Economics of Grid-Supported Electric Power Markets: A Fundamental Reconsideration

Leigh Tesfatsion

Abstract
U.S. centrally-managed wholesale power markets operating over high-voltage AC transmission grids are currently transitioning from heavy reliance on fossil-fuel based power to greater reliance on renewable power. This study highlights four conceptually-problematic economic presumptions reflected in the legacy core design of these markets that are hindering this transition. The key presumption is the static conceptualization of the basic transacted product as energy amounts competitively determined for delivery at designated grid locations during successive operating periods, supported by ancillary services. The study then discusses an alternative “linked swing-contract market design” that appears better-suited for the support of increasingly decarbonized grid operations. This design entails a fundamental switch to a dynamic insurance focus on advance reserve procurement permitting continual balancing of real-time net load. Reserve consists of the guaranteed availability of flexible power-path production capabilities for possible centralized dispatch during future operating periods, offered into linked centrally-managed forward reserve markets by two-part pricing swing contracts in firm or option form.

Key words: U.S. centrally-managed wholesale power markets, grid decarbonization, market design, fundamental conceptual economic issues, linked swing-contract market design, physically-covered insurance, two-part pricing, swing contracts, digital-twin production capability sets
# Table of Contents

1 Introduction ......................................................... 3

2 Legacy Core Design of U.S. RTO/ISO-Managed Markets .... 6

3 Legacy Core Design: Fundamental Conceptual Issues ............. 11
   3.1 Overview: Four Problematic Presumptions (P1)–(P4) ........... 11
   3.2 (P1): Energy is the Basic Transacted Product ................. 12
   3.3 (P2): Two-Part Partition for Total Supplier Cost ............... 16
   3.4 (P3): Energy Functions as a Commodity ...................... 20
   3.5 (P4): Spot Markets Ensure Fixed-Cost Coverage ............. 28

4 Legacy Core Design: Roadblock for Grid Decarbonization . 30
   4.1 Overview ......................................................... 30
   4.2 Proliferation of Participation Models ......................... 30
   4.3 Proliferation of Flexibility Products ........................... 31
   4.4 Proliferation of Out-of-Market Make-Whole Payments ....... 33
   4.5 Growing Revenue Insufficiency Concerns ..................... 35
   4.6 Dangerous Market-Design Conundrum .......................... 36

5 Conceptually-Consistent Alternative Market Design ............. 36
   5.1 Linked Swing-Contract Market Design: Overview ............. 36
   5.2 Key Design Features .......................................... 37
   5.3 Comparison with U.S. RTO/ISO-Managed Markets ............ 40

6 Conclusion ............................................................ 42

Appendices: Quick-Reference Glossaries and Guides ............... 44
   A.1 Acronyms ......................................................... 44
   A.2 Standard Transmission System Terms ......................... 45
   A.3 Standard Economic Terms ..................................... 46
   A.4 Cost Types for Power Markets: Empirical Examples ......... 48
   A.5 Swing-Contract Market Terms .................................. 49

References ............................................................. 50
1 Introduction

The basic purpose of centrally-managed wholesale power markets operating over high-voltage AC transmission grids is to maintain efficient, reliable, just-in-time production and transmission of bulk power to satisfy just-in-time customer power demands.

To achieve this dynamic open-ended purpose, the central manager must continually protect against volumetric grid risk: namely, the possible disruption or collapse of grid operations due to real-time imbalance between the injection of power into the grid and the withdrawal and/or inadvertent loss of power from the grid. Power withdrawals occur when the power usage of customers electrically connected to the grid exceeds their locally-used power generation. Inadvertent power losses occur whenever power flows across transmission grid lines.

In response to private economic incentives and public policy mandates encouraging grid decarbonization [21], U.S. RTO/ISO-managed wholesale power markets are striving to transition from a traditionally heavy reliance on fossil-fuel based power generators to a greater reliance on renewable power facilities. The latter facilities include intermittent power resources (IPRs), such as grid-connected wind farms and photovoltaic solar arrays whose weather-dependent power generation is not fully firmed by storage.

The increasing participation of IPRs in U.S. RTO/ISO-managed wholesale power markets, together with recent initiatives such as FERC Order No. 2222 [14] encouraging more active customer participation in these markets, has increased the uncertainty and volatility of grid net load, i.e., customer power withdrawals and inadvertent power losses minus net non-dispatched power injections. Moreover, many IPRs connect to high-voltage AC transmission grids by means of power electronic inverters that convert DC to AC power, a connection technology that differs fundamentally from the traditional connection technology for fossil-fuel based power generators. At higher IPR penetration levels, this new connection technology can pose new types of system security issues [4].

In consequence, as reported in [15], RTOs/ISOs are finding it harder to procure the dependable availability of RTO/ISO-dispatchable power with sufficiently flexible attributes to maintain reliable real-time balancing of net load.  

---

1 Current U.S. RTO/ISO-managed wholesale power markets consist of energy, ancillary service, and capacity markets whose operations over high-voltage AC transmission grids are managed by a Regional Transmission Organization (RTO) or Independent System Operator (ISO); see [16].

2 An intermittent power resource (IPR) is a grid-connected power resource whose power injections and/or withdrawals are not mediated through some form of aggregator and are not fully controllable by centrally-managed dispatch.

