LECG

RTO Organized Spot Markets
Recognizing the Baby
And the Bath Water

Prepared by John Chandley

for ELCON Workshop

"Going Forward in Organized Markets"

Tampa, Florida

February 6, 2007

What's the Problem?

From ELCON's report, "*Today's Organized Markets – A Step Toward Competition or an Exercise in Re-Regulation?"* (12-04-06)

- "Large industrial electricity consumers have believed for many years that cost-of-service regulation produces limited benefits for consumers. Under cost-of-service, utilities have an incentive to increase their capital expenditures but have little incentive to seek more efficient lower priced generation options or pursue innovative product offerings . . .

- "ELCON continues to believe that 'true" or "real" competition in wholesale and retail electricity markets has the potential to bring significant benefits to consumers and to the overall US economy. The potential benefits include competitive prices, innovative products, and a customer focus. **The problem is that, perhaps with the exception of (present day) ERCOT, the so-called organized markets lack the necessary structure to promote "true" competition at the wholesale level.** And states that have tried to implement competitive retail markets have often included structural flaws (price freezes, politically driven stranded cost awards to generators that don't reflect economic realities, etc.). Those flaws, coupled with poorly functioning wholesale markets, have not provided the necessary foundation for effective competition at the retail level."

What is the "necessary structure" that ELCON sees as missing?

RTOs Are Regional *Open* Power Pools

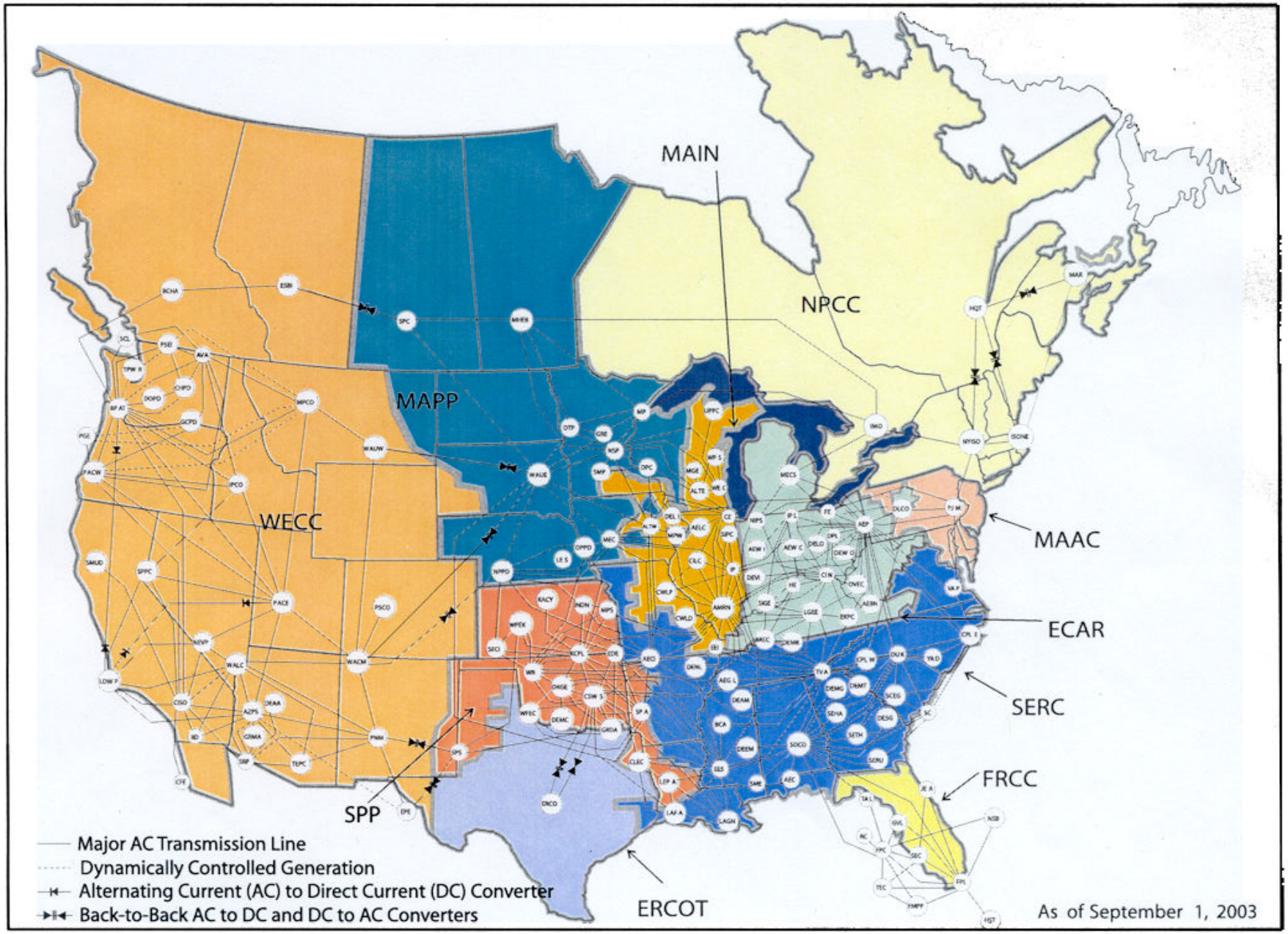
Power pools solve important reliability and network coordination problems that cannot be ignored. *They were inevitable*.

- PJM, ISO-New England and NY ISO started as power pools
- California ISO is a power pool for the three large private utilities
- ERCOT is a power pool for most of Texas' utilities
- MISO is a new power pool for many utilities in the Midwest.
- Large non-RTO utilities created *closed* pools (Southern, Entergy, AEP)

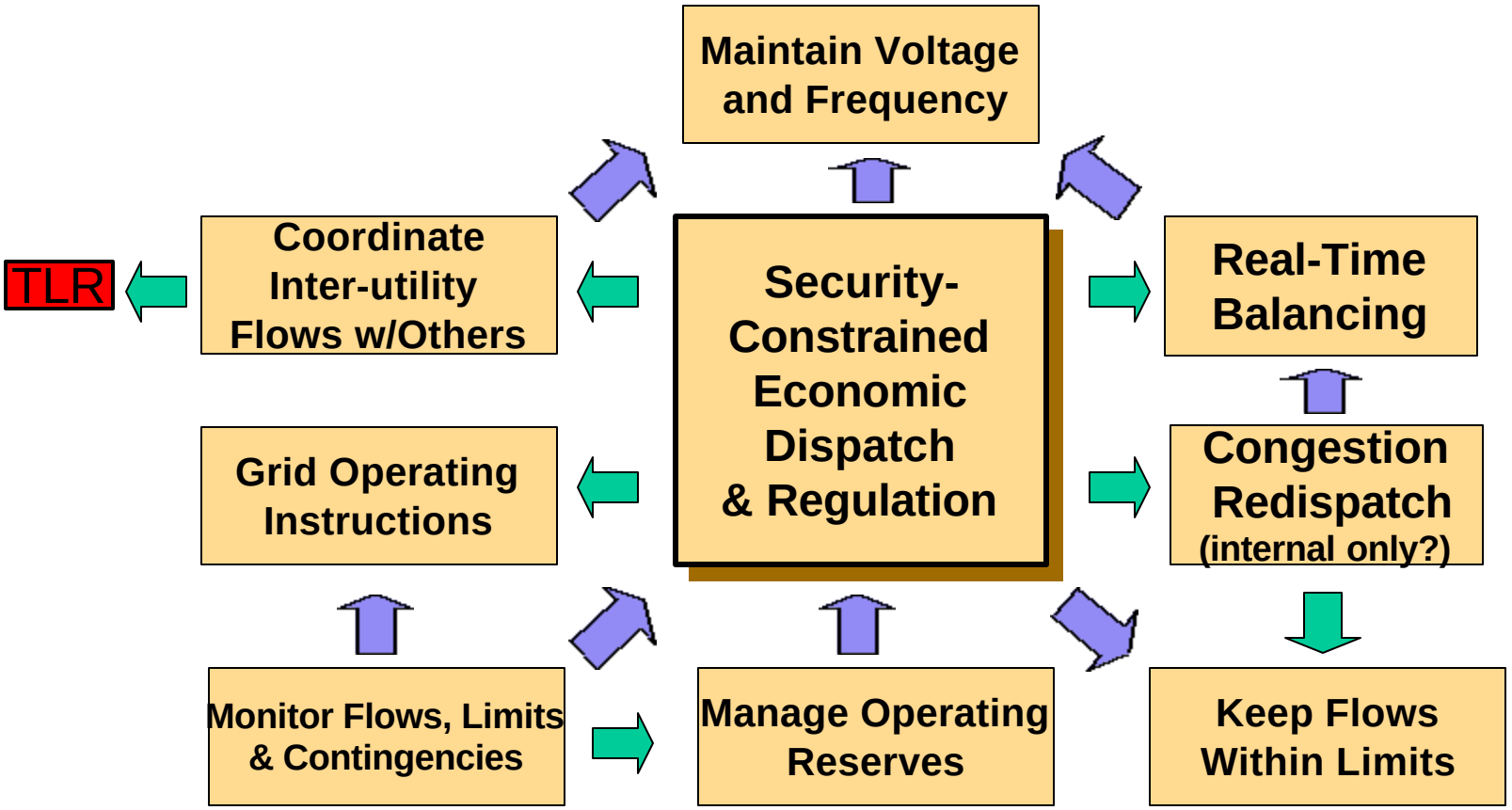
How these pools operate provides the *basic structure* for wholesale electricity markets. Let's examine the basic structure.

Figure 2.6. NERC Regions and Control Areas

Regional coordination => Pools



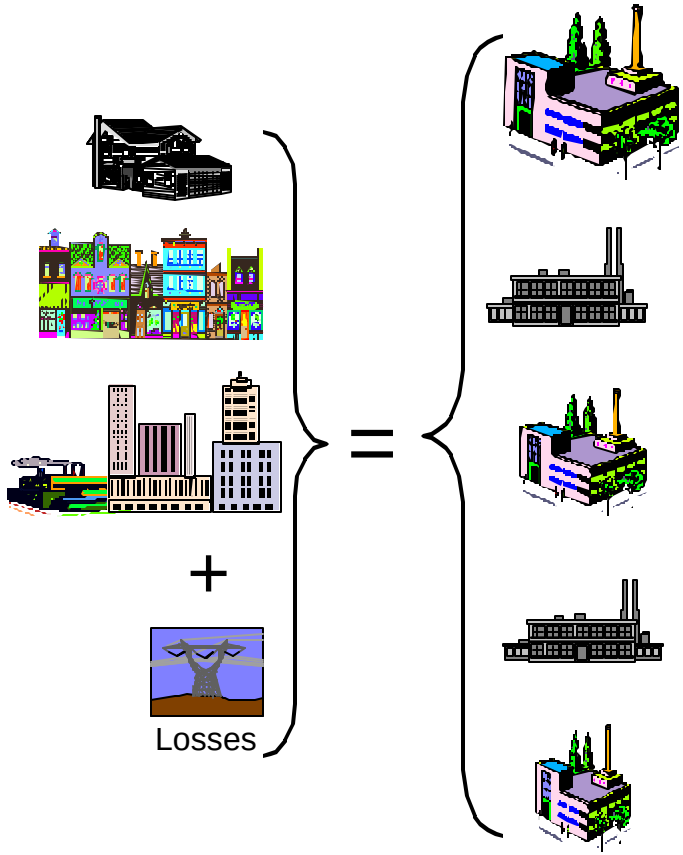
Essential Reliability Functions Center Around Each System Operator's Dispatch



A System Operator's Dispatch Matches Supply and Demand Every Second

- Dispatchers instruct generators how much to generate at each location in each dispatch interval (usually every 5 minutes).

