W displayed a directionally negative relationship with system size in the pooled view, but size alone explained little of the variance in the combined dataset, consistent with broad dispersion and mixed configuration and reporting boundaries. California exhibited the highest mean reported \$/W and the widest dispersion, including the largest tail of high-cost observations. Illinois exhibited the lowest mean reported \$/W and the tightest clustering around low-to-mid \$2/W values. New Jersey exhibited intermediate mean reported \$/W with a small sample and a steep apparent size gradient. Across states, system configuration labels aligned with separable cost bands, with canopy-labeled records appearing more frequently in higher \$/W regions than flat-roof and ground-mounted records. The results section concluded with these observed patterns, which were subsequently interpreted in the Discussion section in relation to benchmarking reliability, comparability of public cost reporting, and implications for development and financing decisions.

## 4. DISCUSSION

### 4.1. Re-stating the research problem and what the results collectively indicate

This study set out to evaluate whether economies of scale and regional variation meaningfully explained differences in reported solar PV cost per watt across California, Illinois, and New Jersey when the underlying dataset reflected public-record project reporting rather than a standardized benchmark survey. The results answer the research question in a disciplined way: system size and geography both matter directionally, but the magnitude and reliability of those signals are highly dependent on configuration mix and reporting boundary consistency. The low explanatory power of the pooled size-to-cost fit is not a failure of the concept of scale economies. It is a direct empirical signature of how mixed-scope, mixed-configuration public cost data behave in the commercial segment.

In practice, the industry uses \$/W as a shorthand for build cost competitiveness, yet this dataset shows that the shorthand is only stable within narrower lanes than is commonly assumed. When records are pooled across states and system types, the \$/W metric becomes a composite of at least three separate processes: physical scale economies, configuration-driven engineering and civil scope differences, and reporting conventions that determine which cost categories are visible in public materials. The contribution of this paper is to separate those processes conceptually and show their fingerprints in the empirical patterns that practitioners routinely encounter but rarely

quantify.

#### 4.2. Economies of scale

Economies of scale appear in the dataset as a downward-sloping central tendency in every plotted view, including the pooled scatterplot and each state-specific plot. That said, the combined plot's low  $R^2$  signals that size alone is not a reliable predictor of reported \$/W in a mixed dataset. This is precisely the condition where decision-makers can make the wrong inference if they treat the average trendline as the story. What the plots actually show is a two-layer structure: a broad vertical spread that remains large at small and mid sizes, plus a shifting density of points toward lower \$/W as size increases.

A useful way to interpret the pooled pattern is to treat it as a market map rather than a regression result. At smaller sizes, the market contains both straightforward rooftop installs and scope-heavy projects that can embed roof work, structural upgrades, canopy civil work, procurement complexity, or other adders. Those scope-heavy projects create a tall variance band. As projects grow into the multi-hundred-kW to multi-MW range, a larger share of observations fall into a lower-cost region, which is consistent with fixed costs being spread across a larger watt base and with more repeatable execution. Yet the persistence of expensive projects at larger sizes, especially in California, demonstrates that scale is not sufficient. If configuration, site constraints, and reporting boundaries change with project size, the scale effect is diluted in observational data.

For veteran reviewers, the key point is not whether scale economies exist. It is whether the observed dataset can support stable rules of thumb. The answer here is conditional. Within Illinois' tighter band, the dataset suggests that larger projects more consistently achieve low-to-mid \$/W outcomes. In California, even large projects can occupy both low-to-mid and high \$/W regions, indicating that scale interacts strongly with configuration and scope. New Jersey's small sample hints at a strong scale gradient, but the evidence cannot carry the weight of generalization. In other words, the dataset supports a practical inference that scale helps, but it does not support a universal inference that size alone predicts cost in mixed public reporting.

#### 4.3. Regional variation: a mix of true structure and compositional effects

The state means and plot shapes demonstrate clear regional separation: Illinois clustered lower,

California spanned wider and higher, and New Jersey sat between with limited observations. A simplistic interpretation would claim that geography drives the difference. The data support a more careful interpretation. Illinois and California are not just different locations. They represent different mixes of project types, different procurement channels, and different frequencies of scope-intensive systems within the public record.