3 In practice, reliable real-time balancing of net load means maintaining net-load balance within acceptable tolerance levels over time.
The remainder of this study is organized as follows. Section 2 reviews the “two-settlement system” design at the core of current U.S. RTO/ISO-managed wholesale power market operations: namely, the daily RTO/ISO-coordinated operation and settlement of a Day-Ahead Market (DAM) and a Real-Time Market (RTM). Four conceptually-problematic economic presumptions reflected in this two-settlement system design are then identified in Section 3:

**Problematic Presumption (P1):**
The basic product transacted in grid-supported U.S. RTO/ISO-managed wholesale power markets is delivered energy (MWh), i.e., flows of power (MW) accumulated at designated grid delivery locations during designated delivery time-periods with duration measured in hourly units (h).

**Problematic Presumption (P2):**
To analyze supplier cost in such markets, it suffices to partition total supplier cost into a “variable” component determined by supplier-delivered energy and a “fixed” component independent of supplier-delivered energy.

**Problematic Presumption (P3):**
Within the context of such markets, energy (MWh) conditional on delivery location and delivery time-period is a commodity with perfectly substitutable units whose uniform price ($/MWh) and transacted amounts (MWh) should be determined in a competitive commodity spot market.

**Problematic Presumption (P4):**
Supplier revenue attained in these competitive commodity spot markets will suffice over time to cover total supplier cost.

Presumptions (P1)–(P4) reflect the misleading view that RTOs/ISOs are fiduciary managers for weakly-correlated collections of competitive commodity spot markets. The reality is far more daunting: RTOs/ISOs are fiduciary “conductors” tasked with orchestrating the availability and subsequent possible dispatch of increasingly-diverse dispatchable power resources to service just-in-time power demands of increasingly diverse grid-connected customers while meeting just-in-time power requirements for reliable grid operation.

Section 3 identifies and analyzes various conceptually-problematic aspects of presumptions (P1)–(P4). Four counter-claims are developed and justified:

**Counter-Claim 1:**
Within the context of a grid-supported RTO/ISO-managed wholesale power market, energy (MWh) is not a commodity, i.e., energy (MWh) is not an asset with a well-defined unit of measurement such that, conditional on delivery location and delivery time-period, each market participant considers each energy unit to be a perfect substitute for each other energy unit.

**Counter-Claim 2:**
Any market process carried out within the context of a grid-supported RTO/ISO-managed wholesale power market must necessarily be a forward market process due to the speed of real-time grid operations.
**Counter-Claim 3:**

Dispatchable power resources participating in grid-supported RTO/ISO-managed wholesale power markets can provide two distinct types of service:

- **Physically-Covered Insurance:** The guaranteed availability of flexible power-path production capabilities for possible RTO/ISO dispatch during future operating periods to protect against volumetric grid risk;

- **Real-Time Power-Path Delivery:** The actual delivery of power-paths in response to RTO/ISO-dispatch signals received during successive operating periods to meet just-in-time customer power demands and grid reliability requirements.

**Counter-Claim 4:**

Two-part pricing contracts, based on a partition of supplier cost into three economically relevant components, can be used for supply offers in grid-supported RTO/ISO-managed wholesale power markets to allow dispatchable power resources to receive separate appropriate compensation for physically-covered insurance and real-time power-path delivery.

Section 4 carefully considers how the retention of the four legacy economic presumptions (P1)-(P4) in the core design of current U.S. RTO/ISO-managed wholesale power markets is hindering the ability of these markets to transition smoothly to decarbonized grid operations. Indeed, it is argued that adherence to (P1)-(P4) has resulted in a dangerous conundrum. Recalling the slow and difficult transition from the earth-centric circular-orbit solar-system model developed by Claudius Ptolemy (circa 100-170 AD), supported by a proliferation of postulated “epicycles,” to the sun-centric elliptical-orbit solar-system model developed by Nicolaus Copernicus (1473-1543), this conundrum can be characterized as follows:

**Ptolemaic Epicycle Conundrum for Power Markets (“Onion Problem”):**

Current net-load balancing issues are addressed by instituting a new rule-layer (“epicycle”) governing trade and settlement for an additional ancillary service product, which in turn gives rise to new net-load balancing issues.

Section 5 reviews key features of an alternative design for grid-supported wholesale power markets, consistent with Counter-Claims 1-4, that appears better suited for decarbonized grid operations: namely, the Linked Swing-Contract Market Design [36]. The latter design fundamentally changes the conceptualization of the basic transacted product from a static focus on energy deliveries dispatched during successive operating periods, supported by proliferating ancillary services, to a dynamic focus on insurance procurement guaranteeing the availability of power-production capabilities for possible dispatch during future operating periods.

More precisely, the Linked Swing-Contract Market Design posits a linked collection of RTO/ISO-managed forward reserve markets M(T) for future operating periods T. Reserve for T consists of the guaranteed availability of flexible power-path production capabilities for possible RTO/ISO dispatch during T, offered into linked RTO/ISO-managed forward reserve markets by two-part pricing swing contracts in firm or option form. As illustrated in Fig. 1, a power-path for T is a sequence of power injections and/or power withdrawals at a single grid location during T.
Fig. 1 One of many possible power-paths $p_m(T) = \{p_m(t) \mid t \in T\}$ consisting of successive injections/withdrawals of power $p_m(t)$ (MW) that an RTO/ISO-dispatchable power resource $m$ with swing (flexibility) in ramp-rate (MW/min) and power capacity (MW) could be dispatched to deliver at its grid location $b_m$ during a future operating period $T$.

