

Project Lab: Definitions and Math Formulas

Brian Nguyen

Summer 2026

Reference doc for the Summer 2026 Project Lab (Fama-Miller funded). Based on the proposal and supplemental info docs. This repository implements the CAISO slice of the program.

1 Workstream 1: Spreads

Collecting and computing price relationships across ISOs. These are the basic objects.

ISO (Independent System Operator). The entity that runs the grid and the wholesale electricity market for a region. Our seven: ERCOT (Texas), CAISO (California), PJM (Mid-Atlantic), NYISO (New York), MISO (Midwest), SPP (Great Plains), ISONE (New England).

DAM (Day-Ahead Market). Auction held the day before delivery. Generators and buyers commit to schedules; produces one price per hour per location.

RT (Real-Time Market). The balancing market. Every 5 minutes the ISO reprices electricity based on actual conditions. Much spikier than DAM because it absorbs forecast errors, outages, and weather surprises in real time.

Nodal Prices (LMPs). Prices are set at thousands of individual grid locations (nodes), not one region-wide price. Two nodes differ because of transmission congestion and line losses between them.

DAM-RT Spread. The financial difference between the hourly Day-Ahead Market (DAM) price and the volume- or time-weighted average of Real-Time Market (RTM) prices for the corresponding hour and electrical node.

Location Spread. Price at node A minus price at node B. Pure congestion and losses.

Time Spread. The price difference between two distinct blocks of time within the same operating day at the same electrical node.

TBx (Top-Bottom x). You take the x most expensive hours of the day, average them out, and subtract the average of the x cheapest hours at that specific spot on the grid. This represents the absolute maximum cash a storage could theoretically make in a day if it has a crystal ball.

2 Workstream 2: Quantos

HDD / CDD (Heating / Cooling Degree Days). Daily weather indices against a 65°F base:

- $HDD = \max(0, 65 - T)$: how cold it was, heat (gas) demand
- $CDD = \max(0, T - 65)$: how hot it was, A/C (electricity) demand

Contracts settle on the cumulative sum over a month or season. Note the $\max(0, \cdot)$ structure: each daily value is already an option-style payoff on temperature.

Heat rate. A gas generator's efficiency: MMBtu of gas burned per MWh of electricity produced. Lower is better. Physical floor is 3.412 (perfect conversion). A generator's marginal cost is heat rate $\times$ gas price.

Location Pairs and Hub Mapping (CAISO). Spark-spread objects are indexed by a market pair (z, g) : a power hub z and the gas hub g whose gas a generator at z actually burns. The baseline mapping $g^*(z)$, with the CAISO fuel-region series that carries each gas leg:

Power hub z	Gas leg $g^*(z)$	Fuel region series
NP15	PG&E Citygate	FRPGE1
ZP26	PG&E Citygate	FRPGE1
SP15	SoCal Citygate	FRSCE1
PACE	PacifiCorp East	FRPACE
PACW	PacifiCorp West	FRPACW

Shorthand: $S^{(z)}(H) := S^{(z, g^*(z))}(H)$. ZP26 sits between gas systems and has no unique hub: PG&E Citygate is the baseline, and Kern River / SoCal Citygate are the admissible alternates for sensitivity runs.

The gas leg is daily. Gas is nominated once per gas day, and each CAISO fuel-region price is a daily index carried forward across the day's hours.

MIHR (Market Implied Heat Rate). Power price divided by gas price:

$$\text{MIHR}_{h,d}^{(z,g)} = \frac{P_{E,h}^{(z)}}{P_{G,d}^{(g)}}$$

Units work out to MMBtu/MWh, a heat rate. Interpretation: the breakeven efficiency line. Any generator with heat rate below the MIHR is profitable to run right now (undefined when the gas print is non-positive).

Spark Spread. The gross margin of a gas generator per MWh, at strike heat rate H :

$$S_{h,d}^{(z,g)}(H) = P_{E,h}^{(z)} - H P_{G,d}^{(g)} = P_{G,d}^{(g)} (\text{MIHR}_{h,d}^{(z,g)} - H)$$

Same information as MIHR, in dollar form instead of ratio form. A gas plant is economically a real option on the spark spread: it runs when positive, idles when negative. Daily aggregation touches only the power leg (the gas leg is already daily):

$$\bar{S}_d^{(z,g)}(H) = \bar{P}_{E,d}^{(z)} - H P_{G,d}^{(g)}, \quad \bar{P}_{E,d}^{(z)} = \frac{1}{n_d} \sum_{h \in d} P_{E,h}^{(z)}$$

with n_d the number of intervals in Pacific trade day d (23-25 at hourly granularity on DST days).

