THOUGHTS ON CAPACITY MARKET DESIGN: OBJECTIVES AND CONSISTENCY OF ASSUMPTIONS

Roy J. Shanker Ph. D

This presentation represents my own views and not necessarily those of any of my clients.

RASTF OCTOBER 11, 2022

CONTEXT

- Not meant to be a proposal nor an exhaustive discussion-high level
- A ***look back*** at part of what we have done in the past and the role/need for consistent assumptions in every stage of a complicated implementation
- The types of problems created by deviating from those consistent assumptions
- *Can this be translated into generic or specific "lessons learned" for new designs?*
- *Some straw ideas suggested by this perspective, not advocated per se, but rather how this leads to strings of features*

WHY HAVE A CAPACITY MARKET?

- “Normal” markets allow recovery of fixed costs via competitive pricing without caps that results in inframarginal rents. The pricing can reflect scarcity and result in economic rationing if necessary. (E.g. parties drop out and can no longer purchase a product as the price rises due to limited supply). (Assume no market power in general discussion, though necessary consideration)
- RTO markets don’t have these basic elements.
  - Energy offers are capped and may not reflect or allow clearing offers. Market prices can be capped. Circuit breakers may further limit price formation and associated rents. Limits on reserve price formation and other limits
  - Capacity reliability requirements are set exogenously., *not* by the market. By design, a marginal energy price unit (e.g. peaker) may not be needed frequently or ever for energy supply, and when offered may be marginal.
  - Rationing by price and knowingly allowing shortages is not viable for setting reliability requirements.
  - Result: the missing money issue and a role for Capacity markets. Marginal unit receives no scarcity rents by design.

MAIN DRIVERS FOR PJM CHANGE RPM OBJECTIVES

- A non-functional existing market with no locational attributes, daily accounting and a vertical demand curve.
- Broad Objectives for change:
  - Maintain Reliability
  - Economic entry and exit signals based on missing peaker capacity recovery and capital energy/substitution across traditional generation resources
  - Economic and Adequacy stability
  - Strong locational requirements/signal
  - Ability to address market power

HIGH LEVEL RPM ATTRIBUTES TO ADDRESS DRIVERS

- RTO wide reliability requirement (peak load and reserve requirement)
- Fungible products
- Locational Delivery Areas with local reliability requirements linked to interconnection/gem deliverability and load deliverability requirements
- Downward sloping demand curve anchored to right of reliability targets and reflecting opportunity for recovery of net CONE for peaker (capital/energy substitution) over time
- 3 year forward market with incremental auctions
- Must offer, must purchase, MSOC, MOPR for market power
- Non-Performance penalties
- Effort to validate long ability to meet design objectives (Hobbs)

HIGH LEVEL RPM ATTRIBUTES TO ADDRESS DRIVERS

- Cost recovery in auction reflects basics of missing money concept and overall general view of capital/energy substitution in the electric market
  - Marginal least capital cost unit of capacity is peaker (CONE)
  - Needed recovery of marginal unit with exogenous reliability requirement is Net CONE of peaker
  - This value stays constant over all marginal units because as soon as the marginal unit has sufficient inframarginal rents to cover conversion to next more efficient but capital intensive unit it converts, still leaving the Net CONE of the peaker not recovered
  - E.g. a peaker would convert 1 Mw of capacity to CC as soon as its energy rents exceeded the conversion costs, regardless of whether the Net CONE was recovered

VERY HIGH LEVEL RPM FLOW (PRE ELCC)

- Develop load and generation modes, World diversity/support
- Set IRM, FPR using PRISM
- Confirm transmission assumptions (CETO/CETL)
- Develop RTO and LDA demand curves (Local Reliability Requirement)
- Sufficient lead time to enter, MSOC, MOPR, LDA curves etc.
- Clear BRA, reflect locational constraints and adders
- Demonstrate ability to recover necessary revenues to support entry (on expectation**) over time** (missing money)
- Demonstrate pricing to exit market
- Settle CTRs and reflect full locational requirements

WALK THROUGH RPM BUILDING BLOCKS AND ASSUMPTIONS

- Overall RTO Reliability Requirement : PRISM (Probabilistic Reliability Index Study Model) Two Area Convolution Model, LOLE, Tie Line to World, CBM, Unlimited Transmission (Copper Plate), IRM Target 1 in 10
  - RTO Wide (load and gen at a single point plus tie)
  - Historical load model on peak and 52 weekly shape (PLOTS) (LAS)
  - Generation model, assumes fungible products (e.g. the entire system could be served by a sufficient number of the designated generators without knowing what the underlying generator was), assumes independent outages (eGADS, MWs (seasonal), EFORd, EEFORd, Planned Maintenance Outage), planned class averages)

