Marginal Value of Energy in an Automated Economy

Matthew Long
The YonedaAI Collaboration
YonedaAI Research Collective
Chicago, IL
matthew@yonedaai.com · https://yonedaai.com

23 July 2026

Abstract

The next unit of electricity is tied to a location and an interval. Its opportunity cost can include energy, congestion, transmission losses, operating reserve, reliability margins, ramp limits, workload deadlines, and marginal emissions. Average electricity use, an average tariff, and an average grid emissions rate do not identify that opportunity cost. This paper gives a time-indexed and node-indexed allocation model for automated loads. Workload benefits are concave and risk adjusted. Electricity prices are fixed from the allocator's perspective. Network, reserve, ramp, deadline, and service constraints are convex. Under these assumptions, the Karush, Kuhn, and Tucker conditions are necessary and sufficient. At an interior optimum, the risk-adjusted marginal value of energy assigned to a workload equals the nodal opportunity cost plus the shadow cost of every binding system and intertemporal constraint. Workloads with the same constraint footprint therefore equalize their marginal net values. The theorem is PROVED only for the stated convex price-taking model.

A dependency-free Python implementation solves small strictly concave quadratic instances by active-set enumeration. Tests compare shadow prices with finite differences and check primal feasibility, dual feasibility, stationarity, and complementary slackness. Illustrative scenarios cover AI inference, batch computation, manufacturing, transport charging, and proof-of-work settlement. They show how rankings can change with location, deadline, reliability, and marginal emissions. The sector response curves are not empirical estimates, so the scenarios do not support a universal sector ranking or a policy prescription.

Keywords: marginal value; electricity allocation; locational marginal price; congestion; flexible load; shadow price; marginal emissions; automated economy

Evidence status. The convex allocation theorem and its corollaries are PROVED under the assumptions in Section 5. The numerical examples and software checks are COMPUTATIONAL. Claims about unobserved sector response curves are OBSTRUCTED. Estimation of those curves is OPEN.

Contents

1 The decision is the next unit, not the average

An automated economy does not allocate an annual national energy total in one stroke. A scheduler places one batch in a data center for a particular hour. A factory advances or delays a process subject to a production window. A fleet controller decides which vehicles charge before morning. A proof-of-work operator runs or curtails machines as electricity prices and settlement revenue change. Each decision concerns an incremental quantity at a node and time.

This observation rules out several tempting calculations. Dividing annual sector output by annual energy use gives an average ratio. It says nothing by itself about the value produced by the next megawatt-hour. A retail tariff may recover fixed network costs through volumetric charges. It need not equal the short-run opportunity cost of one more unit of load. An annual grid emissions factor averages all generation. It does not show which generators respond to an incremental load in a given interval. The differences are conceptual, not rounding errors.

The Federal Energy Regulatory Commission describes a locational marginal price as the marginal cost of serving load at a specific location, given the resources dispatched and transmission limitations. In organized United States markets, the reported price has energy, congestion, and transmission-loss components [5]. That is a useful starting point, but it is not the whole allocation problem. A flexible load may also face a reserve charge, a contractual demand charge, service loss from delay, reliability risk, a carbon price, or an internal deadline. Some of these are market prices. Others are shadow prices generated by the workload's own constraints.

The phrase "marginal value of energy" will mean the derivative of a declared workload benefit with respect to an incremental megawatt-hour, holding fixed a counterfactual and a decision interval. The paper uses megawatt-hours in the equations because system operators and large loads schedule at that scale. Conversion to joules is exact:

$$1 \text{ MWh} = 3.6 \times 10^9 \text{ J}.$$

Nothing in the theory depends on the display unit.

1.1 Research question

The research question is narrow:

> Under what assumptions should the next feasible unit of electricity be assigned to a particular automated workload, node, and interval?

The strongest defensible answer is conditional. If workload benefit functions are concave, costs are evaluated at fixed price-taking terms, uncertainty enters through a declared risk adjustment, and all operational constraints are convex, then an optimum has a marginal equalization rule. The rule is not a claim that markets always satisfy those conditions. It is a benchmark against which nonconvexities, strategic bidding, missing prices, and measurement gaps can be named.

1.2 Contributions

The paper contributes the following:

1. a node-indexed and time-indexed allocation model that separates market energy prices from congestion, losses, reserves, reliability, workload service loss, and marginal emissions;
2. a KKT equalization theorem under explicit convexity, price-taking, information, and feasibility assumptions;
3. a clean distinction among average energy intensity, private marginal cost, nodal system opportunity cost, and social marginal cost;
4. linear representations of capacity, interface, reserve, ramp, deadline, and reliability restrictions;
5. a small standard-library solver with shadow-price, finite-difference, feasibility, and complementary-slackness checks;
6. scenario results whose synthetic status is stated in every interpretation.

1.3 Nonclaims

This paper does not estimate national demand response for AI, manufacturing, transport, or cryptographic settlement. It does not prove that flexible computing benefits a grid. It does not treat a distribution-network tariff as a wholesale nodal price. It does not model voltage, reactive power, transient stability, or unit commitment. It does not rank sectors without matched response data. Those limitations are part of the result.

2 Prior foundations and the scope of the model

2.1 Marginal cost, welfare, and peak load

The welfare logic is older than modern electricity markets. Marginal-cost pricing links an incremental decision to the opportunity cost of the resources it consumes. Samuelson's treatment of constrained optimization and welfare provides the general economic frame [1]. Boiteux's peak-load pricing work showed why electricity costs depend on when capacity is required, not merely on total annual energy [2]. Schweppe, Caramanis, Tabors, and Bohn developed a location-sensitive and time-sensitive account of electricity spot pricing [3]. Stoft connects the same pricing principles to practical electricity-market design and scarcity [17].

These references do not imply that every observed price equals social marginal cost. Retail rates can average costs across time and geography. Wholesale markets contain uplift, scarcity rules, capacity mechanisms, and administrative interventions. External damages may be unpriced. Consumers may not see real-time prices. Borenstein and Holland show that time-invariant retail prices can break the welfare correspondence that competitive wholesale pricing might otherwise provide [4]. The present model therefore keeps the price vector, environmental term, and workload benefit functions separate. Borenstein's review of renewable electricity also distinguishes private incentives from public benefits and costs [18].

2.2 Locational marginal pricing and congestion

FERC's 2024 *Energy Markets Primer* states that organized market software matches supply and demand while respecting physical resource and transmission limits. When a transmission constraint binds, the least-cost generator for the whole region may be unable to serve an additional load at a constrained node. A more expensive local or redispatched unit then affects the nodal price [5]. The LMP decomposition is commonly written

$$\mathrm{LMP}_{n,t} = p_t^E + p_{n,t}^C + p_{n,t}^L,$$

where p_t^E is the reference energy component, $p_{n,t}^C$ is congestion, and $p_{n,t}^L$ is the marginal-loss component.

The formula has a precise domain. It describes a market result produced by a particular network model and dispatch. It does not show that an end user paid the LMP. It also does not make a workload's internal deadline disappear. Our allocation model treats an observed or scenario LMP as exogenous. It then adds internal and external terms without claiming to reproduce the system operator's dispatch engine.

2.3 Reserve, ramping, and reliability

Energy balance alone does not keep a power system reliable. FERC distinguishes energy from ancillary services and notes that regulation follows short-term load changes, while operating reserves must be available within specified times [5]. NERC defines and maintains reliability terminology and standards used by balancing authorities [6]. A load allocator must not consume capacity that has already been withheld for reserve or a declared reliability margin.