California's wide spread is informative in itself. The state contained both many low-to-mid \$/W projects and a heavy upper tail, including several observations above \$5/W. That pattern is consistent with a dataset that includes a higher share of configuration types that are structurally costlier on a \$/W basis, particularly canopy projects, and a higher share of projects where the public record may incorporate broader scopes. Illinois, in contrast, displayed a compact cluster centered in the low-to-mid \$2/W range, which is consistent with a dataset dominated by more repeatable commercial rooftops and ground-mounted builds with fewer scope-heavy records in the public sample.

New Jersey illustrates a third case. The mean is closer to California than Illinois, but the small sample and the steep trendline mean the visual story can be overfit by a few projects. The correct inference from New Jersey in this dataset is not a stable benchmark, but a demonstration of how easily stakeholders can form a strong narrative from a small set of public records. This is one of the most important meta-findings of the paper: when public cost reporting is sparse, the market narrative can be shaped by a handful of atypical projects.

#### 4.4. Configuration and scope signals: canopy effects as an organizing principle

Across the dataset, configuration labels separated into distinct cost bands. Canopy-labeled projects consistently appeared higher in the \$/W distribution than flat-roof and many ground-mounted projects, across states. This is not a subtle effect in the plots. It functions as an organizing principle for interpreting variance. The canopy-heavy segments in California, and the presence of canopy outliers in Illinois and New Jersey, help explain why cross-state comparisons based on state means alone can mislead.

This observation matters because canopy projects are often deployed in exactly the contexts that generate public records: school districts, municipalities, and visible institutional sites. These segments are also where policymakers and public finance stakeholders pay attention to project costs. If canopy-heavy procurements dominate the public

record, an external observer may conclude that a region's solar costs are inherently high, when the dataset may simply reflect a different build mix. The dataset therefore supports a specific recommendation for industry and policy analysis: any public-facing cost discussion should stratify by configuration before drawing state-level conclusions.

#### **4.5. Reporting boundary inconsistency**

The strongest methodological takeaway from the results is that reporting boundary inconsistency behaves like an omitted variable with high variance. It is visible through patterns that would not be expected in purely engineering-driven data, including pronounced horizontal banding, repeated \$/W values across multiple sites within portfolios, and the coexistence of low and high \$/W projects at similar sizes. These signatures suggest that a substantial fraction of the cost dispersion is not random measurement noise but systematic differences in what is being counted as cost.

This matters because the industry routinely compares \$/W across datasets that were never designed to be comparable. Benchmark work such as the NREL cost benchmarks is careful about definitions and category coverage, and the benchmarks are most useful precisely because they control scope (Ramasamy et al. 2023). Public records frequently do not. The implication is not that one is right and the other is wrong. The implication is that they answer different questions. Benchmarks answer what a defined system should cost under defined assumptions. Public records answer what a particular project or portfolio was reported to cost under a particular reporting convention, which may or may not include development margin, acquisition economics, roof work, or interconnection upgrades.

From a financing perspective, this distinction is decisive. Underwriting requires an all-in view, not an EPC-only view, and the contingency allocation required to protect DSCR depends on which cost categories are truly captured in the base case. A dataset like this can still be highly useful if it is treated as an empirical distribution of reported costs that market participants actually see, not as a direct proxy for all-in installed cost.

#### **4.6. Practical implications for developers, financiers, and policymakers**

For developers, the results argue for a disciplined internal benchmarking approach that mirrors the discipline of the best cost benchmarks, even when external datasets do not. Internal performance tracking should separate EPC scope, owner-side

costs, interconnection exposure, structural and roof work, and developer overhead. Without that separation, a developer can mistakenly attribute a high \$/W outcome to execution inefficiency rather than to scope, or can mistakenly claim competitiveness because the reported number excluded important categories. In competitive procurement environments, particularly those involving public entities, the ability to clearly state what is included in the reported project cost becomes a strategic advantage because it reduces bid ambiguity and reduces downstream dispute risk.