Gas basis. How expensive gas is at hub g relative to the defined reference hub:

$$B_d^{(g)} = P_{G,d}^{(g)} - P_{G,d}^{(\text{PG\&E CG})}$$

Convention: all bases quote relative to PG&E Citygate. The anchor only shifts where zero sits: basis differences between hubs are anchor-invariant, and with the g^* mapping in place the anchor never enters any spread or option value.

HRCO (Heat Rate Call Option). The core instrument. Payoff per settlement interval:

$$\text{Payoff} = \max\{0, P_E - \text{HR}_{\text{strike}} \times P_G - \text{VOM}\} \times \text{Volume}$$

Equivalently (at zero VOM), a call option on the MIHR itself:

$$\text{Payoff} = P_G \times \max\{0, \text{MIHR} - \text{HR}_{\text{strike}}\}$$

VOM is a variable O&M adder (dollars per MWh of non-fuel running cost); the stylized version sets it to 0. The cost is zero for simplicity.

Profitability and dispatch. $\text{ITM}(H) = \Pr[S(H) > 0]$ is the share of intervals a heat-rate- H unit is profitable - it says nothing about how much. A unit forced to run every interval earns the average spread; one that runs only profitable intervals earns the average positive part:

$$\underbrace{E[S(H)]}_{\text{always-on}} \leq \underbrace{E[\max(0, S(H))]}_{\text{perfect dispatch}}$$

Settlement frequency defines the instrument (Jensen). $V_\bullet(H)$ is the expected payoff per MWh with $\max(0, \cdot)$ applied at the settlement frequency named in the subscript:

$$V_{\text{hourly}}(H) = E_{h,d}[\max(0, S_{h,d}(H) - \text{VOM})], \quad V_{\text{daily}}(H) = E_d[\max(0, \bar{S}_d(H) - \text{VOM})]$$

Rule: apply the payoff at the contract's settlement frequency, never at the data's observation frequency. Because $\max(0, \cdot)$ is convex, Jensen's inequality orders the conventions (within one price series; DAM and RT are different prices, not averages of each other):

$$V_{5\text{min}} \geq V_{15\text{min}} \geq V_{\text{hourly}} \geq V_{\text{daily}} \geq V_{\text{monthly}}$$

with equality only when the spread holds one sign inside the averaging window. A physical unit (start costs, minimum up/down times) sits between the brackets, $V_{\text{daily}} \leq V_{\text{unit}} \leq V_{\text{hourly}}$, by nesting of feasible dispatch sets - no convexity needed.

Intraday optionality premium. The headline object:

$$\Pi(H) = V_{\text{hourly}}(H) - V_{\text{daily}}(H) \geq 0$$

The value created purely by the ability to pick profitable hours within the day. With hourly power against daily-indexed gas, Π isolates the value of intraday power shape (duck curve, evening ramp) uncontaminated by gas noise.

Margrabe (the pricing layer). Both legs (Power and Gas) are stochastic; they are constantly changing. The pedagogical pricing model is Margrabe, where the key input is the spread volatility:

$$\sigma^2 = \sigma_E^2 + \sigma_G^2 - 2\rho\sigma_E\sigma_G$$

Margrabe prices the ex-ante exchange option on the trading-hub legs (its replication argument needs tradable instruments); the $V_\bullet$ machinery above is the ex-post measurement layer on the fuel-region series. The two meet in the data: a CAISO fuel region prices as its hub's daily index plus tariff adders (transport, plus carbon for the GHG variants), so the adders shift levels but not returns, and σ_E , σ_G , ρ estimated from the fuel-region series are hub-level inputs. **FRPGE1** / **FRSCE1** (non-GHG, backbone transport) are the thinnest wrappers and serve as the hub proxies; non-GHG is deliberate - carbon is a separately hedgeable leg and does not belong inside the gas price. Economic names in the math; series IDs in data sections only.

Quanto Option. Imagine this as a protection contract. It does not just protect you from high prices; it activates when two conditions are met: extreme weather occurs and market prices spike. The pricing consequence:

$$E[XY] = E[X]E[Y] + \text{Cov}(X, Y)$$

Cost of carry / the Zero-Carry Assumption. No-arbitrage forward pricing says

$$F = S \cdot e^{(r + \text{storage} - \text{convenience yield})T}$$