Ramping matters on both sides of the meter. A generator cannot always change output instantly. An automated workload may also have a maximum change between intervals because of thermal cycling, queue stability, material continuity, charging limits, or restart overhead. The mathematical form can be similar even though the engineering causes differ.

Reliability enters the model in two places. First, system operators may withhold deliverable capacity. Second, a workload may value scheduled energy less when service delivery can fail and may face an expected failure loss. A full reliability model would treat correlated outages and state-dependent recourse. The reference implementation uses an explicit reliability factor and loss coefficient. It is an audit-friendly approximation.

2.4 Marginal emissions

An average emissions rate divides total emissions by total generation. A marginal emissions factor asks how emissions change under an incremental load or resource change. Siler-Evans, Azevedo, and Morgan estimated regional and hourly marginal factors for the United States and showed substantial variation [7]. Holland and coauthors demonstrate that the environmental effect of electric driving can vary by location because marginal generation and damages vary [8]. Callaway, Fowlie, and McCormick make the same general point for renewable energy and demand-side resources: location and timing change marginal value [9]. Recent work by Holland and coauthors treats marginal-emissions estimation as a high-dimensional identification problem and uses economic restrictions to regularize the estimate [19].

EPA's AVERT tool estimates emissions impacts from load or clean-energy changes using historical hourly generation and emissions patterns. The current EPA description says that AVERT's statistical module uses observed generation and emissions data, while its main module applies regional scenario changes [10]. EPA also publishes approximate marginal emission rates for AVERT regions [12]. These are more relevant than annual average mix factors for many incremental studies, but they remain model-based regional estimates. They are not exact node-level causal derivatives.

2.5 Convex optimization

The necessary conditions trace to Karush and to Kuhn and Tucker [15, 16]. The mathematical result uses standard convex duality. Boyd and Vandenberghe show that KKT conditions are sufficient for convex programs and are necessary under suitable constraint qualifications [13]. Rockafellar provides the broader convex-analysis foundation [14]. The economic interpretation of a multiplier as a local value of relaxing a constraint follows from the value function and envelope theorem.

This paper does not contribute a new KKT theorem. Its contribution is the disciplined mapping of automated workload value, nodal electricity opportunity cost, emissions, service constraints, and energy boundaries into one model, together with a reference implementation that checks the claimed conditions.

3 Measurement contract

Every application begins with a measurement tuple

$$\mathcal{M} = (b, c, W, \tau, e, a, \pi),$$

following the shared framework of the Joule Standard research program. Here b is the physical boundary, c the counterfactual, W the value or welfare functional, τ the horizon, e the energy convention, a the attribution rule, and π the uncertainty model.

For marginal allocation, the tuple needs a few operational fields as well. Table 1 gives the minimum manifest.

Table 1: Required manifest for a marginal allocation result

Field	Required declaration
Functional unit	Incremental MWh, kWh, or joules delivered to a named load within a stated interval.
Node and interval	Electrical location, market zone or node, clock, interval length, and time zone.
Physical boundary	Meter, facility, network service, and any upstream losses included in the decision.
Counterfactual	What would run, wait, fail, or consume electricity if the incremental allocation were withheld.

Field	Required declaration
Value functional	Producer surplus, expected loss avoided, task service, consumer surplus, or another named outcome and currency base.
Price vector	Energy, congestion, losses, reserve, capacity, and other charges, with a statement of which are marginal.
Flexibility set	Bounds, location choices, deadlines, ramp limits, restart rules, and indivisibilities.
Reliability	Availability model, reserve deduction, service probability, failure loss, and correlation assumptions.
Emissions	Marginal factor, pollutant, geographic scope, counterfactual, valuation or cap, and uncertainty.
Attribution	Rule connecting allocated energy to the incremental workload outcome.
Uncertainty	Scenario distribution, robust set, confidence interval, or explicit deterministic assumption.
Exclusions	Dynamic stability, market power, embodied energy, fixed costs, or other omitted terms.

3.1 Energy convention

The denominator in this paper is marginal grid energy delivered to an incremental workload at node n and interval t. It is not annual primary energy, device telemetry, facility electricity, or lifecycle energy. If a facility-level analysis adds cooling or conversion overhead, the analyst must state whether the nodal price applies before or after that overhead. If embodied energy changes at the margin because a capacity investment is triggered, the short-run model is no longer sufficient.

3.2 Private value and social value

Let V_i be a declared benefit function for workload i. It may describe private surplus or a social welfare component. Those are not interchangeable. A private allocator might maximize expected profit after energy charges. A social planner might add consumer surplus, local pollution damages, reliability externalities, and distributional weights. The equation can accommodate either, but the label must say which.

This paper does not infer welfare from revenue. Revenue is a transfer between parties unless it proxies a separately defined surplus. Likewise, a token, hash, manufactured unit, or vehicle-mile is an output quantity, not a common value unit. Cross-sector allocation requires a declared value functional or a vector report.

3.3 Average and marginal quantities

For a workload with total outcome $W(E)$, average value per unit is

$$\mathrm{AV}(E) = \frac{W(E)}{E},$$

while marginal value is

$$\mathrm{MV}(E) = W'(E).$$

They coincide only for a linear function through the origin. With diminishing returns, average value can remain high after marginal value has fallen below the opportunity cost. Allocating the next unit by average ratios can therefore send energy to the wrong workload even inside a perfectly measured model.

The same distinction applies to cost and emissions:

$$\mathrm{AC}(E) = \frac{C(E)}{E}, \qquad \mathrm{MC}(E) = C'(E),$$

$$\mathrm{AEF}(E) = \frac{G(E)}{E}, \qquad \mathrm{MEF}(E) = G'(E).$$

The derivatives are local to a counterfactual operating point. They may change when a constraint binds, a generator starts, or an investment is triggered.

4 Node-indexed and time-indexed allocation

4.1 Indices and decisions

Let $i \in \mathcal{I}$ index workloads, $n \in \mathcal{N}$ nodes, and $t \in \mathcal{T}$ intervals. The decision

$$x_{i,n,t} \geq 0$$

is incremental MWh assigned to workload i at node n in interval t. Define aggregate workload energy

$$q_{i,t} = \sum_{n \in \mathcal{N}} x_{i,n,t}.$$

Variables can represent alternatives. For example, one batch may run at either of two data centers, subject to a shared completion requirement.

Let ω denote uncertain workload and system conditions. The declared risk-adjusted workload benefit is

$$B_i(x_i) = \mathbb{E}_\pi[V_i(x_i, \omega)] - \rho_i R_i(x_i),$$

where R_i is a risk functional and $\rho_i \geq 0$ is a declared risk weight. In the reference code,

$$B_{i,n,t}(x) = a_{i,n,t}\left(v_{i,n,t}x - \frac{\kappa_{i,n,t}}{2}x^2\right) - (1 - a_{i,n,t}) f_{i,n,t} x.$$

Here a is service reliability, v the intercept of marginal gross value, $\kappa > 0$ curvature, and f failure loss per MWh. For the separable quadratic implementation, the workload-level object in the main objective is

$$B_i(x_i) = \sum_{n \in \mathcal{N}} \sum_{t \in \mathcal{T}} B_{i,n,t}(x_{i,n,t}).$$

4.2 Nodal opportunity cost

The declared external cost per MWh is

$$
\begin{aligned}
c_{n,t} = {} & p^E_{n,t} + p^C_{n,t} + p^L_{n,t} + p^R_{n,t} \\
& + \chi_t g_{n,t} + d_{n,t}.
\end{aligned} \tag{1}
$$

The components are:

- p^E: marginal energy charge;
- p^C: congestion component;
- p^L: marginal transmission-loss component;
- p^R: reserve or reliability service adder allocated to the load;
- g: marginal emissions in kg per MWh;
- χ: declared damage, tax, or shadow value per kg;
- d: another declared marginal damage or network term.