For financiers, the dataset indicates that relying on a single \$/W prior without configuration stratification and without scope adjustments is likely to understate tail risk. The plots show that even within a single state, projects at similar sizes can differ by multiples in reported \$/W. That dispersion translates into variability in construction period interest, contingency draw probabilities, and sponsor equity needs. In practice, the results support two underwriting disciplines. First, cost priors should be segmented by configuration and procurement channel. Second, when a public record reports a low \$/W, the diligence question should be whether scope was narrow rather than whether the project is unusually efficient.

For policymakers, the study reinforces that cost transparency is not an academic preference. It is a prerequisite for setting incentives and evaluating program effectiveness. If public-facing cost data mix canopy-heavy, roof-intensive, and ground-mounted projects without labels, incentive levels calibrated to those numbers will be misaligned. The policy risk is symmetric. Incentives can be set too low if public records understate all-in cost, or they can be set too high if the public record is dominated by scope-heavy projects not representative of the median build. The practical recommendation is straightforward: public programs that collect cost data should require minimum scope disclosure fields, including whether the reported cost includes interconnection upgrades, roof replacement, structural reinforcement, and developer fee. This is the simplest step toward making cost data decision-useful.

#### **4.7. Reconciling this dataset with benchmark literature using limited citations**

The most credible way to triangulate these findings is to compare them conceptually to the structure of benchmark work without forcing numeric equivalence. NREL's cost benchmarks provide a disciplined decomposition framework and

emphasize why scope clarity matters when interpreting system prices (Ramasamy et al. 2023). Berkeley Lab's empirical work on PV pricing trends shows that installed price dispersion persists even in large datasets and that local market conditions and business models affect pricing outcomes (Barbose et al. 2024). The present dataset aligns with those themes but adds a specific commercial-market observation: when public records and transaction summaries are used as data inputs, reporting conventions can behave like a dominant variance driver, enough to suppress simple size-to-cost explanatory power in pooled regressions.

The interconnection environment provides a complementary explanation for why public costs can deviate sharply from benchmark baselines even for seemingly comparable systems. Federal interconnection reforms acknowledge that queue processes and upgrade exposure are material barriers that can affect project economics and timelines (Federal Energy Regulatory Commission 2023). When upgrade work or mitigation is embedded into reported project cost for some projects and excluded for others, the \$/W metric becomes a mixed object. The dataset's dispersion, particularly within California, is consistent with a world in which some reported costs incorporate broader site and grid-facing scope than others.

#### ***4.8. Limitations and directions for future research that follow directly from the findings***

Several limitations follow from the dataset design and also define a clear path for extension. The most important limitation is that the analysis could not fully normalize cost scope across records because public sources rarely provide line-item cost detail. The study therefore treated scope variability as part of the empirical phenomenon. A direct extension would be to create a scope taxonomy and apply it systematically through additional document review, including procurement award packets, EPC scopes, and interconnection documentation, then re-estimate size and state effects within scope-harmonized subsets.

A second limitation is configuration mix, especially canopy representation. The dataset strongly suggests that configuration stratification is necessary for stable bench-marking. Future work should formalize that stratification by estimating separate distributions for roof, ground, and canopy, then quantifying how much of each state mean is explained by composition versus within-configuration differences.

A third limitation is the small New Jersey

sample. The New Jersey figure should be treated as illustrative rather than definitive. Future research should expand New Jersey coverage through additional public procurement sources and market datasets to test whether the steep trend observed here persists.

A fourth limitation is time variation. Records spanned multiple years, a period in which supply chain pricing, rates, and labor conditions changed. A future extension could index costs to a common quarter using a transparent inflation and commodity adjustment method, while clearly stating uncertainty bands.

Finally, future work can translate the present findings into a more finance-native metric set. Because \$/W is a build-cost proxy, the next step is to connect the observed dispersion to PPA pricing, LCOE, and DSCR outcomes by mapping cost bands to financing structures and expected energy yields. That extension would make the findings even more directly usable for project finance underwriting and policy design.