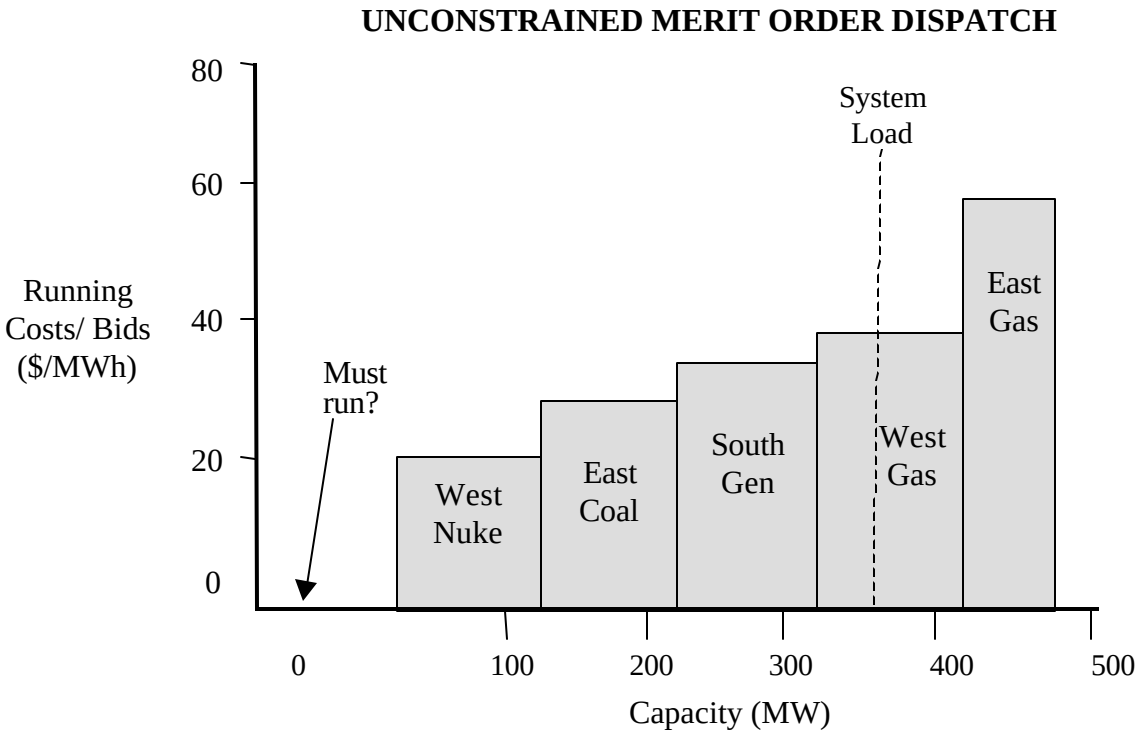
  - There's virtually no "storage" in electricity, so electricity must be generated as it is consumed.
  - Automated "regulation" fine tunes output in seconds to balance supply/demand at all times.
  - Energy dispatch keeps frequency at 60Hz
  - Reactive power dispatch keeps voltage stable
  - These and other actions keep the lights on

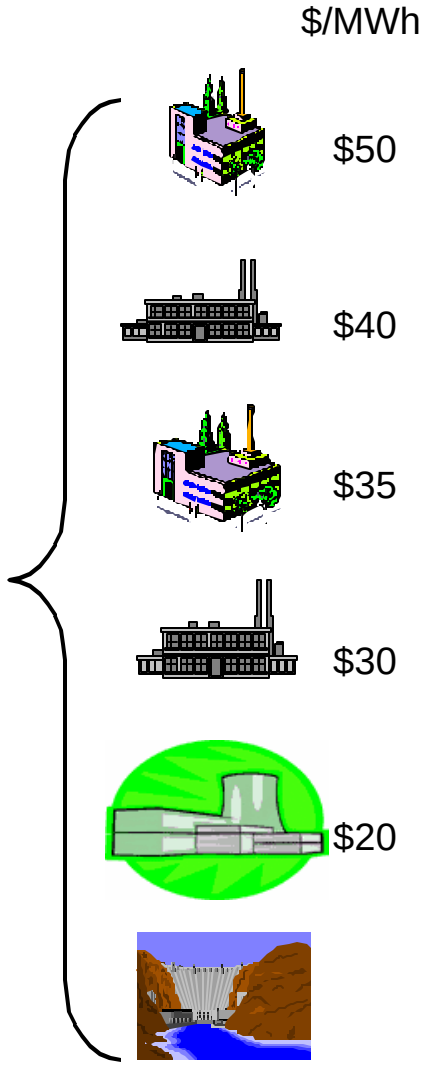


Total MW = Total MW

The System Operator's Dispatch Also Tries to Meet Demand At Lowest Cost

- Operators try to dispatch *economically*.



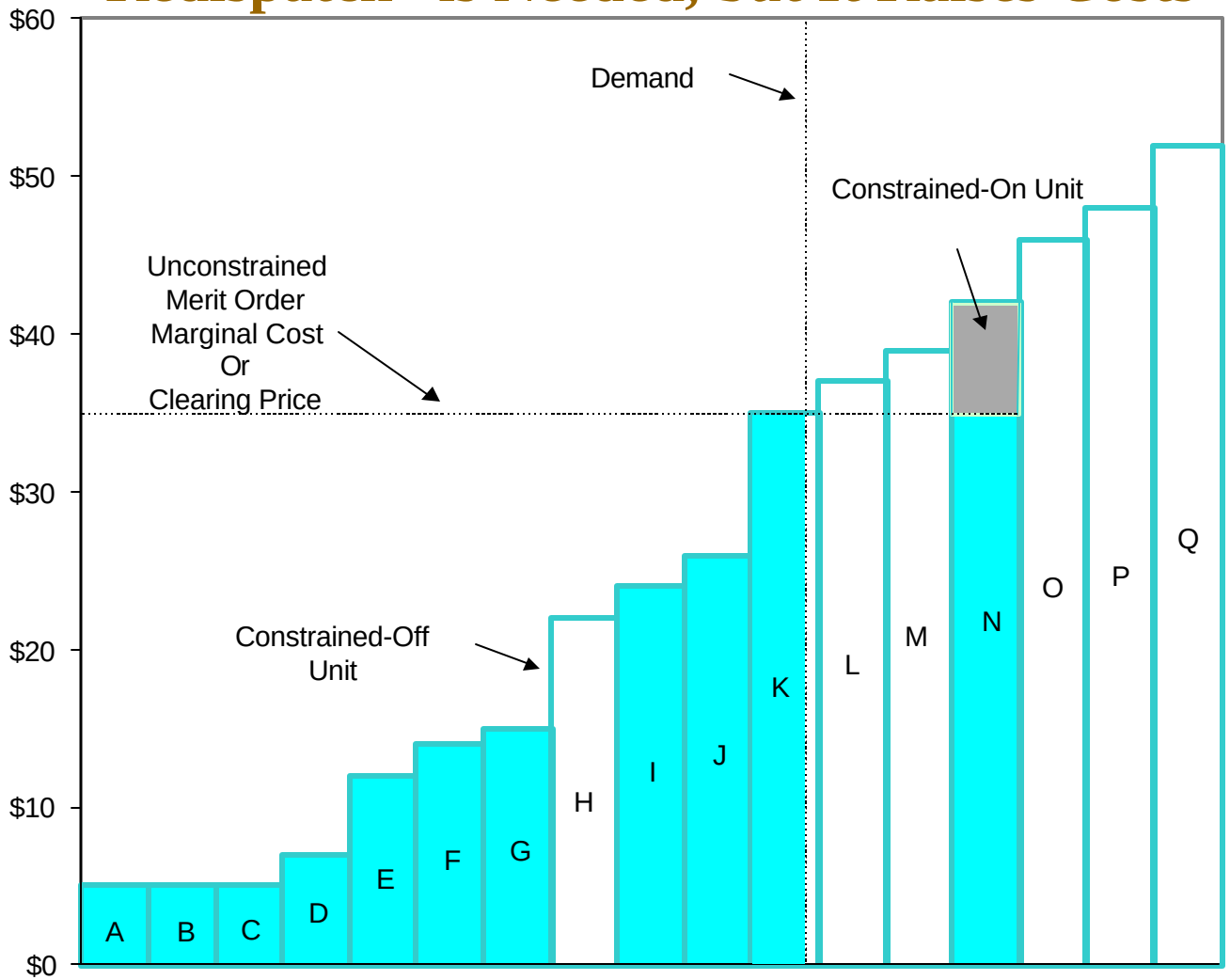


Security-Constrained Economic Dispatch:

Congestion Requires Operators to Dispatch Out of Merit Order to Avoid Overloading Transmission.

LECG

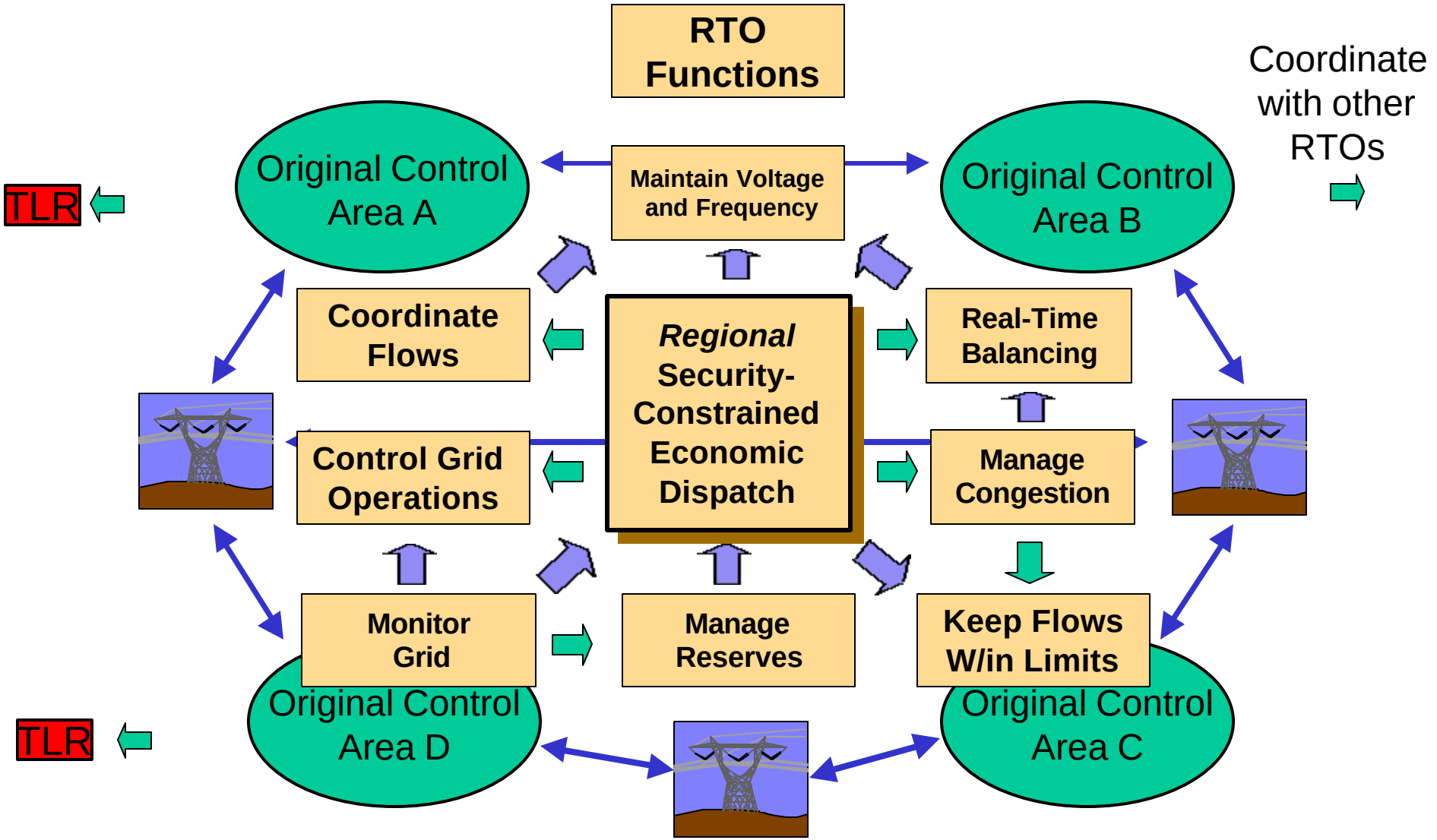
"Redispatch" is Needed, but It Raises Costs



Least cost Redispatch

Units H & N are the most *cost-effective* to constrain off and on to relieve the constraint

An RTO Uses a Regional Dispatch To Replace Local Control Area Dispatches



A Regional Dispatch Creates a Pool

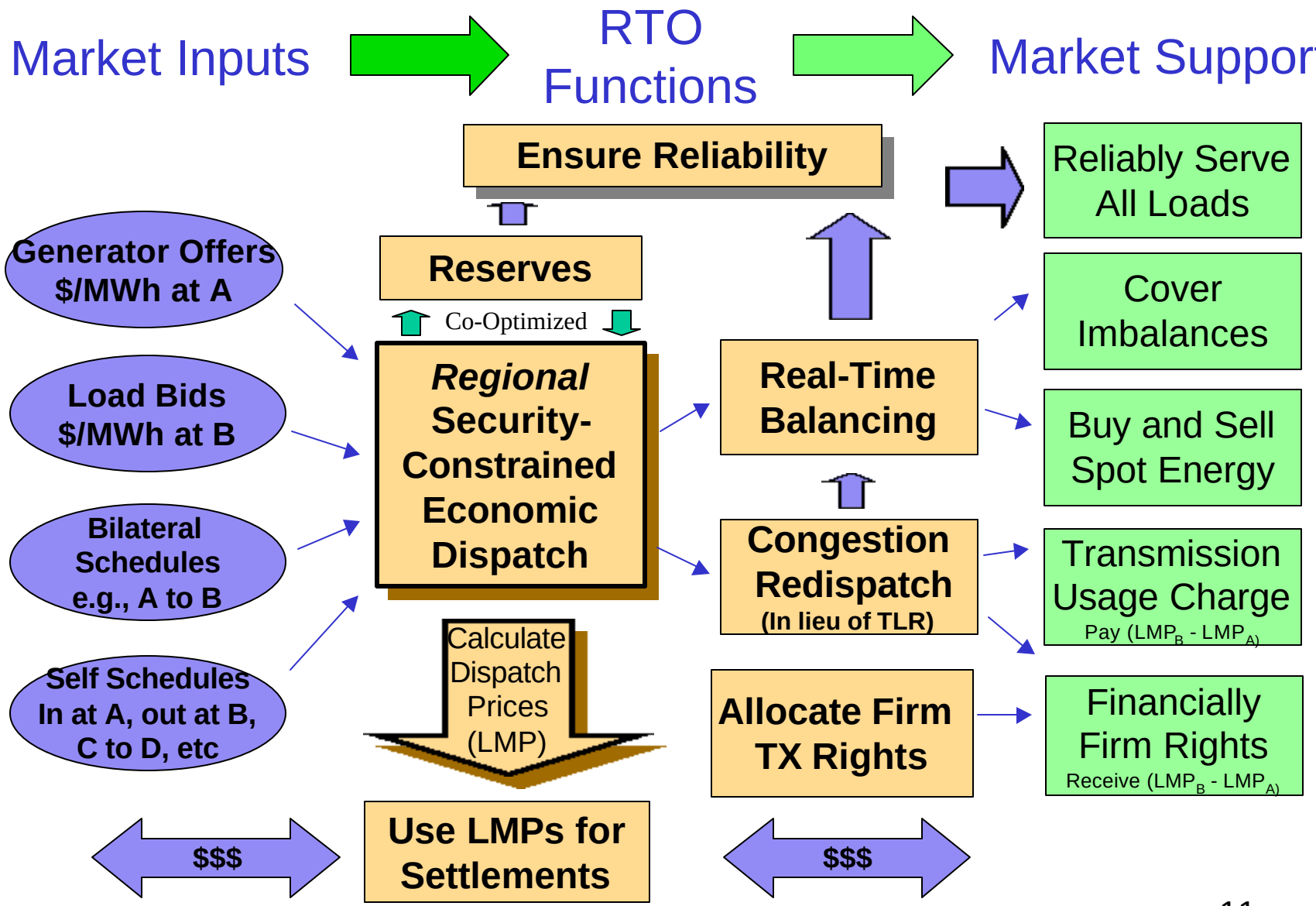
RTO Open Power Pools Create Spot Markets

Once you create a dispatch power pool, and open it to everyone, *it automatically creates a spot market*.