When an observed LMP already contains energy, congestion, and losses, those terms must not be added twice. A capacity payment or fixed network charge belongs in Equation (1) only if the incremental decision changes it. Average bill allocation is not enough. Energy, congestion, and loss price components may be negative. The reference code permits signed market components and keeps any sign restriction on an environmental damage value explicit.

4.3 Capacity, reserve, and reliability margin

Let $\bar{C}_{n,t}$ be incremental deliverable capacity before load-side deductions. Let $r_{n,t}$ be capacity withheld for operating reserve and $m_{n,t}$ a declared reliability margin. Then

$$
\sum_{i \in \mathcal{I}} x_{i,n,t} \leq \bar{C}_{n,t} - r_{n,t} - m_{n,t}.
$$

This is a load-allocation constraint, not a statement about the full bulk system reserve requirement. The terms must be based on the same node, interval, and capacity convention.

4.4 Congested interfaces

A linear interface approximation is

$$
\sum_{i,n,t} h_{\ell,i,n,t} x_{i,n,t} \leq \bar{F}_\ell, \qquad \ell \in \mathcal{L}.
$$

The coefficients h may be simple usage shares or distribution factors from a separate network model. The reference code accepts arbitrary signed coefficients and a limit. It does not derive them and does not perform power flow. If the coefficients are invented, the resulting shadow price is only a scenario artifact.

4.5 Bounds and ramp limits

Each workload option has

$$0 \le x_{i,n,t} \le \bar{x}_{i,n,t}.$$

Aggregate workload ramp limits are

$$-D_i \le q_{i,t} - q_{i,t-1} \le U_i.$$

A historical initial level $q_{i,t_{\text{start}}-1}$ must be fixed when the first modeled interval is ramp constrained. The reference code accepts this value as `initial_energy_mwh`. If it is omitted, ramp limits apply only between modeled intervals and no claim is made about the transition into the first one. A computing workload may have small U_i because abrupt queue changes violate a service target. A continuous industrial process may have small D_i because rapid curtailment damages output. A proof-of-work workload may be electrically curtailable but still face restart, pool, or revenue effects. The coefficients are application data, not properties of the sector label.

4.6 Deadlines and minimum service

Let $d_{i,k}$ be a deadline and $Q_{i,k}$ the minimum cumulative energy required for service milestone k:

$$\sum_{n \in \mathcal{N}} \sum_{t \le d_{i,k}} x_{i,n,t} \ge Q_{i,k}.$$

The constraint can represent energy needed to finish a batch, charge a fleet, or meet a production target. Energy is only a valid completion proxy when the workload has a stable conversion between scheduled energy and service. If efficiency changes with power, temperature, or state of charge, a richer service constraint is required.

4.7 Chance-constrained reliability extension

Suppose delivered energy is $A_{i,n,t}x_{i,n,t}$, where A is random availability. A service reliability constraint can be written

$$\mathbb{P}_\pi \left[\sum_{n \in \mathcal{N}} \sum_{t \le d_{i,k}} A_{i,n,t} x_{i,n,t} \ge Q_{i,k} \right] \ge 1 - \alpha_{i,k}.$$

Equation (4.7) is not automatically convex. It can have convex deterministic equivalents under specific distributional assumptions, or it can be replaced by a conservative robust set. The theorem below applies only when the chosen representation keeps the feasible set convex.

4.8 Objective

The allocator solves

$$\max_{x \in \mathcal{X}} \quad F(x) = \sum_{i \in \mathcal{I}} B_i(x_i) - \sum_{i,n,t} c_{n,t} x_{i,n,t},$$
$$\mathcal{X} = \{x : \text{Equations (4.3) to (4.6) and any convex reliability constraints hold}\}. \tag{2}$$

The objective is incremental surplus relative to the declared counterfactual. Fixed costs cancel only when the decision does not change them. Startup, shutdown, hardware expansion, and discrete staffing can make the objective or feasible set nonconvex.

5 Assumptions

The equalization theorem needs all of the following.

Assumption 5.1 (Price taking)**.** The allocator treats $c_{n,t}$ and all external price components as fixed with respect to its own decision. It cannot influence the market price, its tariff, or the emissions price.

Assumption 5.2 (Concave benefit)**.** Each risk-adjusted benefit B_i is differentiable and concave over the relevant domain. Strict concavity is used when uniqueness is claimed.

Assumption 5.3 (Convex feasibility)**.** The set $\mathcal{X}$ is nonempty, closed, and convex. Inequality constraints are convex and equality constraints are affine. Binary startup decisions, minimum-run times, and other indivisibilities are absent or already convexified.

Assumption 5.4 (Constraint qualification)**.** Slater's condition holds for the nonlinear inequality representation, or a weaker applicable qualification is verified. Linear constraints may use the standard relative-interior form when equalities fix a lower-dimensional affine set.

Assumption 5.5 (Information)**.** The functions, prices, emissions terms, and constraint parameters used in the optimization are known or fixed as scenario inputs. Forecast error is included only through the declared expectation, risk functional, robust set, or chance constraint.

Assumption 5.6 (Matched boundary)**.** Benefits and costs refer to the same incremental decision, time horizon, node, energy stage, and counterfactual. No market or emissions component is counted twice.

The assumptions are strong. They are written separately because each failure has a different consequence. Market power changes the price-taking first-order condition. A binary startup decision breaks convexity. Missing response data blocks empirical implementation. An average emissions rate changes the estimand. One generic caveat would hide these distinctions.

6 Marginal equalization under constraints

Write the feasible set in the compact form

$$g_k(x) \le 0, \quad k = 1, \ldots, K, \qquad Hx = h,$$

and define the minimization objective $f(x) = -F(x)$. Let $\lambda_k \ge 0$ be inequality multipliers and ν equality multipliers.

Theorem 6.1 (Conditional marginal equalization)**.** *Under the assumptions in Section 5, a feasible allocation $x^\star$ is optimal if and only if there exist multipliers $\lambda^\star \geq 0$ and $\nu^\star$ such that*

$$\nabla f(x^\star) + \sum_{k=1}^{K} \lambda_k^\star \nabla g_k(x^\star) + H^\top \nu^\star = 0, \tag{3}$$

$$g_k(x^\star) \leq 0, \tag{4}$$

$$\lambda_k^\star \geq 0, \tag{5}$$

$$\lambda_k^\star g_k(x^\star) = 0. \tag{6}$$

For an interior workload variable $x^\star_{i,n,t}$, stationarity gives

$$\frac{\partial B_i}{\partial x_{i,n,t}}(x_i^\star) = c_{n,t} + \sum_{k=1}^{K} \lambda_k^\star \frac{\partial g_k}{\partial x_{i,n,t}}(x^\star) + (H^\top \nu^\star)_{i,n,t}. \tag{7}$$

Thus risk-adjusted marginal workload value equals the full nodal and intertemporal opportunity cost.