#### ***4.9. Closing synthesis***

Taken as a whole, the results show that \$/W remains a useful metric, but only when used with the segmentation discipline that practitioners often apply implicitly and that public discourse often ignores. The dataset demonstrated that pooled size-to-cost relationships weakened sharply in the presence of mixed configurations and mixed reporting boundaries, even while a directional scale effect remained visible. California's higher mean and wider dispersion reflected a dataset rich in scope-intensive and canopy-heavy records, Illinois' lower mean reflected a tighter and more repeatable cost band, and New Jersey's small sample illustrated how quickly narratives can form from limited public data. The most important contribution of this paper is therefore not a single benchmark number, but a measurement insight: public-record solar cost data are informative precisely because they reveal how reporting practices, configuration mix, and project heterogeneity jointly shape what the industry and policymakers perceive as solar cost.

### **5. CONCLUSION**

This study examined how reported cost per watt varies across commercial-scale solar PV projects in California, Illinois, and New Jersey using a 164-record dataset assembled from public-facing project information. The central question was intentionally practical: when market participants rely on public cost signals, do economies of scale and regional factors reliably explain differences in

cost per watt, or are those signals substantially shaped by configuration mix and inconsistent reporting boundaries. The results answered that question in a way that is directly relevant to developers, financiers, and policymakers who must make decisions under imperfect information. System size and geography showed directionally meaningful relationships with cost per watt, but neither variable, on its own, provided a stable explanatory framework for the dispersion observed in public records. The strongest message of this paper is therefore not a single benchmark number. It is an evidence-based warning and a constructive pathway forward: cost per watt remains useful, but only when used with scope discipline and configuration stratification that reflect how projects are actually delivered and reported.

Several empirical findings stand out. First, the pooled relationship between system size and reported cost per watt was negative but weak, with size explaining only a small fraction of the variance in the combined dataset. This means the dataset did not support a simple rule such as larger projects are cheaper in a way that is sufficiently reliable for benchmarking across mixed project types. Instead, scale effects appeared as a shift in density toward lower cost per watt at larger sizes, while wide dispersion persisted across the full size range. Second, state-level patterns were distinct. California exhibited the highest average reported cost per watt and the broadest dispersion, including a pronounced high-cost tail. Illinois exhibited the lowest average and a comparatively tight clustering around low-to-mid cost per watt values. New Jersey appeared intermediate but could not be generalized due to the small sample size. Third, configuration labels separated into cost bands, with canopy projects appearing consistently higher in cost per watt than many flat-roof and ground-mounted systems. This configuration mix effect provided a coherent explanation for why California's distribution included a larger high-cost tail and why cross-state comparisons based purely on state means can mislead.

These findings matter because cost per watt is used as a shared language across stakeholders who do not share the same level of technical familiarity with what the metric includes. When a public record states a project cost, the value may reflect different accounting boundaries across projects. Some reported costs may resemble EPC-only scope. Others may embed structural upgrades, roof work, electrical upgrades, interconnection mitigation, or portfolio-level pricing allocations.

Some may implicitly include margin or development overhead while others omit it. This dataset displayed signatures consistent with such boundary variation, including visible banding in certain subsets and the coexistence of high and low cost per watt outcomes at similar system sizes. For industry readers, this reinforces a core truth that is sometimes overlooked in public debate: what appears to be a geographic premium or an outlier price can be a scope definition artifact. The discipline required is not simply better regression. It is better disclosure and better segmentation.