For each forward reserve market $M(T)$, the goal of the RTO/ISO is to procure reserve with sufficient swing (flexibility) in power-path attributes to permit maximization of expected period-$T$ total net benefit for market participants subject to system constraints that include nodal net-load balance constraints.

The RTO/ISO procures reserve by clearing reserve offers submitted to $M(T)$ by dispatchable power resources. Each reserve offer is a two-part-pricing swing contract in firm or option form. Two-part pricing enables separate time-consistent settlements for period-$T$ volumetric risk-reduction (reserve availability) provided in advance of $T$ and for period-$T$ performance (dispatched power-path delivery) verified subsequent to $T$. This eliminates the need for separate or co-optimized energy markets; and it also eliminates the need to rely on out-of-market (OOM) make-whole payments in order to ensure supplier revenue sufficiency.

Concluding remarks are given in Section 6. Quick-reference glossaries and guides for key terms and concepts appearing in the main text of this study are provided in appendices.

2 Legacy Core Design of U.S. RTO/ISO-Managed Markets

The development of the legacy core two-settlement system design supporting current U.S. RTO/ISO-managed wholesale power market operations can be traced in a series of reports released by the U.S. Federal Energy Regulatory Commission (FERC), culminating in a 2003 White Paper [8].

In the latter report, FERC envisions delivered energy (accumulations of power) at designated grid locations during designated time-periods to be the basic transacted product. These delivered energy amounts are to be determined in accordance with a two-settlement system [30] consisting of a daily bid/offer-based RTO/ISO-managed Day-Ahead Market (DAM) operating in tandem with a daily bid/offer-based RTO/ISO-managed Real-Time Market (RTM); see Fig. 2.
The overall goal of the DAM/RTM two-settlement system is to permit energy transactions at designated grid delivery locations during designated delivery time-periods to be efficiently determined by the supply offers and demand bids of energy suppliers and buyers. With this overall goal in mind, the DAM/RTM two-settlement system is designed to be in accordance with the determination of market-clearing prices and quantities in competitive commodity spot markets (CCSMs) to an extent consistent with maintaining the reliable support of a transmission grid susceptible to transmission-line congestion.

The purpose of the RTO/ISO-managed DAM held on each day D is to commit RTO/ISO-dispatchable generation units for day D+1 that permit the RTO/ISO to ensure efficient continual net-load balancing during day D+1.

Load-Serving Entities (LSEs) acting on behalf of managed customers can submit demands bids into the day-D DAM for the purchase of energy at designated grid delivery locations for each hour H of day D+1. Each such demand bid can take the form of a fixed (non-dispatched must-service) energy demand. It can also include or take the form of a dispatchable price-sensitive energy demand schedule if: (i) the LSE has installed real-time telemetry permitting the RTO/ISO to incrementally increase or decrease the LSE’s energy demand by dispatch signals; and (ii) the LSE can implement any received RTO/ISO dispatch signals by suitable instructions communicated to its managed customers.

Generation units can submit supply offers into the day-D DAM for the sale of energy at designated grid delivery locations for each hour H of day D+1. Each such supply offer can take the form of a fixed (non-dispatched must-service) energy supply. It can also include or take the form of a dispatchable price-sensitive energy supply schedule if the generation unit has installed real-time telemetry permitting the RTO/ISO to incrementally increase or decrease the generation unit’s energy supply by dispatch signals.
The RTO/ISO conducts a bid/offer-based SCUC/SCED optimization for the day-D DAM in combined or consecutive form. This optimization is conditional on current state conditions, submitted bids and offers, and forecasts for IPR injections and/or withdrawals of power at each designated grid delivery location during each hour H of day D+1; and it is subject to system constraints that include a net-load balancing requirement at each designated grid delivery location for each hour H of day D+1. The optimization determines a binary (yes/no) commitment solution for each dispatchable generation unit for each hour H of day D+1 indicating whether or not this generation unit is required to be available for possible RTO/ISO-dispatch during hour H of day D+1. It also determines anticipated dispatch schedules for price-sensitive energy demands and supplies at each designated grid delivery location for each hour H of day D+1.

An RTM is a daily collection of sub-markets for near-term future time-periods \( \tau \) with relatively short durations (e.g., 15 minutes). These RTM sub-markets are cleared by RTO/ISO-managed SCED optimizations conditional on RTO/ISO-updated forecasts for fixed (non-dispatched must-service) energy demands and supplies for \( \tau \). The purpose of these RTM sub-markets (together with supplemental unit-commitment processes) is to permit the successive modification of previously determined SCUC/SCED optimal solutions to take into account RTO/ISO-updated forecasts as well as unanticipated changes in other relevant factors.