Proof. Concavity of F makes $f = -F$ convex. The feasible set is convex by assumption. Under the stated constraint qualification, convex duality gives necessity of the KKT conditions. Their sufficiency follows from convexity. Equation (7) is the coordinate form of Equation (3), after moving the linear external cost $c_{n,t}$ to the right side. Interior status removes the lower-bound and upper-bound multipliers for that coordinate. No other multiplier is discarded. □

Corollary 6.2 (Same-footprint equalization)**.** *Consider two interior workload variables at the same node and interval. If their coefficients in every binding constraint are identical, their risk-adjusted marginal benefit values are equal.*

Proof. The right side of Equation (7) is identical for both variables. Therefore the left sides are equal. □

Corollary 6.3 (Capacity shadow price)**.** *If the node-interval capacity constraint in Equation (4.3) is binding and regularity holds locally, its multiplier equals the first-order increase in optimized surplus from one additional MWh of dispatchable capacity.*

Proof. Let b be the right side of the capacity constraint. The envelope theorem for the constrained value function gives $\partial F^\star / \partial b = \lambda^\star$, using the sign convention $Ax \leq b$. □

Corollary 6.4 (Deadline opportunity value)**.** *The multiplier on the deadline inequality measures the local cost of requiring one more MWh by that deadline. It enters earlier and later workload first-order conditions with different signs or not at all, so an internal deadline can rationally overturn a ranking based on the electricity price alone.*

6.1 What "equalization" does not mean

The theorem does not say that gross value per MWh is equal across sectors. It concerns marginal, risk-adjusted benefit after accounting for every binding constraint. It does not say that two workloads at different nodes face the same opportunity cost. It does not say that a workload at a lower average tariff should run first. It also does not claim that an observed market allocation is socially optimal.

The theorem is local. If an extra unit triggers a discrete hardware purchase, starts a generator, crosses a demand-charge threshold, or changes a market price, the derivative can jump. A mixed-integer or equilibrium model is then needed.

7 A worked quadratic model

The reference implementation uses one strictly concave term per workload option:

$$B_j(x_j) = a_j \left(v_j x_j - \frac{\kappa_j}{2} x_j^2 \right) - (1 - a_j) f_j x_j.$$

Let c_j be the nodal cost from Equation (1). Define

$$r_j = a_j v_j - (1 - a_j) f_j - c_j, \qquad q_j = a_j \kappa_j > 0.$$

The program becomes

$$\max_x \sum_{j=1}^{J} \left(r_j x_j - \frac{q_j}{2} x_j^2 \right) \quad \text{subject to} \quad Ax \le b. \tag{8}$$

For an active constraint set S, the KKT equations are

$$Qx - r + A_S^\top \lambda_S = 0, \qquad A_S x = b_S.$$

Since $Q = \operatorname{diag}(q_1, \ldots, q_J)$ is positive definite,

$$x = Q^{-1}(r - A_S^\top \lambda_S),$$

and the multipliers solve

$$A_S Q^{-1} A_S^\top \lambda_S = A_S Q^{-1} r - b_S. \tag{9}$$

The software enumerates linearly independent candidate active sets for small problems, solves Equation (9) with Gaussian elimination, and keeps the feasible candidate with greatest surplus. This is transparent and dependency free. It is exponential in the number of constraints and should not be used for operational dispatch.

7.1 Two-workload example

Suppose two interior workloads share 10 MWh at one node. Their gross marginal value curves are

$$80 - 4x_A \quad \text{and} \quad 70 - 2x_B$$

in declared currency units per MWh. The nodal energy price is 20, and there are no other terms. Capacity binds, so $x_A + x_B = 10$. Equalization gives

$$80 - 4x_A - 20 = 70 - 2x_B - 20 = \lambda.$$

Solving yields

$$x_A = 5, \qquad x_B = 5, \qquad \lambda = 40.$$

The two workloads receive the same allocation even though workload A starts with the higher intercept, because its marginal value falls faster. A sector table based on one average ratio would miss this diminishing-return effect.

7.2 Finite-difference verification

Let $F^\star(C)$ be optimized surplus under capacity C. For a small $\epsilon > 0$,

$$\widehat{\lambda}_\epsilon = \frac{F^\star(C+\epsilon) - F^\star(C)}{\epsilon}.$$

The software test uses $\epsilon = 10^{-5}$ MWh and compares $\widehat{\lambda}_\epsilon$ with the KKT multiplier. Agreement within numerical tolerance checks both the dual sign convention and the optimizer. It does not validate the economic inputs.

8 Nodal opportunity cost in detail

8.1 Energy component

The energy component is the cost of supplying another unit at the reference location before congestion and marginal losses. In an organized market the component comes from security-constrained economic dispatch. In a bilateral or vertically integrated setting, the relevant short-run avoided cost may come from a different calculation. The notation p^E does not make the sources interchangeable.

8.2 Congestion component

Congestion is a scarcity effect. When a line or interface is binding, an incremental withdrawal at one node can require redispatch elsewhere. In the allocator, congestion may appear once in the observed nodal price, or through an explicit interface constraint whose multiplier is allocated by a shift coefficient. Including both without reconciliation double counts the same constraint.

The explicit interface form is useful for an internal campus or feeder problem. Suppose a feeder serves two buildings and has 8 MWh of incremental headroom. The constraint

$$x_{\text{north},t} + x_{\text{south},t} \le 8$$

generates one shadow value. If the two loads have different electrical effects, the coefficients change. Any physical claim then depends on the source of those coefficients.

8.3 Loss component

Marginal losses differ from average losses because line losses are nonlinear in current. A nodal price can contain a marginal-loss component. If an analyst instead starts from generator-side energy and converts to delivered load, the loss convention must be explicit. Multiplying an LMP by an additional average loss factor is usually inconsistent.

8.4 Reserve and reliability terms

Reserve can be represented as a capacity deduction, a price, or both when the two objects cover distinct services. The capacity deduction protects physical headroom. The price represents the cost of procuring or consuming a reserve product. A single undifferentiated "reliability adder" can conceal overlap.

The model uses $r_{n,t}$ for withheld capacity, $m_{n,t}$ for a reliability margin, and $p^R_{n,t}$ for a marginal reserve charge. Each is optional. If they come from the same requirement, the analyst must provide a bridge showing why there is no duplicate cost.

8.5 Intertemporal opportunity cost

The true opportunity cost of running now can be larger or smaller than the current electricity price. Running now may consume a limited daily energy budget, make a later ramp infeasible, or satisfy a deadline that would otherwise become expensive. Ramp and deadline multipliers carry this intertemporal value through Equation (7).

This is why "move all computation to the cheapest hour" is not a theorem. The cheap interval may lack network capacity. The workload may need an earlier partial run to respect a ramp-up limit. Delay may reduce completion probability. The cost ranking is only one input.

9 Marginal emissions and environmental opportunity cost

9.1 A derivative tied to a counterfactual

Let $G_{n,t}(\delta)$ be total system emissions when incremental load δ is added at node n and time t, relative to a named dispatch counterfactual. The local marginal factor is

$$g_{n,t} = \left.\frac{\partial G_{n,t}(\delta)}{\partial \delta}\right|_{\delta=0}.$$

Its unit might be kg CO_2 per MWh. A damage value or carbon price χ_t gives the term $\chi_t g_{n,t}$ in Equation (1).