The implications are immediate. For developers, these results support the case for internal benchmarking frameworks that mirror best-practice cost decomposition rather than relying on external public records as direct comparables. Developers that separate EPC scope, owner-side costs, interconnection exposure, structural work, and sponsor overhead can benchmark execution more accurately and defend their pricing more credibly in competitive procurements. For financiers, the observed dispersion highlights why underwriters should treat a single market cost prior as inadequate unless it is stratified by configuration and reporting scope. In a dataset where dispersion persists even among similarly sized systems, contingency planning, reserve sizing, and construction risk assessment must be grounded in distributions rather than point estimates. A low reported cost per watt should trigger diligence about scope completeness, not immediate confidence in superior execution. For policymakers, the dataset reinforces that cost transparency is not merely an administrative preference. It is foundational for designing incentives and evaluating programs. If public-facing cost data are dominated by scope-intensive configurations or inconsistent reporting, then program cost-effectiveness assessments and incentive calibration can drift away from underlying economic reality.

This paper also clarifies what can and cannot be concluded from the study design. The dataset was drawn from public records and therefore was not a random sample of all projects in each state. It likely overrepresents projects that are publicly disclosed because they involve public entities, prominent counterparties, visible installations, or portfolio-level transactions. New Jersey's small sample limited statistical claims for that state and positioned the New Jersey results as illustrative rather than definitive. The analysis was also constrained by incomplete scope transparency in public sources, preventing full normalization of cost

categories across records. The paper treated that limitation as part of the empirical phenomenon rather than a removable nuisance, because inconsistent reporting boundaries are central to how public cost perceptions are formed. Finally, the dataset spanned multiple years, which likely introduced time-based variability tied to supply chain conditions, labor availability, and financing costs that could not be fully controlled without additional structured metadata.

Even with these limitations, the study contributes a useful and publishable insight: when cost per watt is extracted from public deal information across mixed configurations and mixed reporting conventions, the metric behaves less like a clean engineering benchmark and more like a composite indicator of delivery scope, procurement context, and reporting practice. The low pooled explanatory power of system size is not evidence against scale economies. It is evidence that scale economies are often masked in public reporting by configuration and scope variability. This is precisely the sort of insight that experienced practitioners recognize intuitively but benefit from seeing quantified and organized. It also explains why benchmarking conversations can break down when stakeholders assume comparability that does not exist.

The findings point to clear next steps for improving both research and practice. A natural extension is to develop a structured reporting taxonomy that classifies each record by likely cost boundary, such as EPC-only, turnkey installed cost with explicit development components, and portfolio acquisition pricing with blended allocation. With such a taxonomy, future analysis could re-estimate scale and state effects within

boundary-consistent subsets and quantify how much of each state's apparent premium is compositional versus structural. A second extension is configuration-stratified benchmarking, separating flat-roof, ground mount, and canopy into distinct distributions and then testing whether economies of scale become more legible within each configuration. A third extension is deeper linkage between reported cost per watt and economic outcomes such as PPA pricing, LCOE, and underwriting metrics, which would convert cost dispersion into decision-relevant financial implications. Lastly, expanding the New Jersey sample and increasing the representation of comparable project classes across all three states would strengthen external validity and allow more confident inference about regional dynamics.

At the highest level, this research matters because cost narratives shape deployment. When cost perceptions are distorted by inconsistent reporting, stakeholders can make decisions that slow down the energy transition, misallocate incentives, or mis-price risk. This study shows that the industry can still extract useful signal from public cost data, but only by treating disclosure and segmentation as essential analytic steps rather than optional refinements. If the sector can converge on a small set of minimum disclosure fields for public-facing cost reporting, and if analysts consistently stratify by configuration and scope boundary, cost per watt can regain its role as a reliable shared metric. The dataset examined here demonstrates both the risks of assuming comparability and the opportunity that emerges when the comparability problem is addressed directly.