- Quantity/price offers/bids determine who gets dispatched.
- ISO-Pool has to pay generators/sellers for the energy they inject.
- ISO-Pool has to charge loads/buyers for energy they withdraw.
- ISO-pool has to charge/pay everyone for their imbalances.
- ISO must charge/pay for redispatch to relieve congestion.

The spot market/prices flow directly from the dispatch.

RTOs with Standard Core Features Enhance Grid Reliability – And Create Spot Markets



ELCON's Necessary Elements

- Prices must be established through an interaction of supply and demand.
- New capacity must be "incented" through market forces – not administrative re-regulation.
- Market entry and exit should be determined by market forces.
- Consumers must be able to hedge future prices with long-term bilateral contracts.
- There must be adequate transmission infrastructure.
- Market power must be mitigated, but . . .
- Wholesale price caps and bid mitigation may be relaxed.

Price Determination = Market Clearing

ELCON's paper says two different things:

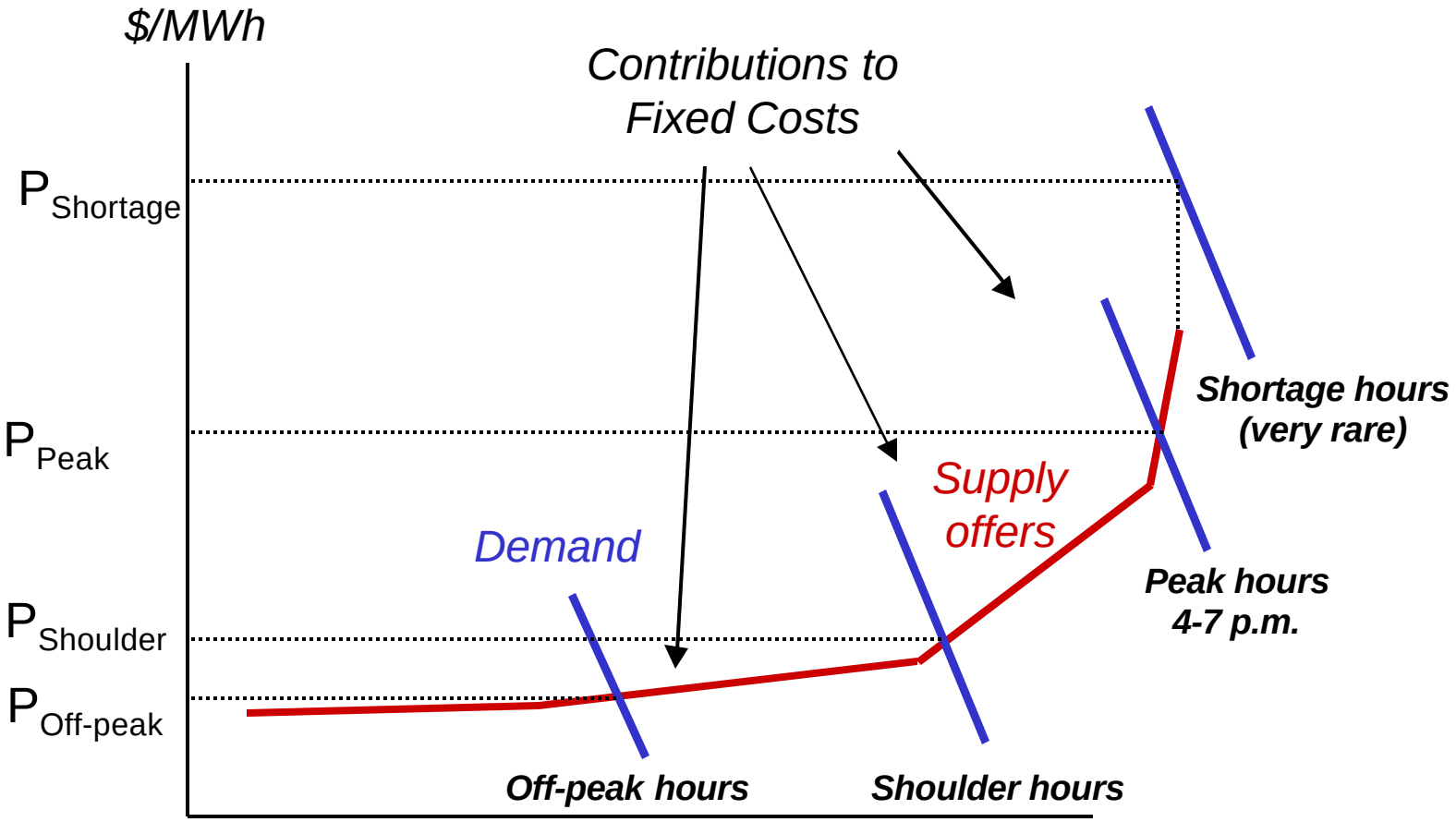
- "Prices must be established through an interaction of supply and demand."

- "A major problem is that today's organized markets have implemented a bidding system relying on a single-price auction that applies to all generators. In general, this allows generators to receive a price that is determined by the highest cost generation unit (usually gas in today's markets) operating at the time (the 'generator on the margin'). This eliminates the benefits of fuel diversity since the low costs of efficient baseload coal and nuclear facilities are never passed onto consumers, but rather produce windfall profits for the owners of generation."

These statements are not consistent.

The real problems with price formation are elsewhere.

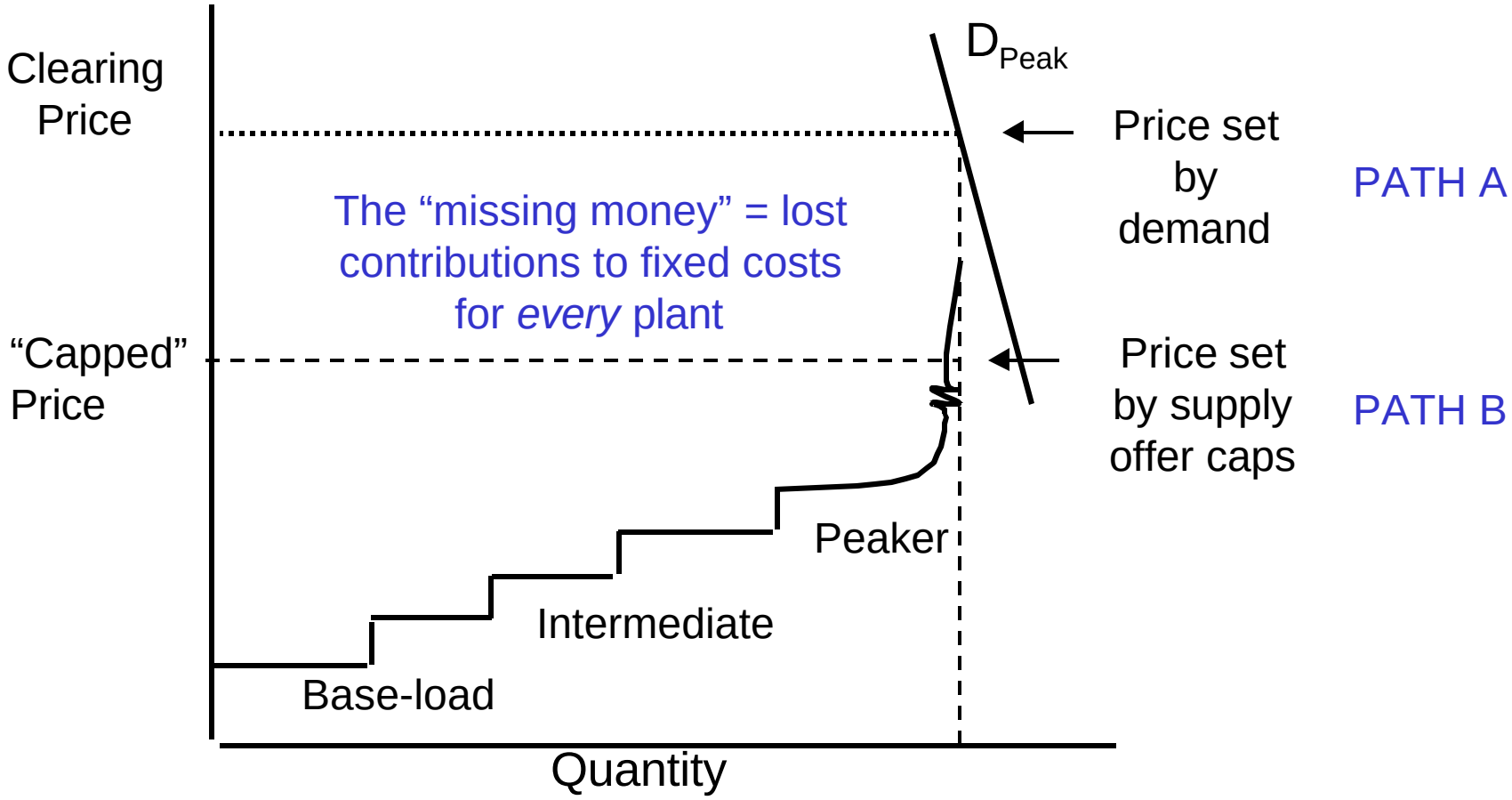
Generators Depend on the Highest-Price Hours To Recover Most of Their Fixed Costs



Low-price hours barely cover operating costs.

All clearing prices are from "interaction of supply/demand."

A "Missing Money" Problem Arises In Shortages If There Is No Shortage Pricing



The "missing money" problem undermines both long-run investment ***and*** short-run reliable dispatch

Path A or Path B?

A key design choice depends on how *energy and operating reserve prices* are set. What happens when the ISO dispatch runs short of plants to supply the total demand for energy and operating reserves?

- Path A: Allow scarcity prices to clear the markets based on *demand for both energy and operating reserves*.
- Path B: Cap the energy prices, but recover the "missing revenues" from capacity payments and other mechanisms.

Some ISOs (except MISO/ERCOT) have chosen Path B

All the First ICAP Markets Failed! Why?

They ignored the underlying incentive problems . . .

- Capped energy prices plus ICAP payments => . . .
  - Poor incentives to gencos to be available when needed in RT.
  - Poor incentives to provide the right operational features.
  - Poor incentives for demand-side response.

Uniform ICAP payments ignored locational differences

- No incentive to build plants at the right locations.
- So all ICAP mechanisms are moving to multi-zonal "LICAP."

And fixed ICAP demand created other problems . . .

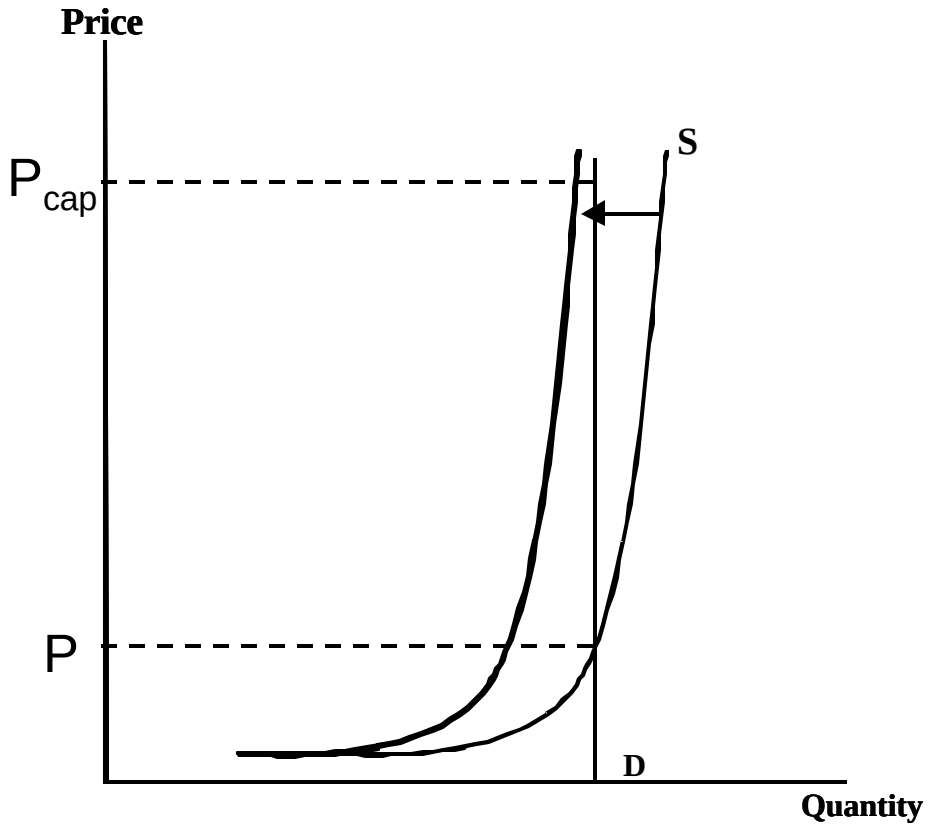
- *Volatility* -- inevitable with fixed capacity requirements.