This scalar is already a compression. Separate pollutants have different locations, damages, and time profiles. A serious social-cost application should report a vector before applying monetary weights:

$$g_{n,t} = \left(g^{\mathrm{CO_2}}_{n,t}, g^{\mathrm{NO_x}}_{n,t}, g^{\mathrm{SO_2}}_{n,t}, g^{\mathrm{PM}}_{n,t}\right).$$

9.2 What AVERT supplies

EPA describes AVERT as a tool for evaluating emissions changes associated with energy policies, programs, and projects [11]. The statistical module uses historical hourly generation and emissions data. EPA's published avoided-emission-rate products provide regional approximations for specified resource categories [12]. These data can support a bounded scenario or screening calculation.

AVERT does not supply a universal node-level factor for any future workload. Its regions are larger than market nodes. Historical relationships may not transport to a system with different fuel prices, retirements, transmission, storage, or weather. A large new load may alter investment, not merely dispatch. The paper therefore labels an AVERT-based term with its region, data year, scenario class, and uncertainty.

9.3 Average-factor ranking reversal

Consider two nodes. Node West has a low electricity price but a high marginal emissions factor. Node East has a higher price but a low factor. With no emissions value, a flexible batch prefers West. Once χg exceeds the price difference, East becomes cheaper in social marginal terms. The code test constructs exactly this reversal.

The example is algebraic. It does not assert that any named region has those numbers. A credible empirical ranking would require matched nodal prices, marginal emissions, workload performance, data-movement energy, and response curves over the same intervals.

10 Reliability, uncertainty, and recourse

10.1 Expected value is not enough in every application

An expected-surplus objective treats upside and downside through their probabilities and values. That may be adequate for many repeated, small jobs. It may be inadequate for a safety-critical industrial batch or a vehicle fleet that must be ready by a fixed hour. A deadline chance constraint or distributionally robust requirement can be more appropriate.

Three choices must remain separate:

1. forecast uncertainty about prices, emissions, availability, and demand;
2. risk preference, such as expected shortfall or a chance constraint;
3. model uncertainty about whether the response curve or probability law is correct.

A confidence interval for an estimated price forecast does not automatically describe tail risk. A robust set does not automatically have a frequency interpretation. The manifest must name the choice.

10.2 Two-stage extension

Let x be a day-ahead schedule and $y(\omega)$ real-time recourse. One extension is

$$\max_x \left\{ B^{DA}(x) - c^{DA\top} x + \mathbb{E}_\pi \left[\max_{y(\omega)\in\mathcal{X}(x,\omega)} B^{RT}(y,\omega) - c^{RT}(\omega)^\top y \right] \right\}.$$

The formulation captures cancellation, migration, or emergency curtailment after uncertainty resolves. It also makes clear that a deterministic shadow price is conditional on the scenario and information set.

10.3 Reliability value versus reserve value

Reliability can produce private and system value. A workload may avoid a service failure. The grid may preserve operating headroom. These are different effects. The reference code's `failure_loss_per_mwh` parameter enters workload benefit. `reliability_margin_mwh` reduces node capacity. Neither parameter estimates loss of load probability or the value of lost load.

10.4 Correlation

Independent availability assumptions can materially understate risk when many workloads depend on one substation, network carrier, cloud region, or weather event. A common-shock model is needed for portfolio reliability. The small solver does not model correlation. Treating its reliability factors as independent probabilities would be an unsupported extension.

11 Illustrative sector scenarios

11.1 Scenario status

The scenarios in this section are synthetic. They test model behavior and expose missing data. The coefficients are not estimates of national or firm response. In particular, the paper has no credible common-scale marginal value curves for AI, cryptographic settlement, manufacturing, and transport. Any sector ranking is therefore OBSTRUCTED.

Table 2: Scenario interpretation and principal missing evidence

Workload	Flexible decision	Binding risk	Evidence needed for an empirical allocation
AI inference	Route requests across nodes or defer low-priority batches	Latency, quality, failed requests, data locality	Causal service value curve, wall energy, routing overhead, completion and rework data
AI training	Shift checkpoints or batch stages across intervals	Deadline, checkpoint loss, cluster ramp, hardware occupancy	Marginal training progress, power trace, restart costs, value of completion time

Workload	Flexible decision	Binding risk	Evidence needed for an empirical allocation
Proof of work	Curtail hashrate or move operations	Settlement revenue, hardware cycling, pool and network conditions	Incremental revenue and assurance under a named threat model, not gross transaction value
Manufacturing	Reschedule or modulate a process	Material continuity, thermal limits, labor and inventory	Plant-specific production function, quality loss, restart cost, contractual delivery value
Transport	Shift fleet charging	State of charge, route readiness, charger limits, battery degradation	Vehicle-level schedules, charging efficiency, degradation and service-loss values

11.2 Scenario A: one node, two smooth workloads

The quadratic example in Section 7 verifies the basic equalization condition. Both workloads are interior and share one binding capacity constraint. The result is a clean theorem fixture. It has almost no sector realism.

11.3 Scenario B: reserve and reliability deductions

Consider three intervals with 8 MWh of gross incremental availability in each. One MWh is withheld for reserve and one MWh for a reliability margin, leaving 6 MWh dispatchable. A batch workload has a 9 MWh completion requirement, a 6 MWh option bound per interval, and a 1 MWh ramp limit between intervals. A shared feeder caps total batch use at 12 MWh.

The executable fixture checks:

- every interval remains below the 6 MWh dispatchable limit;
- adjacent allocations differ by no more than 1 MWh;
- cumulative energy meets the 9 MWh deadline;
- the feeder limit is respected;
- all KKT residuals are below tolerance.

The point is not that 1 MWh is a sensible reserve for a real grid. The point is that reserve, reliability, ramp, and deadline are distinct constraints with distinct multipliers.

11.4 Scenario C: emissions-sensitive location

Two nodes have identical workload response curves. West costs 20 currency units per MWh and has a scenario marginal factor of 800 kg per MWh. East costs 35 and has 100 kg per MWh. An 8 MWh shared compute budget links the choices. Without a carbon value, West receives more energy. With $\chi = 0.1$ currency units per kg, East receives more, and modeled marginal emissions fall.

This is a rank reversal caused by a declared social-cost term. It is not a finding about a real West or East region. Using an annual average factor could produce a different and potentially misleading ranking.

11.5 Scenario D: reliability-adjusted workload value

A fragile service has a high gross value intercept but only 0.5 reliability and a failure loss. A reliable service has a lower gross intercept but a higher risk-adjusted marginal benefit. Under shared capacity, the reliable service receives more energy, while the fragile service remains positive in the chosen fixture. The test confirms that expected failure loss appears separately in the objective decomposition.

11.6 Scenario E: infeasible deadline

A job requires 3 MWh by the only available interval, but both its option bound and node capacity are 2 MWh. The solver returns infeasible. This result matters: an optimization pipeline must not silently violate a service contract because its objective value looks attractive.

11.7 Why no sector league table appears

A sector league table would require common monetary or welfare units, compatible energy boundaries, matched locations and intervals, response curves at the relevant margins, reliability losses, and uncertainty. The current evidence does not supply these data. Assigning plausible-looking curves would replace the research question with invented answers.

The paper instead supplies a protocol. A future empirical study can populate one sector at a time, publish the response data, and report where cross-sector comparison remains obstructed.