CONSISTENCY ISSUES WITH ASSUMPTIONS PRISM

- Assumption: Homogeneous/Fungible Resources
  - Key assumption necessary to support all subsequent steps, particularly BRA which assumed all were fungible and only differentiated by outage and MW (and location)
  - Fundamental to getting meaningful valuations out of the auction structure
  - Past and future modeling struggled with different characterizations of “different” units, e.g. various forms of DR or intermittent
  - Either load gets manipulated or some form of equivalent is estimated, but none meet the simple test of being able to build a full system supply out of the various classes created I.E. not fungible
  - Violation skews results, typically to overstate reliability contribution as most proxies are less reliable than the assumed “unlimited” resources.

CONSISTENCY ISSUES WITH ASSUMPTIONS PRISM

- Assumption-Independent Outages
  - Related to fungible
  - Clearly not met in terms of weather and fuel related correlated outages (thermal)
  - Same difficulty for intermittent (particularly solar) in terms of sun and wind
  - Assumption violation skews results to overstate reliability in general
- Assumption-Outage types
  - Model only represents Maintenance and Forced Outages-allocate Planned outages
  - Not clear how this assumption/adjustment matches other elements

CONSISTENCY ISSUES WITH ASSUMPTIONS PRISM

- Load Model
  - Acknowledged to be best for peak representation
  - Weaker for sequential/multi day events
  - Not clear on coordination with seasonal market structures

CONSISTENCY ISSUES WITH ASSUMPTIONS PRISM

- Unlimited Transmission
  - The modeling assumes unlimited transmission. Load and Generation at same point.
  - This limitation is recognized and addressed through subsequent testing and interconnection requirement
    - CETO-level of imports plus internal generation needed to bring LDA to 1 in 25 (versus assumed infinite connection with rest of pool)
    - CETL-actual import level, consistency check v CETO (desire CETL ≥ CETO)
    - Generator deliverability to assure no gen pocket limitation
    - CETO/CETL inherently reflect lower than 1 in 10 for RTO

MANUALS 20 &14B EXCERPTS WHY 1 IN 25?

- (Manual 20) A fundamental assumption of the PJM Reserve Requirement Study is the absence of any transmission constraints within PJM that could result in “bottled” generation. This assumption is tested by Load Deliverability Analysis based on the Capacity Emergency Transfer Objective (CETO) and Capacity Emergency Transfer Limit (CETL) tests. These tests are applied to electrical areas (called Locational Deliverability Areas or LDAs in the RPM process) within the PJM RTO to ensure that the needed capacity resources are deliverable to load. The CETO is defined to be the import capability required by the area to comply with a Transmission Risk LOLE of one event, on average, in 25 Years. The CETL is defined to be the actual emergency import capability of the test area. The CETO is driven largely by the level of generation reserves, unit performance, and load shape characteristics within the test area. An area passes the deliverability test if its CETL is equal to or greater than its CETO. A detailed description of modeling for these tests is contained in this Manual’s references and summarized below. See PJM Manual 14B, Attachment E for further details. The Load Deliverability Method requires the selection of a transmission risk level to define the CETO.***This risk must be very small when compared to the one day in ten year LOLE applicable to generation risk. A transmission LOLE of 1 D/ 25 Y was judged to be sufficiently small. This risk refers to the probability of having to shed load due solely to insufficient transmission import capability, not a shortage of generation resources***. The one day in 25 year LOLE is subject to periodic review.
- (Manual 14b***) The Transmission System is tested at a LOLE of 1/25 so that the transmission risk does not appreciably diminish the overall target of a 1/10 LOLE for PJM.***
- **Actual transmission limits are a fundamental mismatch with base assumptions. Very large and unknown interactions of this violated assumption grow as the level of stochastic supply increases and the ability to reflect this in reliability and power flow studies gets further and further distorted from the assumption of unlimited transmission. HUGE FUTURE PLANNING RISK.**