Settlements for cleared bids and offers are determined by *locational marginal pricing* [32]; that is, by the pricing of energy (MWh) conditional on its delivery location and delivery time-period, subject to system constraints. The *Locational Marginal Price* \( \text{LMP}(b, H, D+1) \) ($/MWh) determined in a day-D DAM SCED optimization for scheduled energy transactions at a designated grid location \( b \) during some designated hour H of day D+1, conditional on generation-unit commitments determined in a previously conducted day-D DAM SCUC or SCUC/SCED optimization for day D+1, is the dual variable solution for the net-load balancing constraint at location \( b \) for hour H. Any adjustments needed in the scheduled energy transactions determined in the day-D DAM for some hour H of day D+1, as indicated by the solutions for RTM sub-markets conducted after the close of the day-D DAM but prior to hour H, are settled using the LMPs determined in these RTM sub-markets.

Fig. 3 illustrates the determination of an optimal (D=S) solution for a specific hour H by a bid/offer-based RTO/ISO-managed DAM or RTM SCED optimization conducted prior to H. The depicted aggregate demand and supply schedules D and S...

---

4 SCUC is an acronym for *Security-Constrained Unit Commitment*, and SCED is an acronym for *Security-Constrained Economic Dispatch*.


6 This solution consists of a set of optimal power-LMP points \( (p^*, \, \pi^*) \) for H because the optimal \( \text{LMP}^* \) is indeterminate over the indicated range. This indeterminacy arises because the demand bids and supply offers submitted to this DAM take an energy-block (step-function) form that results in flat vertical and horizontal segments for the aggregate demand and supply schedules D and S.
Fig. 3 Illustrative depiction of the optimal market-clearing (demand = supply) solution for the maintained power-withdrawal levels of cleared buyers and the maintained power-injection levels of cleared suppliers for a designated operating hour $H$, as determined by a bid/offer-based RTO/ISO-managed DAM or RTM SCED optimization conducted prior to $H$. The solution is not conditional on grid location, indicating absence of grid congestion at this optimal solution.

$S$ are constructed from the LSE demand bids and generation-unit supply offers submitted to this prior SCED optimization by two LSE buyers (B1, B2) and three generation-unit suppliers (S1, S2, S3).

Note the optimal market-clearing outcomes depicted in Fig. 3 are indeed outcomes for an energy market, despite the appearance of power levels (MW) along the quantity axis. The power levels along the quantity axis are possible choices for a maintained power level (MW) for operating hour $H$. Hence, choice of a power level $p$ is equivalent to choice of the delivered energy amount $p \cdot 1h$ (MWh) for hour $H$.

Finally, participants in an RTO/ISO-managed DAM (or RTM) for a future operating period $T$ are assured, by design, that any energy transactions the RTO/ISO announces it has scheduled for $T$ are supported by scheduled transmission capacity. Traders who determine and settle physically-covered bulk energy trades

---

7 The aggregate demand schedule $D$ in Fig. 3 gives, from left to right, the highest purchase reservation value ($$/MWh$) – i.e., the highest maximum willingness to pay ($$/MWh$) – for each successive unit (MW) increase in the maintained power level $p$ for $H$, where this highest purchase reservation value is calculated across all buyers (here B1 and B2). Conversely, the aggregate supply schedule $S$ in Fig. 3 gives, from left to right, the lowest sale reservation value – i.e., the lowest minimum acceptable payment – for each successive unit (MW) increase in the maintained power level $p$ for $H$, where this lowest sale reservation value is calculated across all suppliers (here S1, S2, and S3). The optimal (market-clearing) solution points, by definition, are then given by the intersection points of these aggregate $D$ and $S$ schedules with all flat and vertical segments included. Compare, for example, the demand bid, supply offer, and market-clearing concepts used by the Midcontinent Independent System Operator (MISO) for its day-ahead and real-time energy markets [29]. For an extended discussion of these fundamental economic concepts, see [35] and [36, Ch. 12].
for T through other venues, such as privately-negotiated bilateral trades, must secure transmission-capacity support for these physically-covered trades by self-scheduling them as fixed-form energy bids and offers in a DAM (or RTM) conducted for T. In addition, these traders might need to procure supporting contracts (e.g., Financial Transmission Rights and Contracts-for-Difference) to ensure that the settlement terms they agreed to in these other venues are not disrupted by obligatory DAM/RTM LMP price settlements; see [36, Sec. 13.9].

Seven regions of the U.S. are currently operating RTO/ISO-managed wholesale power markets in accordance with FERC’s proposed two-settlement system design; see Fig. 4. As seen in Fig. 5, these seven RTOs/ISOs operate over a physical high-voltage AC transmission grid consisting of three separately synchronized parts.

Fig. 4 North American RTO/ISO-managed wholesale power markets. (Public domain: [13])

Finally, FERC’s proposed two-settlement system design did not include guidance for provision and settlement of ancillary services. Rather, FERC explicitly delegated ancillary service aspects of power system management to the individual states participating in each RTO/ISO [8, p. 11]. Consequently, as reported in [6, Tables 1-2] and [15, Table 1, p. 6], ancillary service procurement and settlement processes differ widely across the seven U.S. RTOs/ISOs.