- *Market power* -- it's easy to exercise in a capacity market.

Market Power Incentives in ICAP Markets

The combination of a vertical demand curve and a near-vertical supply curve presents strong incentives to exercise market power.

- Any withheld supply shifts the Supply curve to the left.
- A slight shift of the supply curve to the left causes a large increase in price.
- This can easily offset the revenues foregone as a result of withholding.



Two Approaches To Lessen Market Power

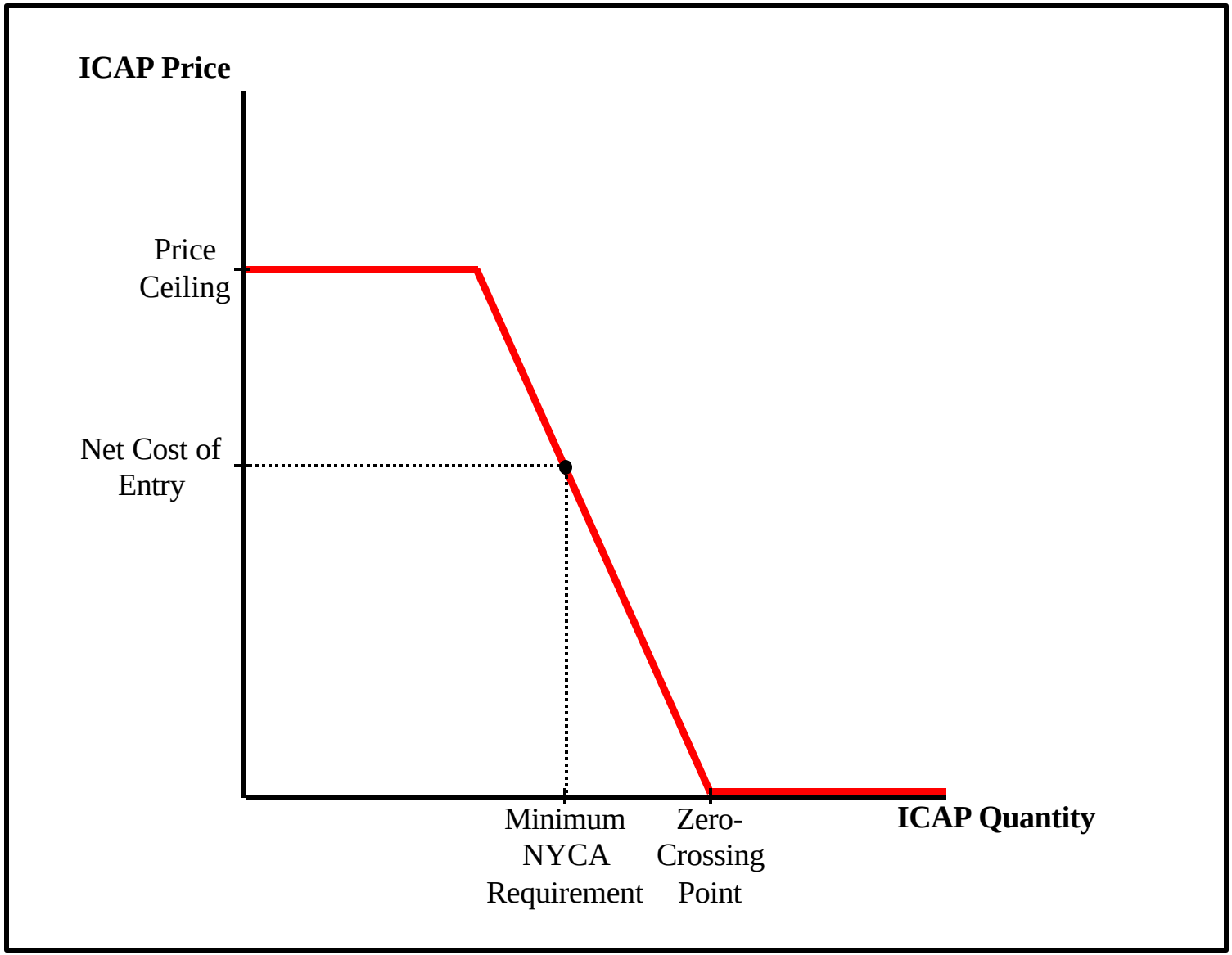
Use a downward-sloping curve for ICAP demand.

- A sloping demand curve reduces the incentives for market power.
- Recognizes that capacity beyond the “target” has value.
- Recognizes that capacity is more valuable when short.

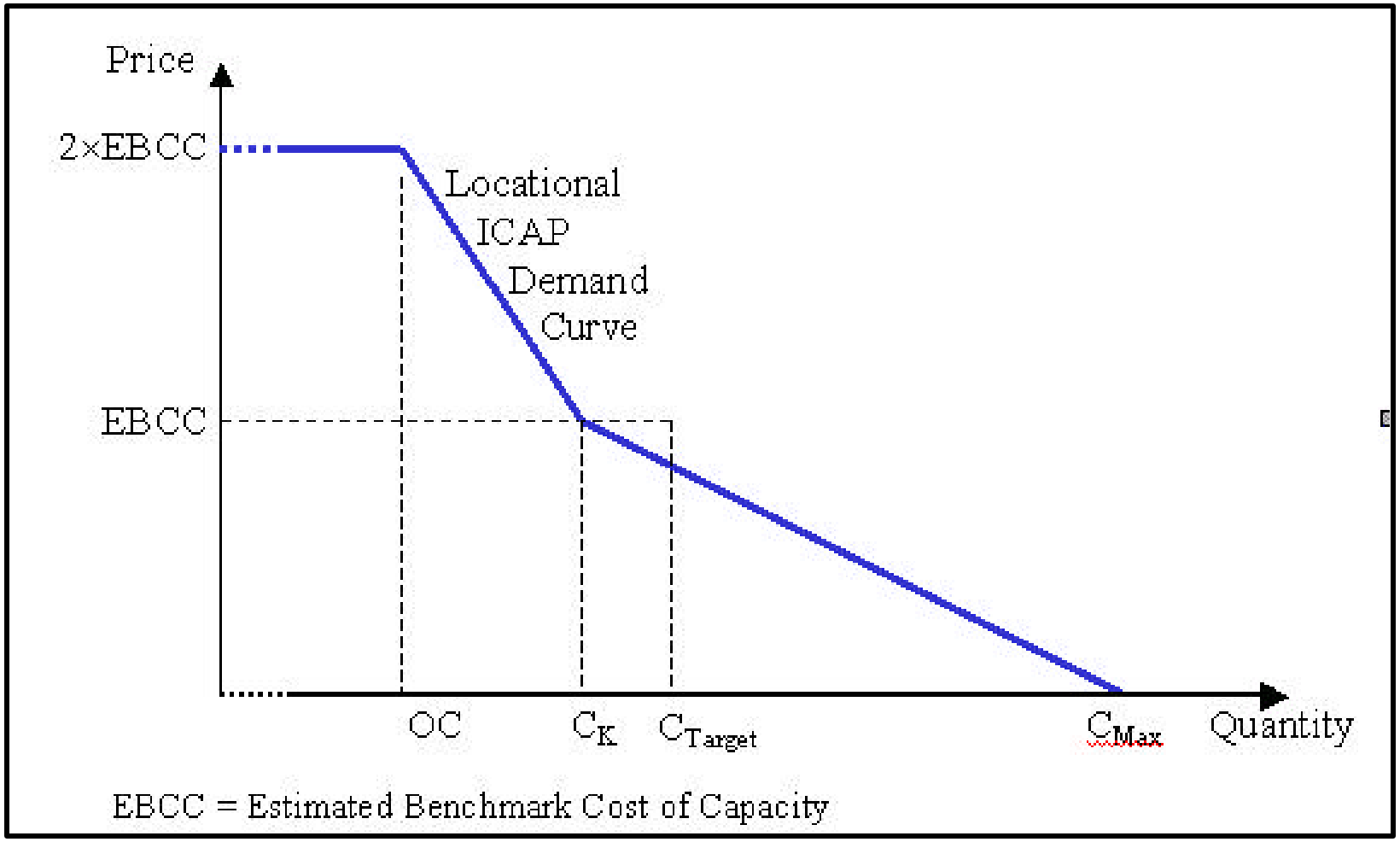
Use a forward auction for ICAP (extend the supply curve).

- Define “product” to be delivered 3-4 years from now.
- New entrants can compete with existing resources.
- May facilitate competition with transmission?

NY ISO ICAP Demand Curve



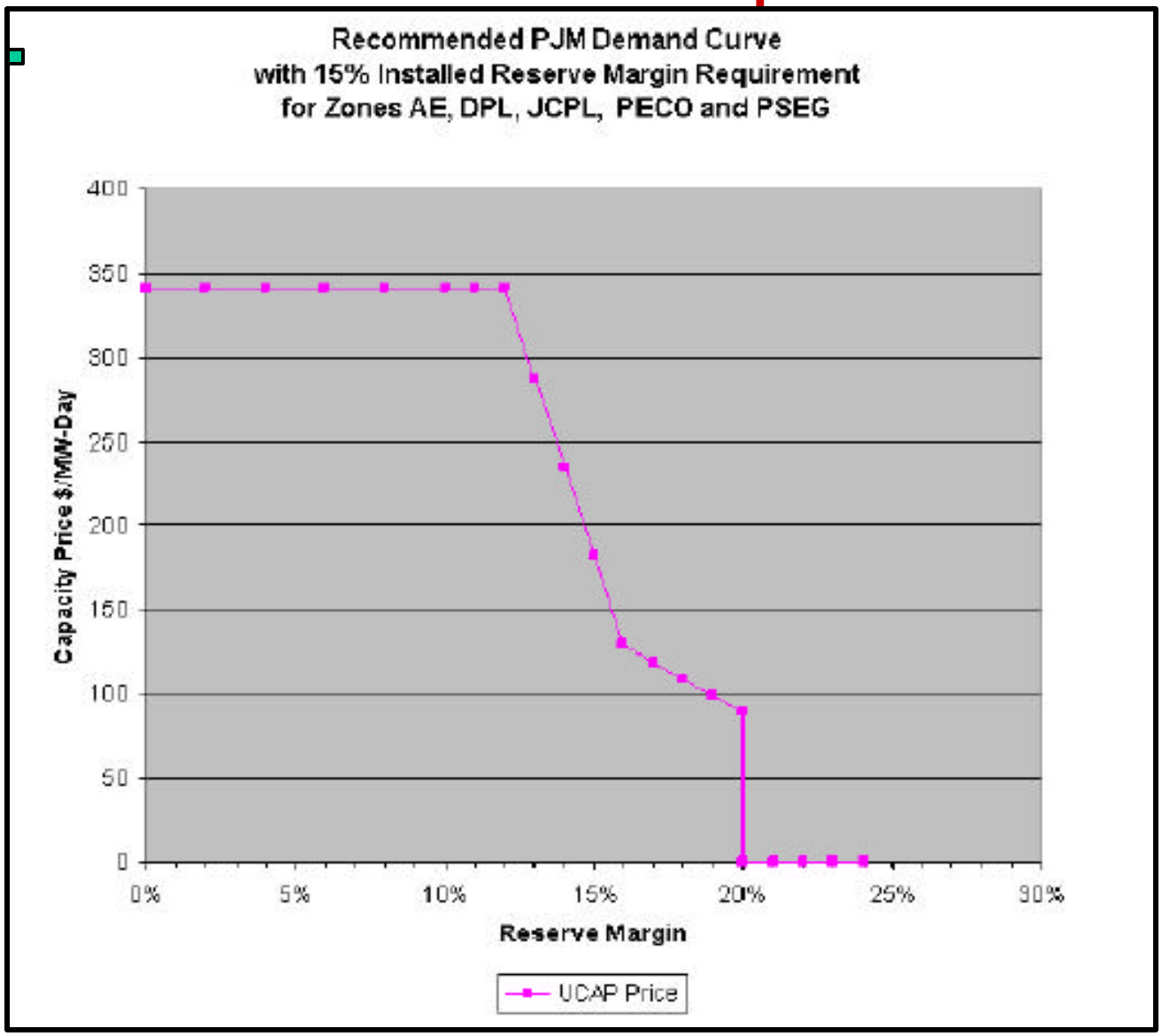
Proposed ISO-NE Demand Curve (Not included in Settlement)