CONSISTENCY ISSUES CETO/CETL

- CETO/CETL are Vital Elements to Entire Historic Construct
  - Validates basis FPR, PRISM results, Local Reliability Requirement, LDA demand curves, CTR locational adder
  - Inherent is their compatibility with PRISM assumption of fungible products characterized by MW, outages
  - *These assumptions are not true for historic or future system*, in particular in presence of intermittent resources. *My understanding is that **thermal equivalents** are used for intermittent resources in both analyses. Clearly not fungible and also stochastic output is not realistically captured for tests. The more intermittent resources the greater the distortion*
  - *In particular for CETL, a power flow must be analyzed. Thermal equivalents are not representative of stochastic output of intermittent resources. (CETO also uses thermal equivalent which is problematic, but in my mind less risky to PJM reliability than in CETL)*
  - CETO is however inherently understated by use of thermal equivalent by ignoring stochastic production.
  - NET EFFECT OF INCONSISTENCY: CETO TOO LOW AND CETL TOO HIGH--BOTH HARM RELIABILITY AND WERE NOT REFLECTED BY PJM.

CETO/CETL INCONSISTANCY AND LDA CHARACTERIZATION IN RPM

- Local reliability requirement is CETO plus LDA generation resources. This plus the LDA applicable net CONE set LDA demand curve as defined.
  - If mismatch of assumptions mentioned above occurs, then the local reliability is too low and curve misplaced((not reflecting enough demand and associated price)
- CETL sets the transfer limits in the auction optimization
  - If mismatch to assumptions is as discussed, then CETL too high and not enough resources are procured in LDA and price too low (adverse reliability and price)
- Under procures for LDA, Misprices LDA curves, LDA clearing prices and nominally CTR's.
- ***All stem from a break in the "string" of consistency of assumptions***

OVERALL CONCLUSIONS

- The Resource Adequacy Modeling and Implementation is highly dependent on assumptions
- Historically assumed “little” issues and fixes have been used to cover mismatches between modeling assumptions and reality AND preferences
- As little issues grow, the fixes become worse and worse, the distortions and reliability risks larger
- A corollary is that the distortions create a constituency to perpetuate the “fixes” and assumption violations making bad situations worse and harder to fix
- Downhill spiral leading us to the range of RA issues we are facing today: little accommodations breed very large future distortions and barriers to real fixes/accurate reliability assessment and pricing.
- **We still don’t have a uniform test/performance approach to long term viability of a design v entry, exit and cost recovery in presence of subsidies.**

DO THESE CONCLUSIONS SUGGEST POTENTIAL FIXES?

- Yes. I don't want to get into specific proposals, but at a high level several options might work if we start from objective maximizing efficiency and minimizing any deviations from assumptions of the underlying basic complete model and assumptions:
- E.g. Straw 1.
  - a) Product definition, make everything look like what we now call "unlimited" based on a physical performance definition. This could accommodate all types of generation but force broadly defined aggregations.
  - b) Keep current auction, no need to modify approach to clearing auction, validate LDA logic and associated transmission. Average is marginal
  - c) Find a way to address locational issues with intermittent resources (possibly easier if the defined product aggregate looks like "unlimited", but still major task (continuing major problem)

STRAW 2

- Keep current auction structure but:
  - Uniform fungible product
  - Fix transmission constraints
  - Put any resources that can't meet fungible product definition "on the load side", i.e. their capacity value is reflected by load reductions based on xxx hour peak load reduction allocation of the non-traditional capacity resources
  - Shifts limitations of ability to define uniform product into risk on ability to actually modify load's billing determinants. Something like partitioning between uniform product and ELCC but not exactly the same.

STRAW 3

- Full Marginal ELCC
  - No Prism, the ELCC tool (including a locational solution that works) does it all
  - All resources are subject to evaluation but with refined class definitions and limited energy up to CIR, new deliverability tests etc.
  - We need to have a solution to locational ELCC characterization **and** possibly different notions or approach to concepts like CETO/CETL
  - Need definition to deal with extra MW and obligations to perform as ELCC accreditation declines (notionally it always must s at the pre-ELCC accreditation level)
  - Also need a way to reflect simultaneous interaction of quantity, price and ELCC value.
  - We need solution to transmission limits with high levels of stochastic supply (common problem but likely most important here). THIS IS A KEY ELEMENT TO ANY PACKAGE OF CONSISTENT RA SOLUTIONS GOING FORWARD

OBSERVATION ON STRAWS

- I am not favoring any of these example straws.
- The intent is to show a range of alternatives might work and be consistent within the scope of SOME of the full set of relevant assumptions
- The intent is to show how such package might differ and require very different implementation
- The intent is that right now my concern is very high concerning the interaction with respect to locational capacity issues/constraints and the testing/characterization of the transmission system in the presence of stochastic supply options/intermittent resources.