---

Ancillary services are support services for grid reliability [15, Appendix]. Examples include: “black-start” services for restoration of power flow to a collapsed grid; reactive power support for voltage control; and net-load balancing services provided by on-line generation units with unencumbered capacity or by off-line relatively quick-start generation units.
3 Legacy Core Market Design: Fundamental Conceptual Issues

3.1 Overview: Four Problematic Presumptions (P1)–(P4)

The DAM/RTM two-settlement system design reviewed in Section 2, originally proposed by FERC [8] in 2003, constitutes the core design of all seven current U.S. RTO/ISO-managed wholesale power markets. As noted in Section 1, this core design reflects four conceptually-problematic economic presumptions, as follows:

**Problematic Presumption (P1):** The basic product transacted in grid-supported RTO/ISO-managed wholesale power markets is delivered energy (MWh), i.e., flows of power (MW) accumulated at designated grid delivery locations during designated delivery time-periods with duration measured in hourly units (h).

**Problematic Presumption (P2):** To analyze supplier cost in such markets, it suffices to partition total supplier cost into a “variable” component determined by supplier-delivered energy and a “fixed” component independent of supplier-delivered energy.

**Problematic Presumption (P3):** Within the context of such markets, energy (MWh) conditional on delivery location and delivery time-period is a commodity with perfectly substitutable units whose uniform price ($/MWh) and transacted amounts (MWh) should be determined in a competitive commodity spot market.

**Problematic Presumption (P4):** Supplier revenue attained in these competitive commodity spot markets will suffice over time to cover total supplier cost.

Specific conceptually-problematic aspects of each of the legacy economic presumptions (P1) through (P4) will next be taken up in turn. 
3.2 Presumption (P1): Energy is the Basic Transacted Product

3.2.1 (P1): Two Key Conceptual Concerns

Presumption (P1) is consistent with the focus of current U.S. RTO/ISO-managed wholesale power markets on amounts of energy bought and sold at designated grid locations during designated operating periods. This section stresses two important conceptual concerns regarding this presumption.

- **Physical Reliability Concern:** Energy transactions in U.S. RTO/ISO-managed wholesale power markets must be supported by the physical operations of underlying transmission grids. The necessary requirements for the reliable operation of these grids over successive operating periods \( T \) cannot be expressed solely in terms of transacted and delivered amounts of energy.

- **Benefit and Cost Valuation Concern:** How power (MW) is injected at designated grid locations during successive operating periods can matter greatly to RTOs/ISOs, power producers, and power customers. Delivered energy amount is only one of many possible valued attributes of this flow of power.

3.2.2 (P1): Physical reliability Concern

In order for a transmission grid to operate reliably over time, the grid must be in continual net-load balance. Roughly stated, this means that the injection of power into the grid must balance the withdrawal and/or inadvertent loss of power from the grid at each point in time.

More carefully stated, continual net-load balance for a transmission grid means that Kirchhoff’s Current Law must hold for this grid. Applied to any electrical network at a given point in time, this law asserts the following: The algebraic sum of all currents entering a network node \( n \) must equal the algebraic sum of all currents exiting node \( n \), where current \( i \) is measured in amperes (A).

U.S. RTO/ISO-managed wholesale power markets operate over high-voltage alternating-current (AC) transmission grids. Consider the analytical modeling developed in [36, Ch. 6 & Sec. 9.2] for an RTO/ISO-managed wholesale power market \( M(T) \) for a future operating period \( T \), where \( M(T) \) operates over a high-voltage AC transmission grid with buses \( b \) in a bus-set \( B \). The participants in \( M(T) \) consist of the following entities. For each bus \( b \in B \),

- a collection \( M(b) \) of dispatchable generation-units \( m \) with unique electrical connection to the transmission grid at transmission bus \( b \);
- a collection \( LS(b) \) of LSEs \( j \), each of whom manages power-usage for a distinct collection \( C_j(b) \) of customers with unique electrical connection to the transmission grid at transmission bus \( b \);
- a collection \( NG(b) \) of non-dispatchable generation-units \( n \) with unique electrical connection to the transmission grid at transmission bus \( b \).

The relative timing of \( M(T) \) and \( T \) are depicted in Fig. 6.
The AC power-flow operations of the transmission grid for M(T) are approximated in [36, Ch. 6 & Sec. 9.2] as direct-current (DC) power-flow operations. Given this DC power-flow approximation, losses are zero; and Kirchhoff’s Current Law expressed in terms of current $i$ measured in amperes (A) can equivalently be expressed in terms of (active) power $p$ measured in megawatts (MW) using $p = v_o \cdot i$, where $v_o$ measured in volts (V) denotes the constant voltage magnitude assumed for the DC power-flow approximation.\footnote{See [33, Sec. 3.1] for a careful discussion of the standard assumptions used to derive a DC power-flow approximation for the AC power-flow of a high-voltage AC transmission grid.}

The net-load balance constraints for this DC-approximated transmission grid for a given operating period $T$ can then be summarized as follows, where the qualifier fixed is used as a short-hand expression for non-dispatched must-service. For each bus $b \in B$ and time $t \in T$, the total dispatched power injection at bus $b$ by the dispatchable generation-units $m \in M(b)$, plus the total net line-power inflow at bus $b$ from buses $b' \in B$ with $b' \neq b$, must equal the total forecasted net load at bus $b$, calculated as the total dispatched customer load at bus $b$ for customers of the LSEs $j \in LS(b)$ plus the total forecasted fixed customer load at bus $b$ for customers of the LSEs $j \in LS(b)$ minus the total forecasted fixed power injection at bus $b$ by the non-dispatchable generators $n \in NG(b)$.