OC = Objective capability = RM Requirement

PJM Initial Demand Curve for Variable Resource Requirement

LECG



**Settlement Curve Reverses "Kink"*

Market Power Solution II: Change the Supply Curve

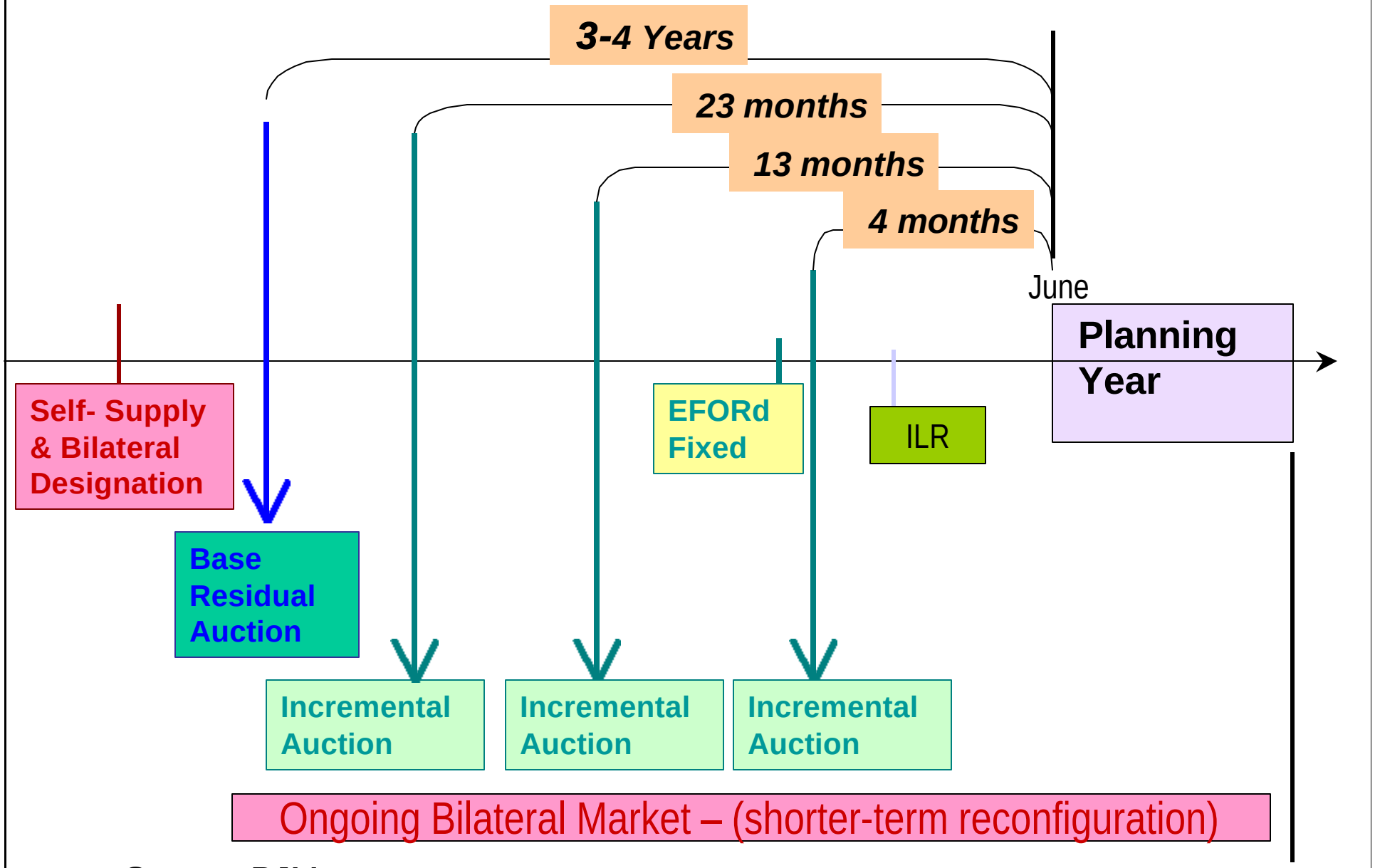
A joint PJM-NY-NE study developed the concept of holding the ICAP auctions 3-4 years in advance.

- If the ICAP "product" does not have to be delivered for 3-4 years, then new entrants could compete against existing plants.

New entry/competition would limit the ability of existing plants to exercise market power, such as by withholding capacity from monthly auctions.

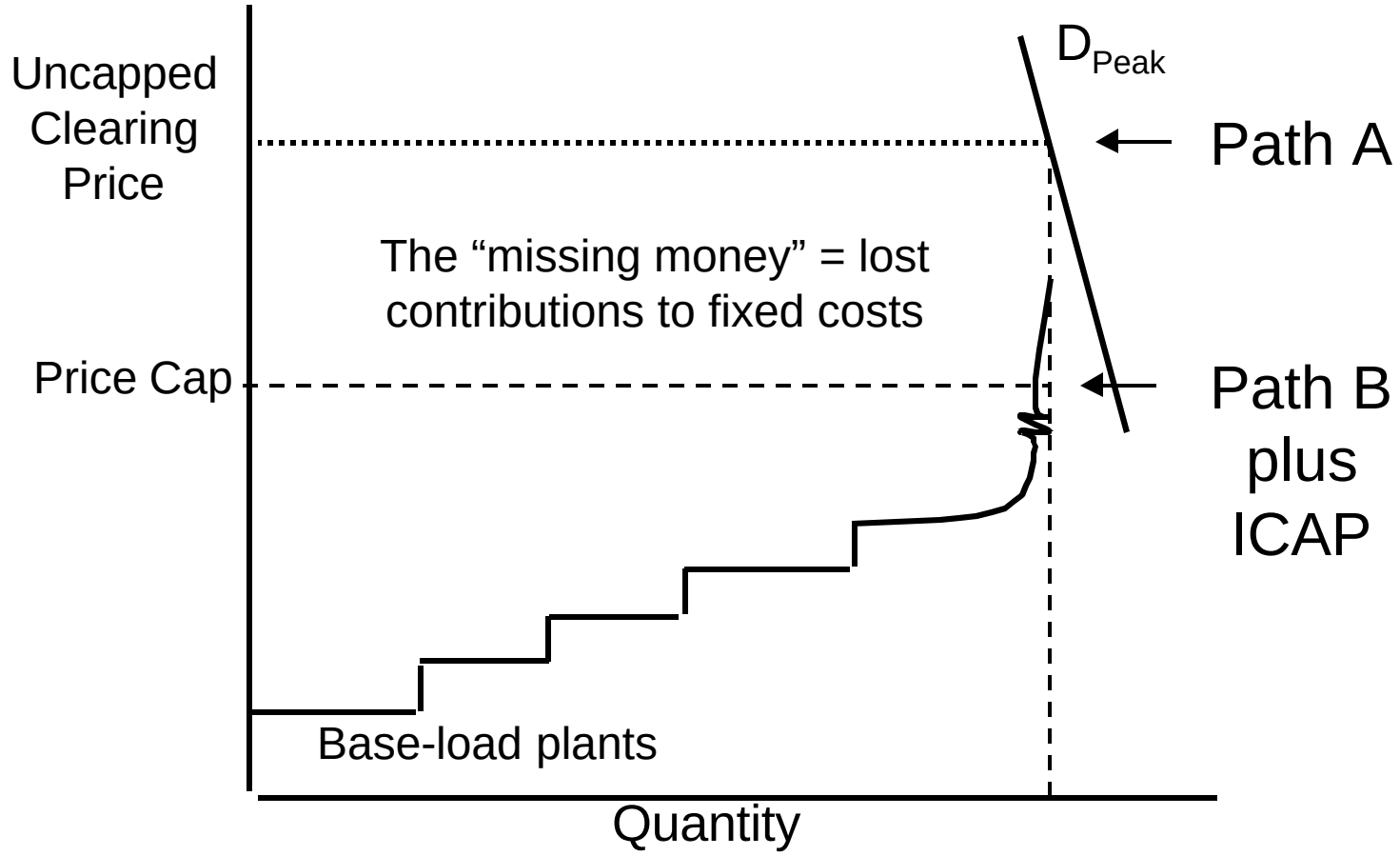
PJM's Reliability Pricing Mechanism and ISO-NE's LICAP Settlement use this forward auction approach.

Proposed Timing of RPM Auctions



Source: PJM **Structure Provides Contract Equivalent*

Revisit the Strategic Decision: Avoid Missing Money With Shortage Pricing



Path A = Energy and Op. Reserve Market with Shortage Prices
Path B = Capped Energy Market with (L)ICAP, curves, RPM

What the Midwest ISO Is Proposing

Apply shortage-cost pricing in the spot markets (PATH A).

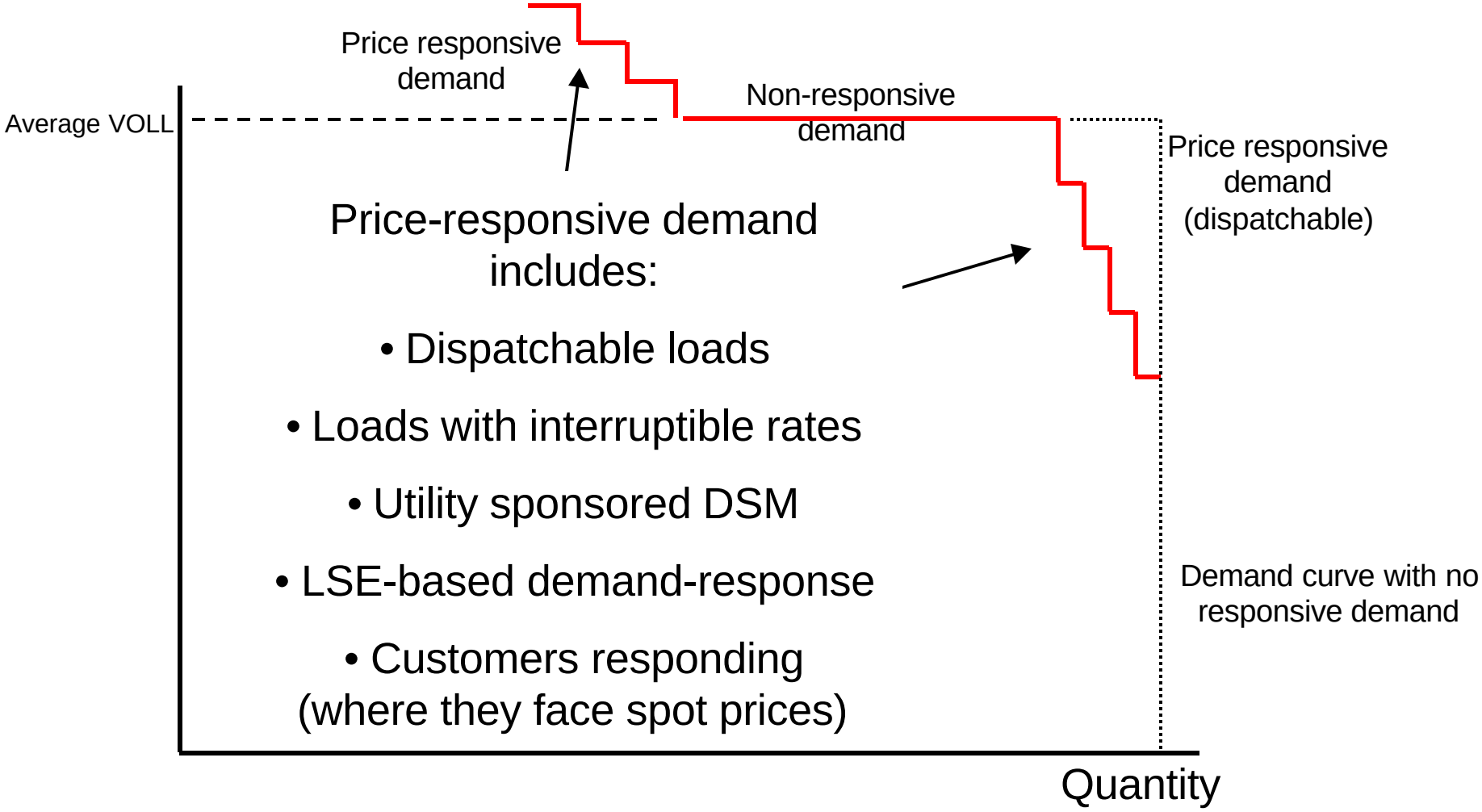
- To improve real-time reliability, MISO plans to fix the way it prices energy and operating reserves in the real-time (and day-ahead) spot markets.

In "shortage-cost pricing" (aka "scarcity pricing") . . .

- When supplies are short, prices can be set by demand – the willingness of consumers to pay -- not merely by the offers/bids of generators.