12 Empirical protocol

12.1 Step 1: state the decision

Write the intervention as a quantity, location, interval, and alternative. For example: "allocate an additional 0.5 MWh to batch inference at node N17 between 02:00 and 02:15, rather than defer the batch to the next admissible window." An annual sector label is not a decision specification.

12.2 Step 2: measure the response curve

Estimate workload value as a function of incremental energy over the relevant range. Randomized throttling, queue experiments, staggered price exposure, or engineering response tests may identify part of the curve. Record quality, failure, latency, and rework. Do not regress revenue on electricity and call the coefficient causal without a design.

For AI, the service unit should be a verified task outcome or another declared quality measure, not tokens alone. For proof of work, revenue and settlement assurance must remain separate. For manufacturing, record scrap and quality changes. For transport, record service readiness and battery effects.

12.3 Step 3: obtain marginal electricity terms

Collect interval-matched energy, congestion, and loss components where they exist. State whether the load actually faces those prices or whether they are social opportunity-cost inputs. Reconcile reserve, capacity, demand, and fixed charges. A bill total divided by MWh is an average.

12.4 Step 4: model physical constraints

Record deliverable headroom, feeder or interface limits, ramp capability, location restrictions, downtime, restart cost, and deadlines. Verify that energy is a valid service proxy. If not, express constraints in the actual service unit and link energy through a measured conversion function.

12.5 Step 5: estimate marginal emissions

Choose a method matched to scale and horizon. AVERT may support regional screening based on historical dispatch patterns. Unit-level econometric, dispatch, or production-cost models may be needed for other questions. Investment-induced emissions require a long-run model. Publish the pollutant vector and uncertainty before monetization.

12.6 Step 6: solve and audit

Report primal decisions, objective decomposition, every constraint slack, and every shadow price. Check KKT residuals for a convex model. Perturb each important right-hand side and compare finite differences with its multiplier. For a mixed-integer model, report the integrality gap and local sensitivity without labeling a discrete jump as a derivative.

12.7 Step 7: test sensitivity

At minimum vary:

- workload value-curve parameters and failure loss;
- marginal rather than average price and emissions inputs;
- reserve and reliability deductions;
- deadline and ramp bounds;
- price-taking versus endogenous price response;
- short-run versus capacity-expansion horizons;
- environmental weights and pollutant scope.

If a ranking changes under credible alternatives, report the range or Pareto set. Do not hide the reversal in an appendix.

13 Reference implementation

13.1 Objects and units

The module `joule_standard.marginal_allocation` contains:

- `NodalPrice`, with energy, congestion, losses, reserve, marginal emissions, and emissions-price components;
- `NodeInterval`, with available capacity, reserve, and reliability deductions;
- `WorkloadSlice`, with node, interval, concave benefit, upper bound, reliability, and failure loss;
- `DeadlineRequirement`, `RampLimit`, and `InterfaceLimit`;
- `solve_allocation`, the active-set solver;
- `finite_difference_shadow_price` and `check_kkt`.

Energy is in MWh. Monetary coefficients use one declared currency base per MWh. Emissions use kg per MWh, and the emissions value uses currency per kg. The code checks finiteness, nonnegativity where required, matching node-time keys, unique slice identifiers, and valid constraint references.

13.2 Active-set enumeration

Bounds and operational requirements are compiled into $Ax \leq b$. For each candidate active set with at most J linearly independent constraints, the solver:

1. forms the Gram matrix in Equation (9);
2. solves for candidate multipliers;
3. rejects negative multipliers;
4. reconstructs the allocation;
5. rejects primal violations;
6. keeps the feasible point with largest surplus.

The routine stops if the number of candidate sets exceeds a declared cap. This guard prevents an apparently small example from turning into an accidental combinatorial workload.

13.3 KKT audit

The audit computes four residuals:

$$
\begin{aligned}
r_{\text{primal}} &= \max_k [A_k x - b_k]_+, \\
r_{\text{dual}} &= \max_k [-\lambda_k]_+, \\
r_{\text{stationarity}} &= \left\| Qx - r + A^\top \lambda \right\|_\infty, \\
r_{\text{complementarity}} &= \max_k \left| \lambda_k (b_k - A_k x) \right|.
\end{aligned}
$$

All must lie below the requested tolerance.

13.4 Test contract

The focused test file contains ten tests:

1. interior marginal values equal the binding capacity shadow price;
2. the capacity multiplier matches a finite-difference value derivative;
3. reserve, reliability margin, ramp, deadline, and interface constraints are feasible;
4. an impossible deadline raises an infeasibility error;
5. marginal emissions pricing reverses a two-node dispatch;
6. reliability and failure loss change the allocation;
7. binding and slack constraints satisfy complementary slackness;
8. slice-level decisions aggregate correctly by workload;
9. signed energy, congestion, and loss components are accepted;
10. an optional historical initial ramp state constrains the first modeled interval.

The tests establish COMPUTATIONAL evidence for the implementation. They do not prove that a real plant, fleet, mine, or data center has the scenario coefficients.

14 Failure modes outside the theorem

14.1 Market power

A large flexible load may influence the market price or anticipate its effect on scarcity. Then $c_{n,t}$ depends on x, and price taking fails. A strategic bidder maximizes against a residual supply curve, not a fixed vector. The KKT machinery can still apply to a suitably convex private problem, but the welfare interpretation changes.

14.2 Nonconvex operations

Unit commitment, minimum-run constraints, startup costs, integer servers, minimum batch sizes, and all-or-nothing deadlines create nonconvexities. An interior marginal equalization rule can fail because no nearby feasible point exists. Convex-hull prices, uplift, or mixed-integer methods may be needed.

14.3 Endogenous capacity

The short-run model treats node capacity as given. Persistent automated load can trigger generation, transmission, distribution, or data-center investment. Long-run marginal cost then includes capital and embodied effects. A short-run LMP should not be extrapolated into an investment theorem.

14.4 Distribution network limits

A wholesale node may hide distribution constraints. Thermal feeder limits, voltage, phase imbalance, and protection settings can determine whether an incremental load is deliverable. A single interface proxy does not validate those conditions.

14.5 Dynamic stability

The reference model does not simulate frequency, voltage, inverter controls, or transient behavior. Even a full DC optimal power flow would not prove dynamic stability. Claims about grid support require separate engineering analysis.

14.6 Baseline error

Marginal value and emissions require a counterfactual. A deferred batch might run later rather than disappear. Curtailed mining may shift to another site. Delayed manufacturing may incur overtime. Vehicle charging may move into a more polluting interval. Counting only the immediate reduction overstates the effect if the load reappears.

14.7 Telemetry and attribution

Facility meters may combine many loads. Allocating shared cooling or network energy by server time, peak power, or direct metering can change the result. The attribution rule must be published. An unexplained proportional allocation is not a measurement.

15 Policy interpretation

15.1 A benchmark, not an automatic mandate

The equalization rule describes an optimum inside a declared model. It does not tell a regulator which value functional to adopt or how to weight pollution, reliability, employment, privacy, or distribution. Those choices enter the numerator, constraints, or prices.

The model can still clarify policy disagreements. If two analysts disagree because one uses an average retail rate and the other uses nodal marginal cost, that is an estimand disagreement. If they disagree about the carbon weight, that is a valuation disagreement. If one includes an industrial deadline and the other does not, that is a feasible-set disagreement. These should not be blended into one unexplained score.