## REFERENCES

- Barbose G, Darghouth N, OShaughnessy E, Forrester S (2024) Tracking the Sun, 2024 Edition: Pricing and design trends for distributed photovoltaic systems in the United States. Lawrence Berkeley National Laboratory. <https://emp.lbl.gov/sites/default/files/2024-10/Tracking%20the%20Sun%202024%20Report.pdf>
- Barbose G, Darghouth N, OShaughnessy E, Forrester S (2024) Tracking the Sun, 2024 Edition: Executive Summary. Lawrence Berkeley National Laboratory. <https://emp.lbl.gov/sites/default/files/2024-08/Tracking%20the%20Sun%202024%20Executive%20Summary.pdf>
- Brattle Group, Grid Strategies (2024) Generator interconnection scorecard. Advanced Energy United. <https://www.brattle.com/wp-content/uploads/2024/03/Generator-Interconnection-Scorecard.pdf>
- Department of Energy (2024) i2X Trans-mission interconnection roadmap. U.S. Department of Energy. <https://www.energy.gov/sites/default/files/2024-04/i2X%20Transmission%20Interconnection%20Roadmap.pdf>
- Department of Energy (2024) Transmission impact assessment. U.S. Department of Energy. [https://www.energy.gov/sites/default/files/2024-12/DOE Transmission Impact Assessment Report 120 524.pdf](https://www.energy.gov/sites/default/files/2024-12/DOE%20Transmission%20Impact%20Assessment%20Report%20120524.pdf)
- Department of Labor (2024) Inflation Reduction Act prevailing wage and apprenticeship requirements. U.S. Department of Labor. <https://www.dol.gov/agencies/whd/government-contracts/inflation->

- reduction-act
- Farmer JD, Lafond F (2016) How predictable is technological progress? *Res Policy* 45(3):647–665. <https://doi.org/10.1016/j.respol.2015.11.001>
- Federal Energy Regulatory Commission (2023) Order No. 2023: Improvements to generator interconnection procedures and agreements. <https://www.ferc.gov/media/order-no-2023>
- Federal Energy Regulatory Commission (2023) Generator interconnection rule explained. <https://www.ferc.gov/media/generator-interconnection-rule-explained>
- Federal Energy Regulatory Commission (2024) FERC Order No. 2023-A. <https://www.ferc.gov/media/ferc-order-no-2023-a>
- Federal Energy Regulatory Commission (2024) Order 2023 rehearing and clarification explained. <https://www.ferc.gov/media/order-2023-rehearing-and-clarification-explained>
- Federal Energy Regulatory Commission (2024) Order No. 2023-A final rule, public inspection PDF. Federal Register. <https://public-inspection.federalregister.gov/2024-06563.pdf>
- Federal Energy Regulatory Commission (2024) FERC Order No. 1920: Building for the future through electric regional transmission planning and cost allocation. <https://www.ferc.gov/media/ferc-order-no-1920>
- Federal Energy Regulatory Commission (2024) Transmission planning rule explained. <https://www.ferc.gov/media/transmission-planning-rule-explained>
- International Energy Agency (2024) Renewables 2024: Analysis and forecast to 2030. IEA. <https://www.iea.org/reports/renewables-2024>
- International Energy Agency (2024) Renewables 2024: Analysis and forecast to 2030 (PDF). IEA. <https://iea.blob.core.windows.net/assets/493a4f1b-4031-49c9-a7c1-995ef7770d45/Renewables2024.pdf>
- International Energy Agency (2024) World Energy Outlook 2024. IEA. <https://www.iea.org/reports/world-energy-outlook-2024>
- International Renewable Energy Agency (2024) Renewable power generation costs in 2023. IRENA. <https://www.irena.org/Publications/2024/Jul/Renewable-power-generation-costs-in-2023>
- International Renewable Energy Agency (2024) Renewable power generation costs in 2023 (PDF). IRENA. [https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2024/Jul/IRENA\\_Renewable\\_Power\\_Generation\\_Costs\\_in\\_2023.pdf](https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2024/Jul/IRENA_Renewable_Power_Generation_Costs_in_2023.pdf)
- Internal Revenue Service (2024) Notices and guidance under the Inflation Reduction Act. IRS. <https://www.irs.gov/credits-deductions/notices-and-guidance-under-the-inflation-reduction-act>
- Internal Revenue Service (2024) Prevailing wage and apprenticeship requirements. IRS. <https://www.irs.gov/credits-deductions/prevailing-wage-and-apprenticeship-requirements>
- Internal Revenue Service (2022) Notice 2022-61: Additional guidance on prevailing wage and apprenticeship requirements. IRS. <https://www.irs.gov/pub/irs-dro/p/n-22-61.pdf>
- Kavlak G, McNerney J, Trancik JE (2018) Evaluating the causes of cost reduction in photovoltaic modules. *Energy Policy* 123:700–710. <https://doi.org/10.1016/j.enpol.2018.09.015>
- Lazard (2025) Levelized cost of energy, levelized cost of storage, and levelized cost of hydrogen. Lazard. <https://www.lazard.com/research-insights/levelized-cost-of-energy-levelized-cost-of-storage-and-levelized-cost-of-hydrogen/> Lazard (2025) Lazard's LCOE+ (June 2025), Version 18.0 (PDF). Lazard. <https://www.lazard.com/media/j4oo5xhg/lazards-lcoeplus-june-2025-vf.pdf>
- Masson G, Kaizuka I, Smets A et al (2024) Trends in photovoltaic applications 2024. IEA PVPS. [https://iea-pvps.org/wp-content/uploads/2024/04/IEA\\_PVPS\\_Trends\\_Report\\_2024.pdf](https://iea-pvps.org/wp-content/uploads/2024/04/IEA_PVPS_Trends_Report_2024.pdf) doi:10.69766/JNEW6916
- National Renewable Energy Laboratory (2024) Annual Technology Baseline 2024: Utility-scale PV. NREL. <https://atb.nrel.gov/electricity/2024/2023/utility-scale-pv>
- National Renewable Energy Laboratory (2024) Fall 2024 solar industry update (PDF). NREL. <https://www.nrel.gov/docs/fy25osti/92257.pdf>
- Raj K, Roy A, Sapkota J (2026) Comparative analysis of UK vs US developer for pricing solar, lighting, and BESS PPAs. *Future Technology*
- Rand J, Manderlink N, Bolinger M, Wiser R, Jeong S (2024) Queued Up: 2024 Edition, characteristics of power plants seeking transmission interconnection as of the end of 2023. Lawrence Berkeley National Laboratory. <https://emp.lbl.gov/queues>
- Lawrence Berkeley National Laboratory (2025) U.S. interconnection queue data through 2024. <https://emp.lbl.gov/publications/us-interconnection-queue-data>