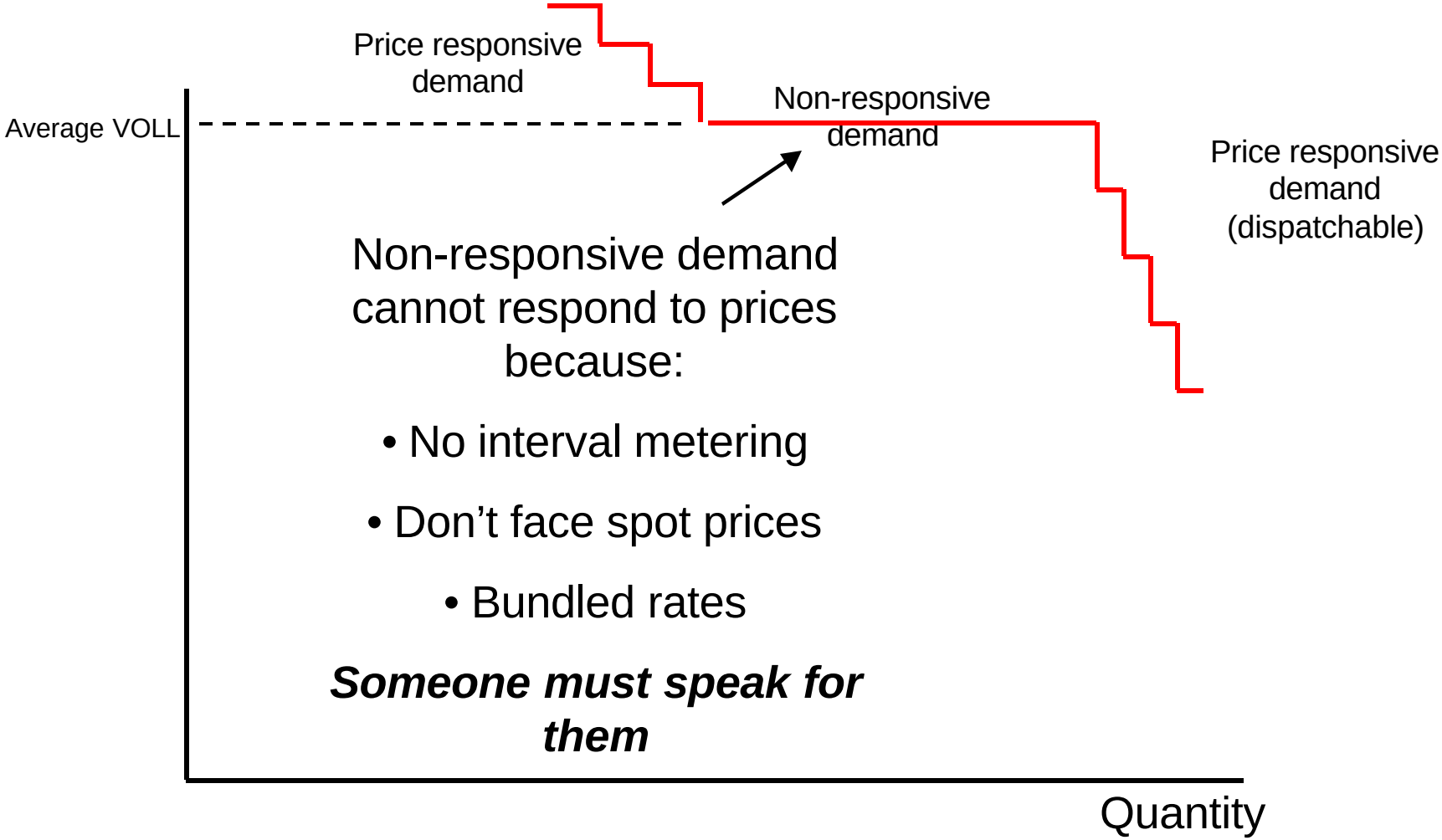
- And energy prices are affected by the level of operating reserves. If the ISO falls short of operating reserves, jeopardizing reliable dispatch, energy prices would rise to reflect that shortage.

Energy Demand Curve Components



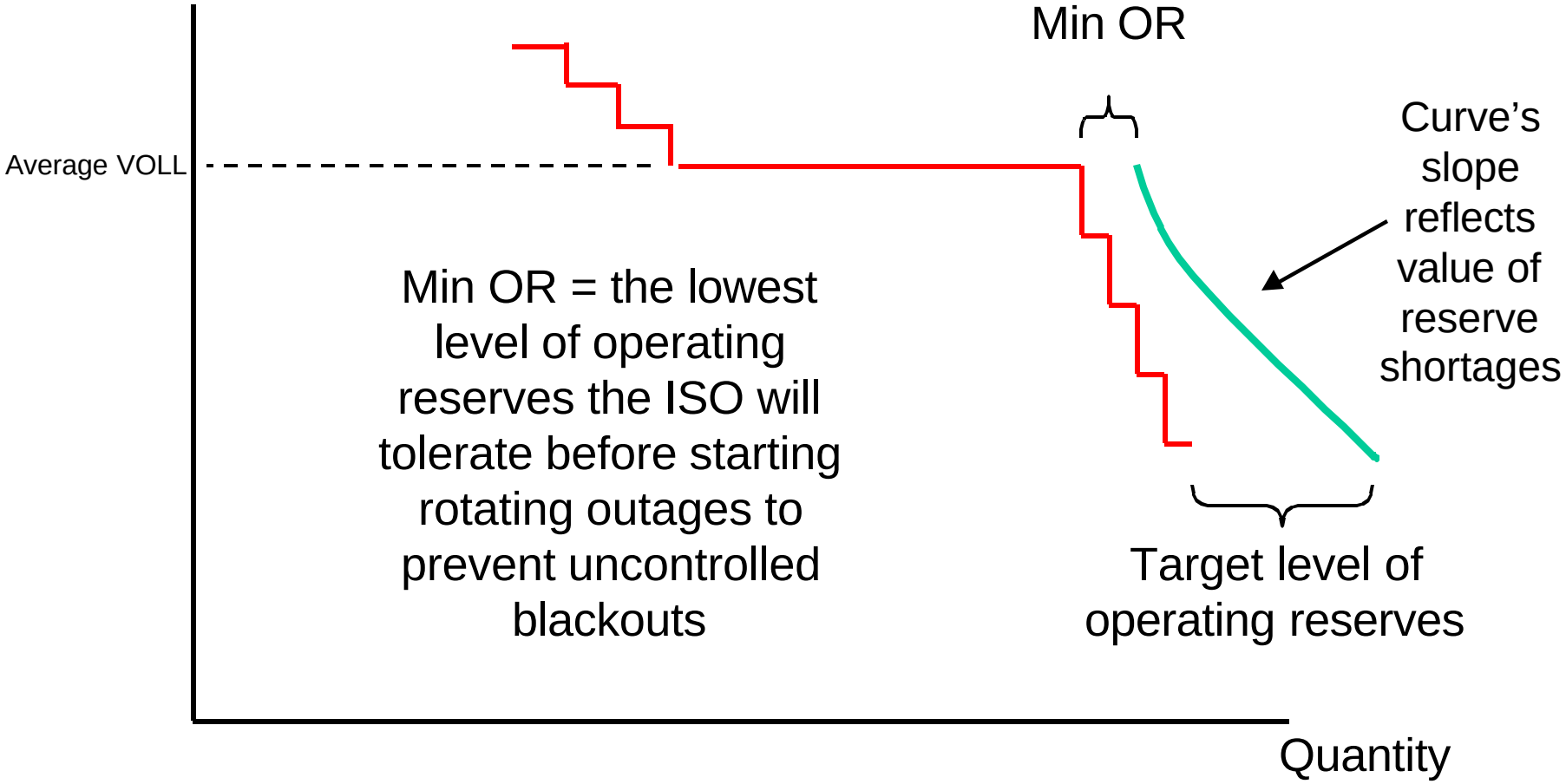
Price responsive loads (or their states, utilities, LSEs) decide what price they are willing to pay.

Some Portion of Demand Cannot Easily Respond to Price



Some entity (states? ISO?) estimates what price these customers would be willing to pay to avoid curtailment.

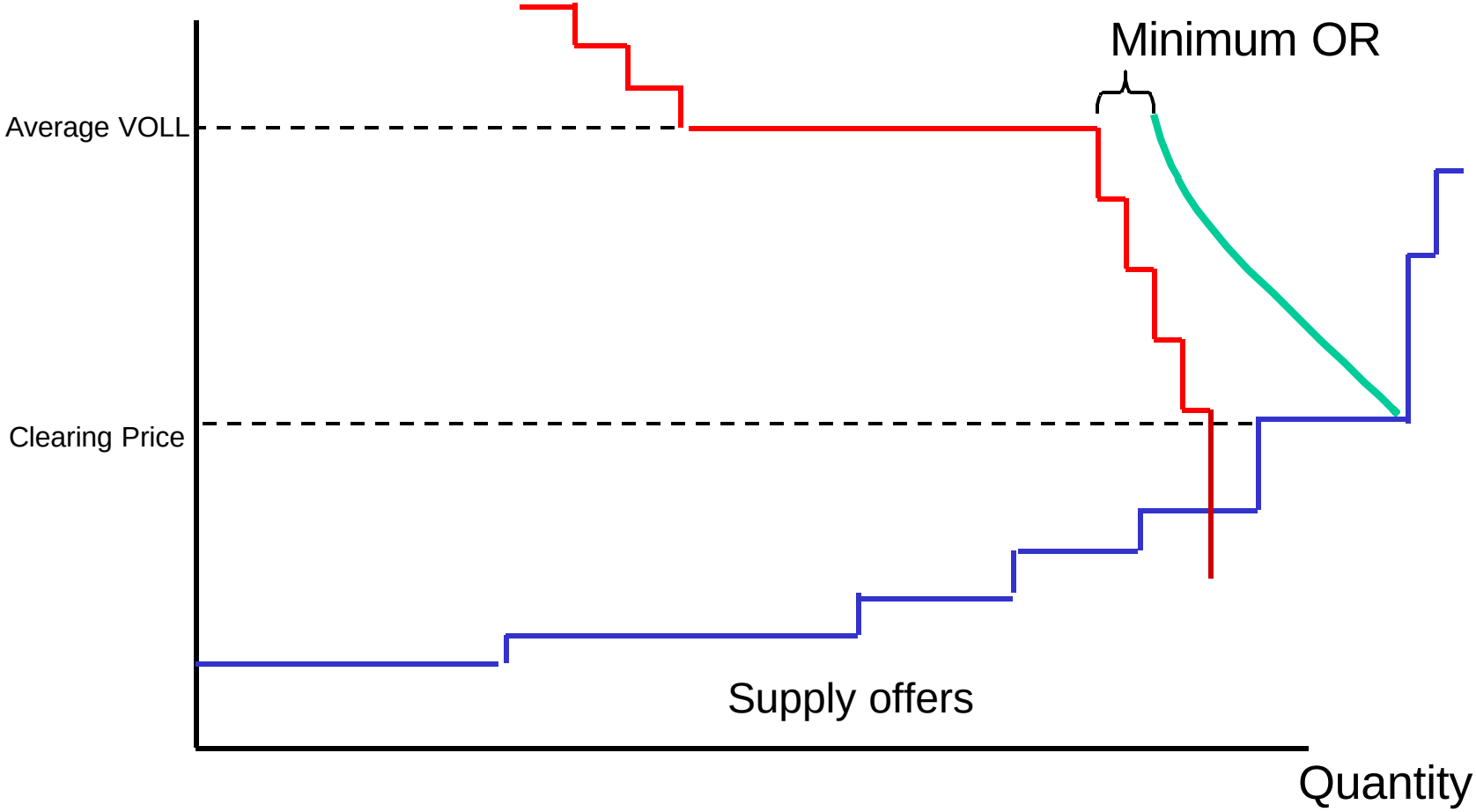
Composite Demand Curve for OR



Every ISO/SO has a curve. It's not explicit or transparent.
Proposal: make it transparent and allow it to affect prices.

Prices With Ample Supplies
These Apply Most of the Time

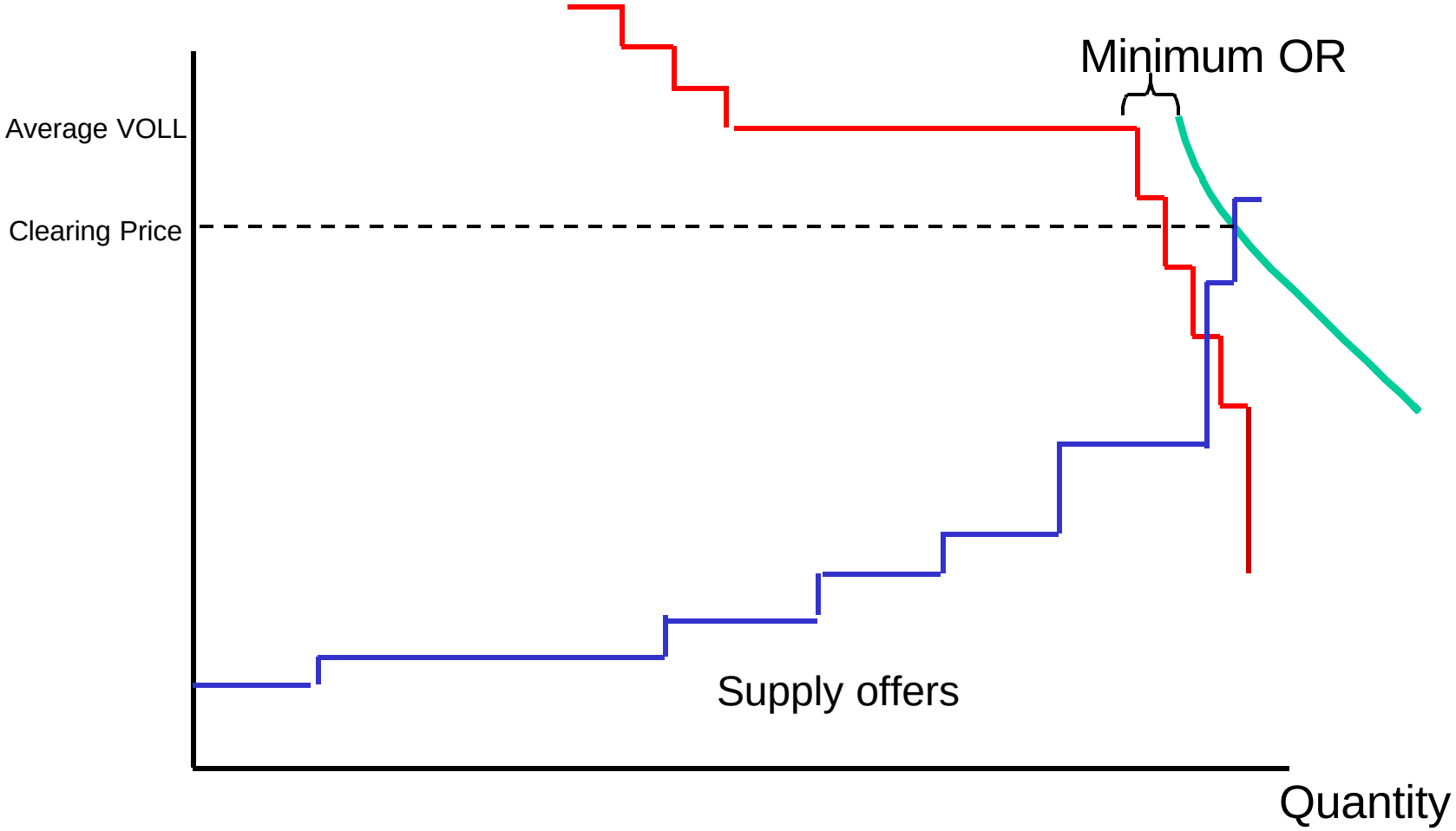
LECG



There is ample supply to meet demand for energy and operating reserves at low prices.