15.2 Flexible load programs

A flexible-load program can improve dispatch only if the load responds when and where flexibility is useful, respects telemetry and control rules, and does not shift costs or emissions outside the measured window. Compensation should be tied to an auditable baseline and delivered service. The paper does not design that baseline or settlement rule; Paper 8 in the program addresses market design.

15.3 Environmental dispatch

Adding χg to nodal cost is one transparent way to represent marginal emissions damages. Other approaches include emissions caps, clean-energy constraints, or multi-objective frontiers.

A monetized scalar can be useful, but the underlying pollutant vector and weights should remain visible.

15.4 Reliability and essential service

Some workloads may receive minimum allocations because they provide essential service. That policy appears as a lower bound or deadline, and its shadow price reports the local opportunity cost. The multiplier does not decide whether the requirement is legitimate. It makes the resource tradeoff inspectable.

16 Limitations

Limitation 16.1 (No estimated sector response curves). The paper has no matched marginal benefit functions for AI, proof of work, manufacturing, and transport. The numerical sector examples are synthetic.

Limitation 16.2 (Price-taking benchmark). The theorem does not cover strategic demand, endogenous tariffs, or a load large enough to change market prices.

Limitation 16.3 (Linear operational constraints). The implementation uses linear capacity, interface, ramp, deadline, and bound constraints. It does not solve AC power flow, unit commitment, or dynamic stability.

Limitation 16.4 (Simplified reliability). The code uses one service-reliability factor, one expected failure-loss coefficient, and explicit capacity deductions. It omits correlated outages, recourse, loss of load probability, and state-dependent recovery.

Limitation 16.5 (Marginal emissions uncertainty). EPA AVERT and published marginal-factor studies support better screening than an annual average factor, but they do not remove model error, geographic aggregation, or long-run investment effects.

Limitation 16.6 (Short-run energy boundary). The model allocates marginal operational electricity. It does not include embodied energy or capacity investment unless those enter a separately stated long-run extension.

Limitation 16.7 (No distributional welfare result). A sum of monetized benefits can hide who gains and who bears cost. The paper does not supply distributional weights or a social welfare function.

Limitation 16.8 (Small solver). Active-set enumeration is suitable for auditable fixtures only. Large operational problems require a tested optimization package and independent engineering validation.

17 Research agenda

17.1 Estimating workload response

The central empirical task is to estimate marginal service value under controlled energy variation. For AI workloads, randomized queue throttling can measure completion, latency, quality, and rework. For manufacturing, plant experiments need product-quality safeguards. For fleet charging, route readiness and battery degradation belong in the outcome. Proof-of-work studies must distinguish private mining revenue from network assurance.

17.2 Linking flexible loads to grid nodes

Public workload data rarely identify electrical nodes. Data-center region labels are too coarse for congestion studies. Partnerships with utilities, system operators, and large users could produce privacy-preserving data that join meter location, interval load, price components, constraints, and workload service.

17.3 Long-run investment

Persistent automated demand changes generation, transmission, and data-center capacity. A long-run model should include discrete investment, lead times, embodied energy, and demand uncertainty. Short-run and long-run marginal value should be reported separately.

17.4 Distribution and governance

An automated allocation rule can move pollution, congestion, and economic activity across communities. A vector account should report local emissions, reliability exposure, consumer effects, and workload outcomes before any weighted scalar is formed.

17.5 Robust optimization

Marginal emissions, prices, and workload value curves are estimated. Robust sets or distributionally robust methods can show which allocations survive credible misspecification. Their conservatism should be calibrated against held-out operating periods, not chosen for aesthetic smoothness.

18 Conclusion

The next unit of electricity has an address and a time. Its allocation depends on what the workload produces at the margin, what the grid gives up at that node and interval, which constraints bind, how reliability is represented, and whether marginal emissions are valued. Annual energy intensity and average tariffs do not answer that question.

Under explicit concavity, convexity, price-taking, information, and boundary assumptions, the allocation problem has a precise KKT rule. Risk-adjusted marginal workload value equals nodal and intertemporal opportunity cost. Shadow prices report the local value of relaxing capacity, congestion, ramp, and deadline constraints. The result is useful because its failure conditions are visible.

The executable model verifies the algebra on small scenarios. It also exposes the empirical gap. A cross-sector allocation cannot be inferred from labels such as AI, manufacturing, transport, or settlement. It requires credible response curves and matched marginal grid data. Until those exist, sector rankings remain obstructed.

A Full KKT derivation for linear constraints

Consider the quadratic program

$$\max_x \quad r^\top x - \frac{1}{2}x^\top Qx \quad \text{subject to} \quad Ax \le b,$$

where Q is diagonal and positive definite. Convert it to minimization:

$$\min_x \quad \frac{1}{2}x^\top Qx - r^\top x.$$

The Lagrangian is

$$\mathcal{L}(x, \lambda) = \frac{1}{2}x^\top Qx - r^\top x + \lambda^\top(Ax - b), \qquad \lambda \ge 0.$$

Stationarity gives

$$Qx - r + A^\top\lambda = 0.$$

For an active set S, inactive multipliers are zero and

$$x = Q^{-1}(r - A_S^\top\lambda_S).$$

Substitution into $A_S x = b_S$ gives

$$A_S Q^{-1} A_S^\top \lambda_S = A_S Q^{-1} r - b_S.$$

If the active rows are independent, the system has a unique multiplier vector. A candidate is optimal when it also satisfies

$$\lambda_S \ge 0, \qquad Ax \le b.$$

Strict convexity of the minimization objective gives a unique primal optimum. Dual multipliers need not be unique when active constraints are redundant. The solver can assign zero to a redundant binding constraint and retain a valid KKT certificate.

A.1 Bounds

The upper bound $x_j \le \bar{x}_j$ is a row of A. The lower bound $x_j \ge 0$ is written $-x_j \le 0$. If neither binds, both multipliers are zero. If $x_j = 0$, the lower-bound multiplier can absorb a negative unconstrained marginal surplus. Thus zero allocation is compatible with the KKT rule and does not require equal marginal value with an interior workload.

A.2 Deadline sign convention

A minimum requirement

$$\sum_{j \in D} x_j \ge Q$$

is compiled as

$$-\sum_{j \in D} x_j \le -Q.$$

Relaxing the compiled right side means lowering the required minimum. The associated multiplier is the local increase in optimized surplus from that relaxation. Conversely, tightening the physical deadline requirement by one MWh reduces surplus by the same first-order amount.

B Unit and boundary crosswalk

Table 3: Units and common category errors

Object	Unit	Common error
$x_{i,n,t}$	MWh per interval	Treating power in MW as energy without multiplying by interval duration
$v_{i,n,t}$	Currency per MWh	Calling an average revenue ratio marginal value
$\kappa_{i,n,t}$	Currency per MWh squared	Omitting the range over which the local quadratic is plausible
p^E, p^C, p^L, p^R	Currency per MWh	Adding LMP components twice
$g_{n,t}$	kg pollutant per MWh	Substituting annual average grid intensity for a marginal factor
χ_t	Currency per kg pollutant	Hiding the damage value or policy weight
λ_k	Currency per unit of constraint RHS	Interpreting a local shadow value as an average price
$a_{i,n,t}$	Probability or availability fraction	Ignoring correlation or treating a score as a probability
$f_{i,n,t}$	Currency per failed MWh	Calling gross protected value an expected loss

C Illustrative input ledger

Table 4: Synthetic fixtures used by the test suite

Fixture	Capacity	Price or cost	Key constraint	Expected computational behavior
Two workloads	10 MWh	20 per MWh	Shared node capacity	Interior marginal net values equal capacity multiplier
Three intervals	6 MWh dispatchable each	10 per MWh	9 MWh deadline, 1 MWh ramp, 12 MWh feeder	Feasible schedule and KKT residuals below tolerance
Impossible job	2 MWh	10 per MWh	3 MWh minimum	Infeasibility exception

Fixture	Capacity	Price or cost	Key constraint	Expected computational behavior
Two nodes	8 MWh each	20 versus 35 per MWh	8 MWh shared compute budget	Carbon value reverses location preference
Reliability	5 MWh	20 per MWh	Shared node capacity	Failure loss changes risk-adjusted allocation

The ledger is deliberately incomplete as empirical data. It lists only the parameters needed to reproduce software behavior. A publishable field study would add provenance, timestamps, meter boundaries, price-source identifiers, emissions-model version, workload sample, uncertainty, and counterfactual.