For the purposes of this section, however, it is important to express these net-load balance constraints in their explicit mathematical form [36, Ch. 6 & Sec. 9.2]:

\textit{Net-load balance constraints:} For each bus $b \in B$ and time $t \in T$,

$$G^\text{dis}_b(t) + \text{NLPI}_b(t) = L^\text{dis}_b(t) + \hat{NL}^f_b(t)$$

(1)

where:

- $G^\text{dis}_b(t) = \sum_{m \in M(b)} p^\text{dis}_m(t)$ (total dispatched power injection);
- $\text{NLPI}_b(t) = \left[ \sum_{\ell \in L^E(b)} w^{\ell}_b(t) - \sum_{\ell \in L^O(b)} w^{\ell}_b(t) \right]$ (total net line-power inflow);
- $L^\text{dis}_b(t) = \sum_{j \in LS(b)} p^\text{dis}_j(t)$ (total dispatched customer load);
- $\hat{NL}^f_b(t) = [\hat{L}^f_b(t) - \hat{G}^f_b(t)]$ (total forecasted net fixed load);

Fig. 6 Time-line for a grid-supported RTO/ISO-managed wholesale power market M(T) conducted for a future operating period $T$. 

\[Fig. 6\] Time-line for a grid-supported RTO/ISO-managed wholesale power market M(T) conducted for a future operating period T.
Step functions are universal approximators for the class of all continuous real-valued functions. Specifically, constraints (1) impose joint restrictions on the following four types of power-paths at each bus \( b \in \mathbb{B} \) during \( T \):

- the dispatched power-path \( p_{m}^{\text{dis}}(T) = \{ p_{m}^{\text{dis}}(t) \mid t \in T \} \) at \( b \) during \( T \) for each dispatchable generation-unit \( m \in M(b) \);
- the dispatched power-path \( p_{j}^{\text{dis}}(T) = \{ p_{j}^{\text{dis}}(t) \mid t \in T \} \) at \( b \) during \( T \) for the managed customers of each LSE \( j \in \mathbb{L}(b) \);
- the forecasted fixed power-path \( p_{j}^{\text{f}}(T) = \{ p_{j}^{\text{f}}(t) \mid t \in T \} \) for fixed power withdrawals at \( b \) during \( T \) for the managed customers of each LSE \( j \in \mathbb{L}(b) \);
- the forecasted fixed power-path \( p_{n}^{\text{f}}(T) = \{ p_{n}^{\text{f}}(t) \mid t \in T \} \) for fixed power injections at \( b \) during \( T \) by each non-dispatchable generator \( n \in \mathbb{N}(b) \).

As demonstrated in [36, Ch. 7], given mild regularity conditions and a finite-duration operating period \( T = [t^{s}, t^{e}] \), it is possible to approximate each of these power-paths as closely as desired by a step-function consisting of a discretized sequence of "energy blocks" \( E_{k} \) for a collection \( \mathbb{K}(T) = \{ k_{n} \mid n = 1, \ldots, N(T) \} \) of half-open sub-periods \( k_{n} = [k_{n}, k_{n}^{-}] \) constituting a suitably-refined partition of \( T \). However, as demonstrated in [36, Chs. 7,16] and [24], it is then important to use market optimization formulations expressed in run-time variables to ensure system constraints are imposed with sufficient accuracy. For example, the constraints imposed at the start-time \( k_{n}^{+} \) of each successive sub-period \( k_{n} \) of \( \mathbb{K}(T) \) should be expressed in terms of run-time min/max limits on power-capacities and ramp-rates for sub-periods \( k_{n}^{-} \geq k_{n} \), i.e., min/max limits that are conditional on the state of each market participant at the start-time \( k_{n}^{+} \) for sub-period \( k_{n} \), given the specific solution trajectory assumed through sub-period \( k_{n-1}^{-} \).

In addition, it could be advantageous, or even necessary, to use partitions of \( T \) with different sub-period lengths \( \Delta k \) for different types of market participants in

---

10 As defined in [36], a power-path \( p(T) \) for a grid operating over a time-interval \( T \) is a sequence of injections and/or withdrawals of power (MW) that take place at a single grid location during \( T \).

11 Step functions are universal approximators for the class of all continuous real-valued functions \( f : [a, b] \rightarrow \mathbb{R} \) defined over compact intervals \( [a, b] \) of the real line. For example, given any such function \( f \), and any \( \varepsilon > 0 \), it is straightforward to establish the existence of a step function \( f_{\varepsilon} : [a, b] \rightarrow \mathbb{R} \) with infinitely many time-steps for which the maximum absolute approximation error \( |f(x) - f_{\varepsilon}(x)| \) over \( x \in [a, b] \) is less than \( \varepsilon \). This assertion follows immediately from the Heine-Cantor Theorem, which establishes that any continuous function \( f : X \rightarrow Y \) between metric spaces \( X \) and \( Y \), with \( X \) compact, is uniformly continuous.
order to capture adequately the specific static and dynamic attributes of their power-paths. For a related discussion, see [20, Sec. 3.1.1].

3.2.3 (P1): Benefit and Cost Valuation Concern

A more serious conceptual concern regarding presumption (P1) is that it prevents a comprehensive high-fidelity valuation of benefits and costs.