Prices With Tighter Supplies These Apply On Rare Occasions

LECG



With ISO unable to meet OR target at lower prices, prices rise. Some price-responsive load may reduce demand.

Market Power Mitigation Still Applies

In an energy plus operating reserve market with scarcity pricing, market power mitigation would still apply.

- Offer caps would still apply to prevent price gouging bids, just as they do today.
  - Unit-specific conduct and impact tests still apply
  - $1000/MWh overall cap still applies.

- Must offer rules would prevent physical withholding, just as they do today.

Shortage cost pricing does NOT mean removal of market power mitigation.

Reminders: Limited Spot Price Exposure

Most retail customers have little or no exposure to volatile spot prices. Only those who choose to rely on spot prices are exposed.

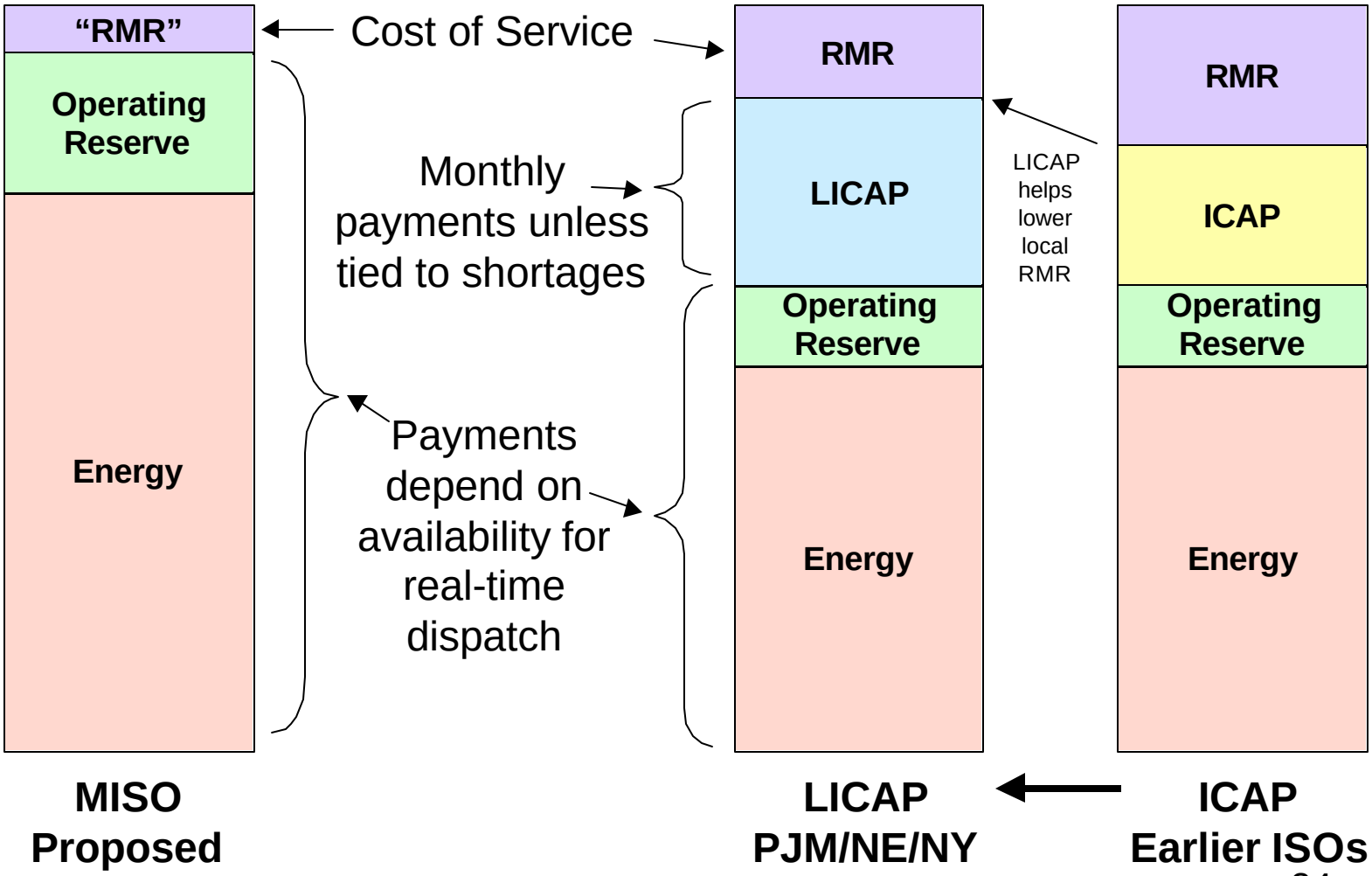
In regulated states (no retail choice) . . .

- If a utility covers full load requirement, *retail customers are not exposed*.
- If utility purchases any energy from ISO spot market, it pays spot prices only for that amount; *but retail customers are hedged by fixed rates*.

In states with retail choice . . .

- Largest customers might face hourly spot prices, but they can be hedged through contracts, own generation (self supply), demand response, to the degree they choose. Several states already doing this.
- For smaller customers: Regulators would ensure utility or competitive LSEs hedge default customers with longer-term contracts. (e.g., New Jersey/Illinois default supply model)

Different Models Determine Where Generators Get Their $$$, But the Total Revenue Required Is Similar



Different Models Cost About the Same

First, the total revenue requirements are about the same for all three approaches, for a given RA target.

- It means aggregate retail rates should be about the same.

Second, for the same level of revenues, all three approaches achieve about the same level of RA.

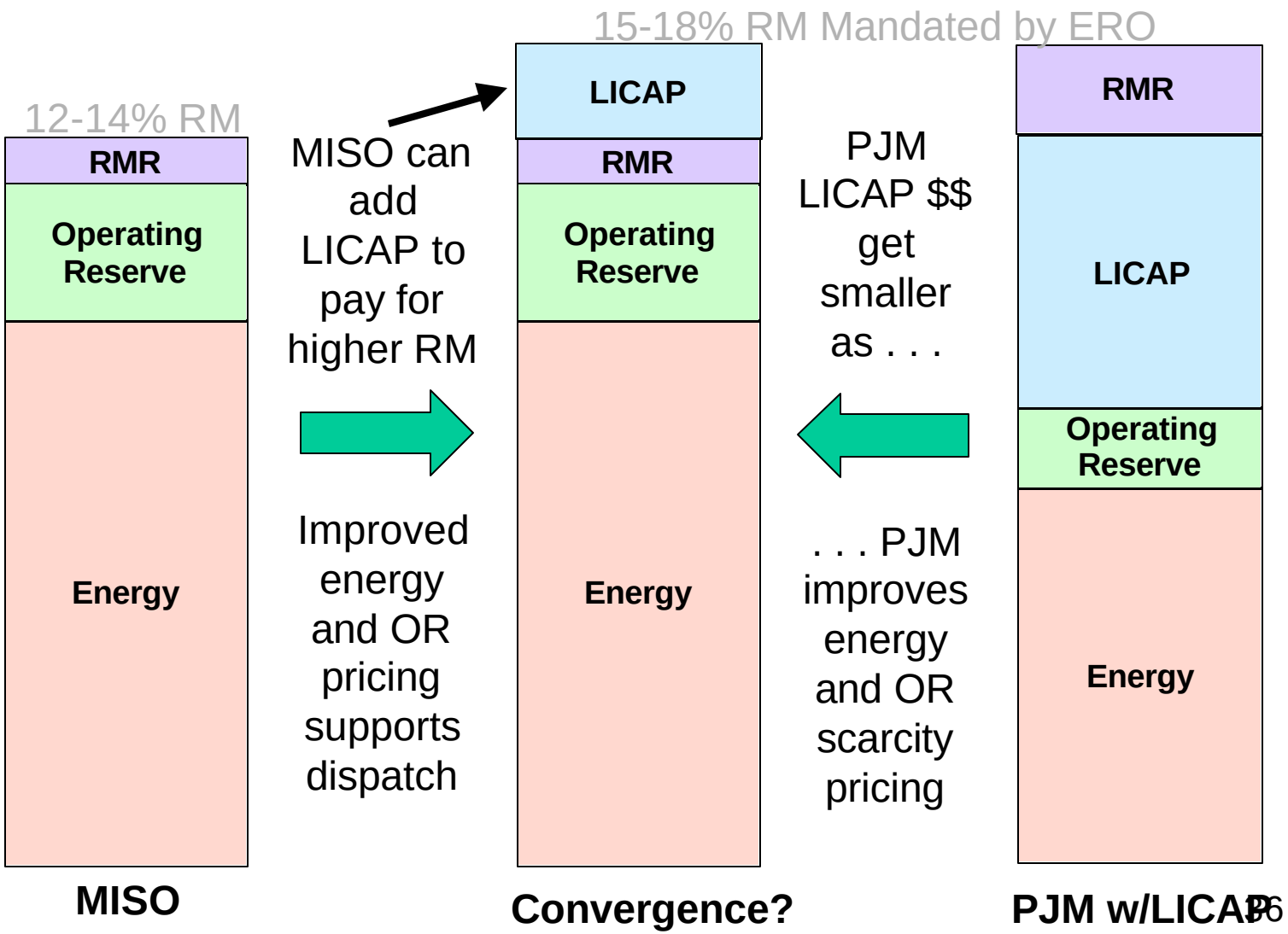
- That's because, if total revenues are about the same, the total level of investment is also about the same.

- (The mix of investments may be a little different, because investment incentives depend partly on where the revenues come from – i.e., what are we rewarding? Reliable energy producers? Or just countable capacity?)

Bottom line: different models change *the source* of where generators get their revenues, *but not the total*.

If All ISOs Must Meet Higher Reserves Then MISO and PJM May Converge

Setting the RM Drives Total Costs!!



Some Final Observations

The Basic Structure of RTOs is an Open Regional Pool.

- Accommodates contracts, competition, retail choice, DG, DR,. . .
- Only structure that provides non-discriminatory open access.
- Non-RTO utilities don't allow open dispatch, redispatch, FTRs, etc.

Spot Price formation is classic economics = market-clearing

- Clearing prices are based on marginal costs, but that is not the same as "the highest cost unit" running. Prices can be higher/lower than any bid.
- LMP = marginal cost of the dispatch => no subsidies, no discrimination

Original capacity payments were clearly problematic, but…

- 1-3 years of capacity payments reduces need for ICAP contracts.
- Reforms are changing ICAP to look more like "energy-only" markets.

Other troublesome price issues are in *the structure of retail rates* and controlled by states, not by RTOs.

- *Retail access to the RTO spot price* is the key to solving these issues.

Extra Slides

Some Incorrect Statements

- "Nodal pricing actually creates a greater disincentive to build,[transmission] as the congestion often protects a high-cost generator from low-cost competition and thus provides sustained profits."

- "The larger the number of "nodes" in a LMP environment, the greater the opportunities for the exercise of local market power because there simply are fewer suppliers at each node."

Both statements are incorrect. The problem is congestion, not nodal pricing. Insufficient transmission leads to congestion, and congestion prevents the competition from external suppliers. That would be true even if we used zonal or uniform pricing.

Nodal pricing (LMP) merely reveals the marginal cost of redispatch; it does not create or increase congestion nor create or increase market power. It *reveals congestion costs* so that market participants can do something about those costs.