D Reproduction checklist

1. Use Python 3.11 or newer.
2. Run `python3 -m pytest tests/test_marginal_allocation.py -q`.
3. Confirm that all focused tests pass.
4. Run the full repository test suite.
5. Inspect every reported KKT residual.
6. Perturb each binding right-hand side and compare the optimized-value difference with the stored shadow price.
7. Re-run the emissions reversal with the emissions price set to zero.
8. Increase an impossible deadline and confirm that the solver remains infeasible rather than violating a constraint.
9. Keep all scenario inputs labeled synthetic.

E Audit questions for an empirical allocation

Table 5: Pre-publication audit

Question	Failure status
Is the decision incremental, node specific, and time specific?	If no, average accounting has replaced marginal allocation.
Does the value curve share the same counterfactual as the energy change?	If no, the numerator and denominator do not describe one intervention.
Are fixed and marginal bill components separated?	If no, nodal opportunity cost is not identified.
Are congestion and losses already inside the LMP?	If unknown, double counting is possible.

Question	Failure status
Are reserve capacity and reserve charges distinct?	If no, reliability may be counted twice.
Is the emissions factor marginal for the stated horizon and geography?	If no, environmental ranking is obstructed.
Are deadlines, ramps, bounds, and location choices measured?	If no, the feasible set is speculative.
Does the KKT theorem's convexity assumption hold?	If no, report a nonconvex method and do not claim marginal equalization.
Can the load influence price?	If yes, price taking fails.
Are all sector value curves estimated in compatible units?	If no, a cross-sector ranking is obstructed.
Does the model include dynamic grid stability?	If no, make no stability claim.

F Formal claim ledger

Claim JS-C010. In a convex price-taking allocation model, an interior optimum equalizes risk-adjusted marginal net value with nodal and intertemporal opportunity cost. Status: PROVED under Section 5.

Shadow-price sensitivity. A regular binding constraint multiplier is the local value of relaxing its right-hand side. Status: PROVED.

Finite-difference agreement. The reference capacity multiplier agrees with an optimized-value finite difference within test tolerance. Status: COMPUTATIONAL.

Scenario feasibility. The reserve, ramp, deadline, interface, and reliability fixtures satisfy their declared constraints. Status: COMPUTATIONAL.

Marginal-emissions reversal. A synthetic two-node case reverses its location ranking when a declared emissions value is introduced. Status: COMPUTATIONAL.

Universal sector ranking. The available evidence identifies a common marginal-value ordering across AI, proof of work, manufacturing, and transport. Status: OBSTRUCTED.

Dynamic grid benefit. The convex load allocation proves frequency or voltage stability. Status: OBSTRUCTED.

G Data template

An empirical release should publish one row per workload option, node, and interval with at least:

- stable workload, node, and interval identifiers;
- interval start, duration, and time zone;
- energy decision unit and meter boundary;
- gross marginal benefit estimate and uncertainty;

- curvature or nonparametric response representation;
- maximum and minimum feasible energy;
- service reliability and failure-loss definition;
- energy, congestion, loss, and reserve price components;
- marginal emissions by pollutant and model version;
- capacity, reserve, reliability, ramp, interface, and deadline inputs;
- counterfactual, attribution rule, and exclusion list.

Publishing only the optimized allocation is insufficient. The result cannot be audited without the inputs, active constraints, slacks, multipliers, and uncertainty model.

References

[1] P. A. Samuelson. *Foundations of Economic Analysis*. Harvard University Press, 1947.

[2] M. Boiteux. Peak-load pricing. *The Journal of Business*, 33(2):157 to 179, 1960.

[3] F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R. E. Bohn. *Spot Pricing of Electricity*. Kluwer Academic Publishers, 1988. https://doi.org/10.1007/978-1-4613-1683-1.

[4] S. Borenstein and S. Holland. On the efficiency of competitive electricity markets with time-invariant retail prices. *RAND Journal of Economics*, 36(3):469 to 493, 2005.

[5] Federal Energy Regulatory Commission. *Energy Markets Primer*. 2024. https://www.ferc.gov/sites/default/files/2024-01/24_Energy-Markets-Primer_0117_DIGITAL_0.pdf.

[6] North American Electric Reliability Corporation. *Glossary of Terms Used in NERC Reliability Standards*. 2022. https://www.nerc.com/pa/Stand/Glossary%20of%20Terms/Glossary_of_Terms.pdf.

[7] K. Siler-Evans, I. L. Azevedo, and M. G. Morgan. Marginal emissions factors for the U.S. electricity system. *Environmental Science & Technology*, 46(9):4742 to 4748, 2012. https://doi.org/10.1021/es300145v.

[8] S. P. Holland, E. T. Mansur, N. Z. Muller, and A. J. Yates. Are there environmental benefits from driving electric vehicles? The importance of local factors. *American Economic Review*, 106(12):3700 to 3729, 2016. https://doi.org/10.1257/aer.20150897.

[9] D. S. Callaway, M. Fowlie, and G. McCormick. Location, location, location: the variable value of renewable energy and demand-side efficiency resources. *Journal of the Association of Environmental and Resource Economists*, 5(1):39 to 75, 2018. https://doi.org/10.1086/694179.

[10] United States Environmental Protection Agency. How AVERT works. Updated 2 March 2026. https://www.epa.gov/avert/how-avert-works.

[11] United States Environmental Protection Agency. AVERT User Manual, version 4.3. 2024. https://www.epa.gov/avert/avert-user-manual.

[12] United States Environmental Protection Agency. Avoided emission rates generated from AVERT. Updated 3 March 2026. https://www.epa.gov/avert/avoided-emission-rates-generated-avert.

[13] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. https://web.stanford.edu/~boyd/cvxbook/.

[14] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[15] W. Karush. *Minima of Functions of Several Variables with Inequalities as Side Constraints*. Master's thesis, University of Chicago, 1939.

[16] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481 to 492, 1951.

[17] S. Stoft. *Power System Economics: Designing Markets for Electricity*. IEEE Press and Wiley, 2002.

[18] S. Borenstein. The private and public economics of renewable electricity generation. *Journal of Economic Perspectives*, 26(1):67 to 92, 2012. https://doi.org/10.1257/jep.26.1.67.

[19] S. P. Holland, E. T. Mansur, V. Verdier, and A. J. Yates. Regularization from economic constraints: a new estimator for marginal emissions. NBER Working Paper 32065, 2024. https://doi.org/10.3386/w32065.