- Raj K, Walton M (2025) Which assumptions really set power purchase prices and returns in United States solar projects. *Advances in Consumer Research* 2(5):1960–1978. <https://acr-journal.com/>
- Raj K (2025) Impacts of Investment Tax Credit on public schools in the United States: A paired analysis of four Connecticut projects. *International Journal of Environmental Sciences* 11(24s):3320–3331. <https://doi.org/10.64252/tz4ra011>
- Ramasamy V, Feldman D, Fu R, Woodhouse M (2023) U.S. solar photovoltaic system and energy storage cost benchmarks, Q1 2023 (NREL/TP-7A40-87303). National Renewable Energy Laboratory. <https://www.nrel.gov/docs/fy23osti/87303.pdf> doi:10.2172/2005540
- Ramasamy V, Woodhouse M, Feldman D (2025) Documenting 15 years of solar PV cost reductions: Historical and near-term outlook (PDF). National Renewable Energy Laboratory. <https://www.nrel.gov/docs/fy25osti/92536.pdf>
- Solar Energy Industries Association, Wood Mackenzie (2025) U.S. Solar Market Insight: 2024 year in review, executive summary (PDF). SEIA. <https://seia.org/wp-content/uploads/2025/03/SMI-2024-YIR-ES.pdf>
- Seel J, Mulvaney Kemp J, Cheyette A et al (2024) Utility-scale solar, 2024 edition: Executive summary. Lawrence Berkeley National Laboratory. [https://emp.lbl.gov/sites/default/files/2024-10/Utility-Scale%20Solar%202024 Executive Summary.pdf](https://emp.lbl.gov/sites/default/files/2024-10/Utility-Scale%20Solar%202024%20Executive%20Summary.pdf)
- Wright TP (1936) Factors affecting the cost of airplanes. *J Aeronaut Sci* 3(4):122–128