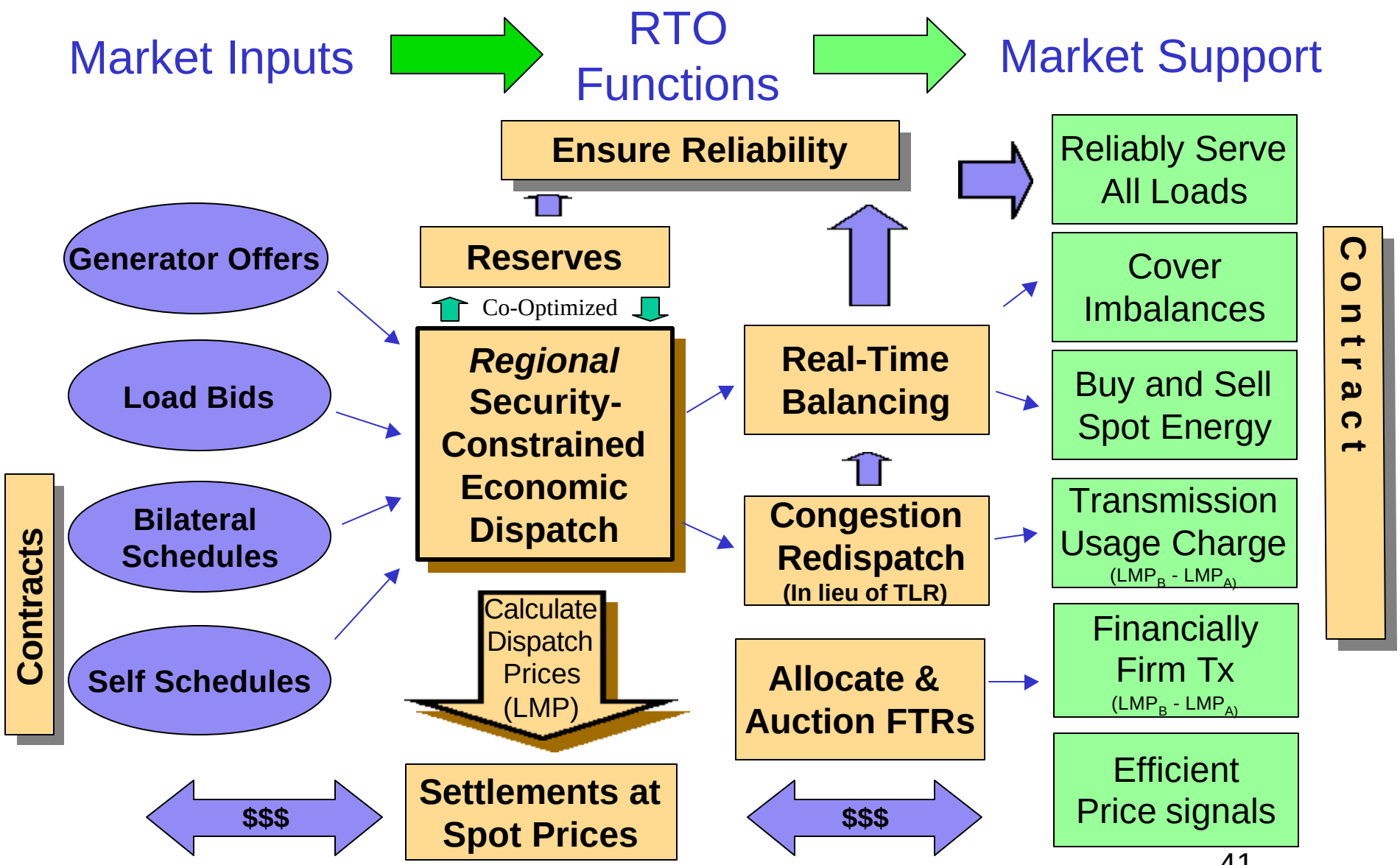
Are Loads Forced to Pay "the Higher of Cost-of-Service or Market"?

Not exactly, and these terms do not mean what you think they mean.

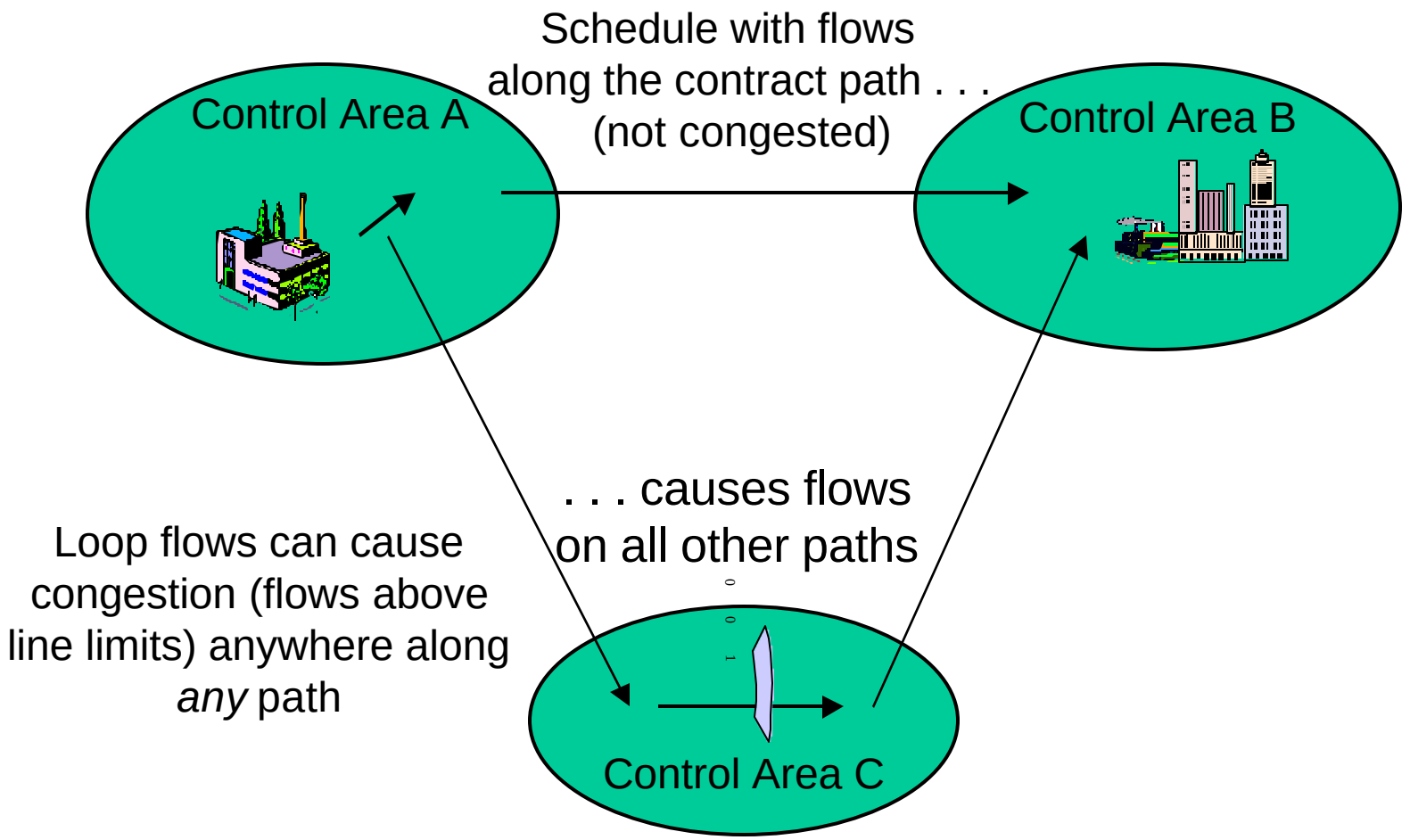
- ICAP payments are defined as "net" of profits from energy and operating reserve markets. Total revenues from all markets – energy, OR, ICAP – are thus capped at cost-of-service, as defined by the ISO.
- This ensures that ICAP payments recover only the "missing money" but no more.
- However, traditional notion of "cost-of-service" includes cost overruns for *overbuilding and overcharges approved through regulatory inefficiency*.
- But "cost of service" in ICAP structure is based on administrative determination of the cost of service *for the target level of capacity* – no overbuiding. And ISO COS may (or may not) be better than states? It's the lowest cost for 15% RM.

So the reality is, we get total market prices capped at a level of cost-of-service for exactly the level of reliability set by the RM requirement – roughly what we wanted from markets.

RTOs with These Core Features Support Reliability, Renewables, DR and Contracts

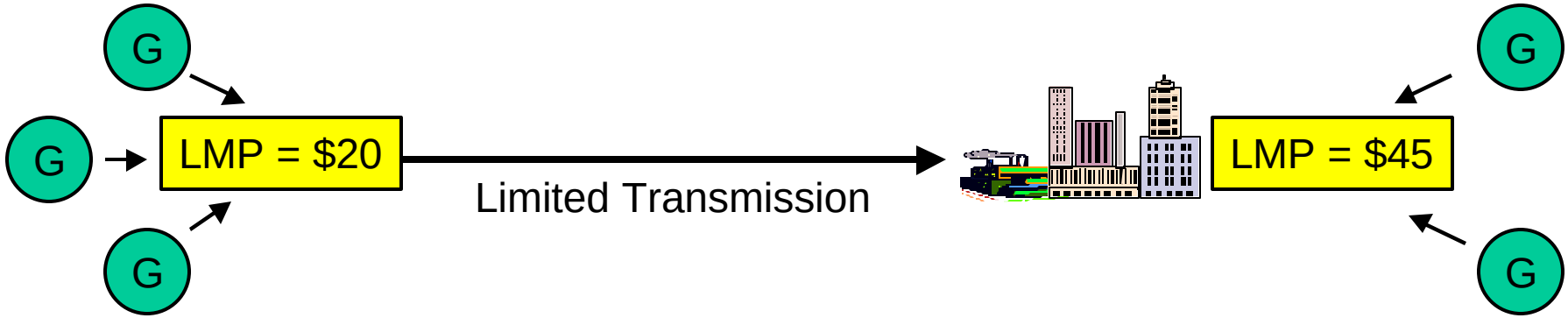


Contract Path Scheduling Is Flawed Because It Ignores the Actual Flows/Physics



Contract path scheduling needs curtailments (TLRs) to "unschedule" the grid to get flows within security limits

Locational Marginal Prices Also Define Transmission Usage Charges



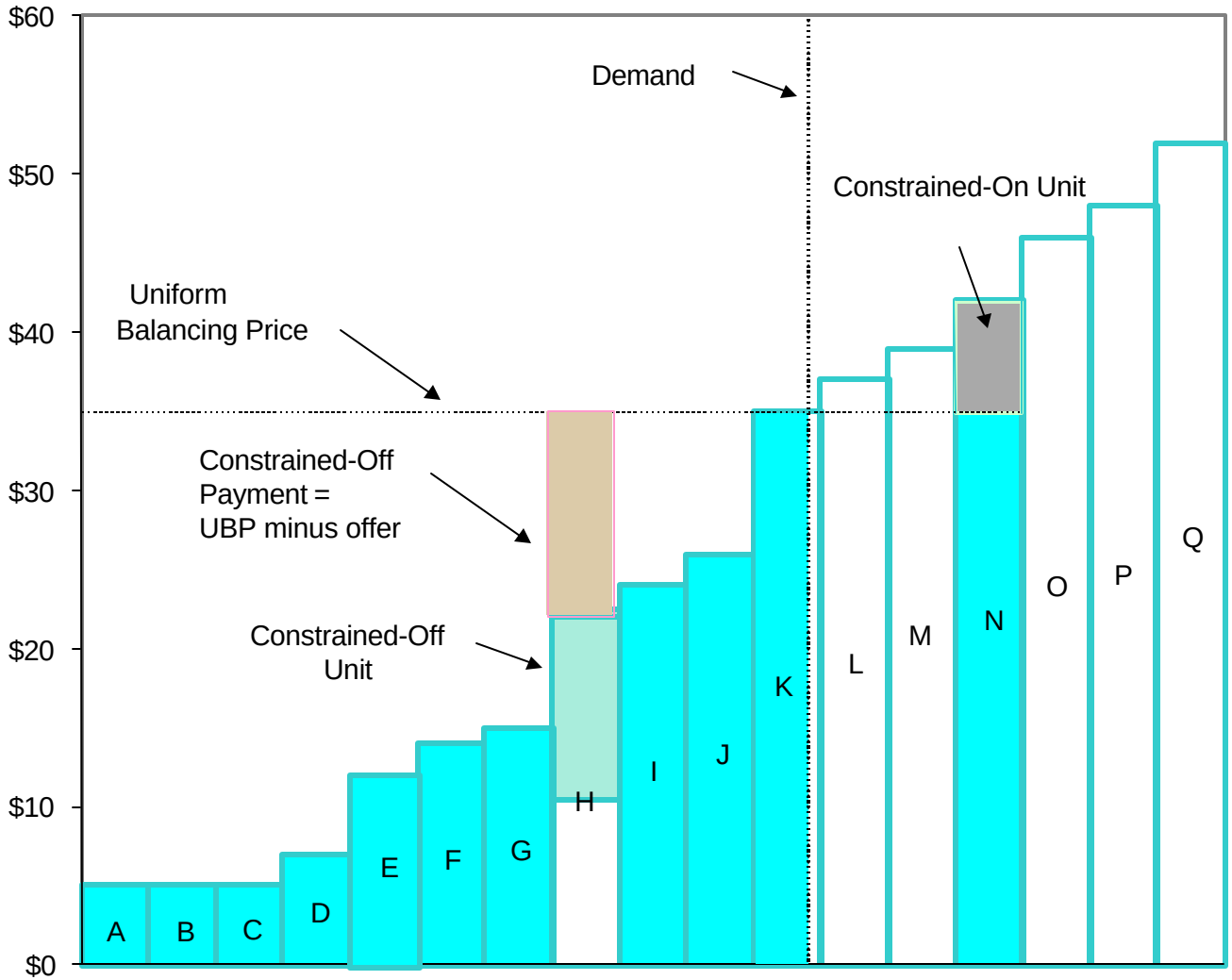
The difference in LMPs between two locations defines the price charged to transmission users.

- The difference between the LMPs at any two locations reflects the marginal cost of the RTO's least-bid-cost, security-constrained redispatch needed to move power from one location to the other, given constraints on the system.

- If there were no congestion, no redispatch would be needed. Locational Prices would differ only slightly because of losses.

Why Not Use the Same Price Everywhere?

Ignoring Price Differences Creates Problems



Constrained-on and –off units must be paid

Non-LMP Prices Encourage Bid "Games" . . .

LMP avoids these "games"