Let $M(T)$ denote an RTO/ISO-managed wholesale power market $M(T)$ operating over a high-voltage AC transmission grid for a future operating period $T = [t^*, t^e]$. Let $p(b, T) = \{p(b, t) \mid t \in T\}$ denote a power-path for $T$ that consists of a sequence of power injections and/or withdrawals $p(b, t)$ (MW) at a particular grid location $b$ during times $t \in T$. Let $m(b)$ denote an RTO/ISO-dispatchable generation unit electrically connected to $b$ during $T$, and let $C_j(b)$ denote a collection of customers $c_j(b)$ electrically connected to $b$ during $T$ who are serviced by an LSE $j \in LS(b)$.

Suppose $T$ has finite duration; and suppose the power-path $p(b, T)$ has a continuous extension over $\bar{T} = [t^*, t^e]$, the compact closure of $T$. It then follows from the discussion in Section 3.2.2 that $p(b, T)$ can be approximated arbitrarily closely over $T$ by a suitably-constructed step function. Plotted in a time-MW plane, this approximating step-function consists of a finite sequence of one or more “energy blocks” $E_k$ for a collection $K(T) = \{k_n \mid n = 1, \ldots, N(T)\}$ of half-open sub-periods $k_n = [k_{sn}, k_{en})$ that partition the operating period $T = [t^*, t^e]$.

Consequently, if power-path $p(b, T)$ is dispatched at $b$ during $T$, its total energy delivery at $b$ during $T$ can be calculated in close approximate form by adding up the energy-blocks for its approximating step-function. However, this in no way guarantees that the actual value assigned to $p(b, T)$ by a producer $m(b)$, a customer $c_j(b)$, or the RTO/ISO can be expressed solely as a function of this delivered energy.

A generation unit $m(b)$ dispatched by the RTO/ISO to deliver power-path $p(b, T)$ would presumably care about the dynamic attributes of this power-path as well as its static attributes. For example, $m(b)$ might be concerned about equipment depreciation cost incurred during $T$ from ramping wear and tear, and the fuel costs incurred during $T$ for power production.

Moreover, what each customer $c_j(b)$ would presumably value in advance of $T$ is a guaranteed ability to determine their power withdrawals at $b$ during $T$ in a flexible just-in-time manner to run their personally-owned electrical devices for locally-determined purposes. The value they would attach in advance of $T$ to any one pre-specified power-path $p(b, T)$ would presumably be low, simply because of its inflexibility.

Finally, what the RTO/ISO would presumably value in advance of $T$ is the guaranteed availability of a collection of suitably diverse RTO/ISO-dispatchable power-paths enabling the RTO/ISO to balance uncertain net-power withdrawal at $b$ and other grid locations during $T$ by suitable just-in-time dispatched net-power injections. The value that the RTO/ISO would attach in advance of $T$ to the availability of any one specified power-path $p(b, T)$ would presumably be low, simply due to its inflexibility.
The key implication of the above observations is that the benefits and costs of producers, customers, and the RTO/ISO itself in RTO/ISO-managed wholesale power markets cannot properly be assessed solely in terms of transacted energy amounts.

3.3 Presumption (P2): Two-Part Partition for Total Supplier Cost

3.3.1 (P2): Conceptual Concern Overview

Presumption (P2) has two troublesome aspects. First, it reflects the traditional economic overly-simplistic partition of total supplier cost into only two parts: fixed and variable. Second, its focus on delivered energy as the sole determinant of supplier variable cost is based on the conceptually-problematic presumption (P1). These two aspects of (P2) are separately addressed in the next two sub-sections.

3.3.2 Needed Partitioning of Total Cost Into Three Components

For reasons carefully articulated in seminal work by Baumol et al. [3], the traditional economic partition of total cost into two components, “fixed” and “variable,” is conceptually incomplete and empirically problematic for many U.S. industries. Rather, total cost should be partitioned into three economically-distinct types of costs:

\[
Total\ Cost = Sunk\ Cost + Avoidable\ Fixed\ Cost + Variable\ Cost \tag{2}
\]

As will be stressed in subsequent sections of this study, the need for the three-part partition (2) is particularly critical for the conceptually-coherent design and operation of grid-supported centrally-managed wholesale power markets.

Consider a Decision-Maker (DM) at a current time \( t \) who must decide now whether or not to commit to undertaking an action of type \( A \) at a future time \( t + \Delta t \). The DM’s total cost at time \( t \) can be partitioned into three components – sunk cost, avoidable fixed cost, and variable cost – as follows:

\[
Sunk\ Cost =: Non-Avoidable\ Fixed\ Cost \tag{3}
=:\ Cost\ C^o\ that:\
\begin{align*}
(i) & \text{ the DM incurs whether or not the DM commits at time } t \\
& \text{ to undertaking a type-}A \text{ action at time } t + \Delta t; \\
(ii) & \text{ does not depend on the specific form of type-}A \text{ action }
\end{align*}
\]

\[
Avoidable\ Fixed\ Cost =:\ Cost\ C^o\ that:\
\begin{align*}
(i) & \text{ the DM incurs if and only if the DM commits at } \\
& \text{ time } t \text{ to undertaking a type-}A \text{ action at time } t + \Delta t; \\
(ii) & \text{ does not depend on specific form of type-}A \text{ action.}
\end{align*}
\]
**Variable Cost** \( = \): \( \text{Cost } C(a) \) that:

(i) the DM incurs **if and only if** the DM commits at time \( t \) to undertaking a type-A action at time \( t + \Delta t \);

(ii) **does** depend on specific form \( a \) of type-A action.

\[ \text{Avoidable Cost} = [\text{Avoidable Fixed Cost}] + [\text{Variable Cost}] \] (6)

For illustration, consider a **currently off-line** dispatchable thermal generation unit \( m \) at the start-time \( t \) of a day-D U.S. RTO/ISO-managed DAM that the RTO/ISO is conducting to prepare for grid operations during day D+1. An example of a **sunk cost** for \( m \) would be an amount of money that \( m \) previously spent to purchase a piece of generation equipment that now has no resale value. An example of an **avoidable fixed cost** for \( m \) would be the start-up cost that \( m \) would have to incur in order to transition from its currently off-line state to a synchronized state\(^{12}\) by the start of day D+1 if \( m \) submits an offer into the day-D DAM to provide positive power injection during Hour 1 of day D+1 and the RTO/ISO clears this offer.\(^{13}\) An example of a **variable cost** for \( m \) would be the fuel cost \( VC_m(p,H,D+1) \) that \( m \) would have to incur in order to maintain a specific positive power-injection level \( p \) during some hour \( H \) of day D+1.

Consider, once again, the decision problem articulated at the beginning of this subsection: namely, a decision-maker DM at a current time \( t \) must decide **now** whether or not to commit to undertaking an action of type A at a **future** time \( t + \Delta t \). We conclude this sub-section by expressing several decision principles for DM in terms of the three-part partition (2) for total cost.

By definition, the DM’s sunk cost at time \( t \) is unavoidable, hence incurred whether or not the DM agrees to the commitment at time \( t \). The DM’s sunk cost at time \( t \) should therefore play no role in the DM’s time-\( t \) commitment decision.

**“Sunk Cost is Sunk” Dictum:**

A decision-maker at time \( t \) who is required to make a decision at time \( t \) should make this decision independently of his sunk cost at time \( t \).

Define the DM to be **risk-averse**\(^{14}\) if the DM is **not** willing to participate:

- in any **risky** undertaking with **zero expected payoff**, where the qualifier “risky” means there is some dispersion of positive-probability payoffs **around** the zero expected payoff, hence a positive probability of a negative payoff;
- in **any** undertaking (risky or certain) that has a **strictly negative expected payoff**.

---

12 A thermal generator is said to be in a **synchronized state** if it is an operating state that permits it to inject power into the grid, even if no such power injection is currently being undertaken.

13 Kirschen and Strbac [22] refer to start-up cost as a “quasi-fixed cost”.

14 In economic theory, survival is often guaranteed **a priori**, e.g., by postulating zero subsistence needs. The risk aversion of a decision-maker DM is then characterized as a preference attribute of DM: namely, the degree to which the DM’s utility function expressing preference orderings over possible payoffs exhibits concave curvature properties. In reality, strictly negative payoffs can pose grave **survival risks** for any person or commercial entity without deep financial pockets that simply wishes to avoid starvation or insolvency, inducing them to behave in a “risk averse” manner.
If the DM is risk-averse, the DM should agree to the commitment at time $t$ if the DM believes this would result for sure in a strictly positive net benefit, where:

$$\text{net benefit} = \text{benefit} - \text{avoidable cost} \quad (7)$$

Note a strictly positive net benefit (7) would permit the DM to pay down at least part of his time-$t$ sunk cost, assuming this sunk cost is positive.

Conversely, a risk-averse DM should agree to the commitment at time $t$ only if the DM’s expected net benefit from this commitment is non-negative, where net benefit is again defined as in (7). Otherwise, commitment at time $t$ would be an agreement to participate in an undertaking with a strictly negative expected payoff.

These observations are summarized in the form of two commitment principles:

**Commitment Principles:** A risk-averse DM should agree to commit at time $t$

- if the DM believes this commitment would result for sure in a net benefit (7) for himself that is strictly positive;
- only if the DM believes his expected net benefit from this commitment is non-negative, where net benefit is again defined as in (7).

### 3.3.3 Variable Cost Compensation in U.S. RTO/ISO-managed DAM/RTMs

Multiple types of variable costs incurred by power resources participating in U.S. RTO/ISO-managed wholesale power markets are listed in Appendix A.4. In addition to the commonly-considered fuel-cost category, the list includes: labor cost; intermediate good (supply chain) cost; equipment/software rental cost; equipment depreciation cost; transmission service charges; variable-cost offsets for sales of valuable bi-products; and disposal costs for waste bi-products.

The key concern raised in this sub-section is whether the standardized forms of supply offers required in current RTO/ISO-managed wholesale power markets permit suppliers to receive appropriate conceptually-coherent compensation for their incurred variable costs. As will next be shown, these supply-offer forms force suppliers to express their variable costs as functions of delivered energy (MWh), with no consideration of dynamic power-path implementations. It is difficult to understand how any of the variable-cost categories listed in Appendix A.4 – including fuel cost – can be accurately measured and reported solely as a function of delivered energy, with no consideration of power-path implementation.

As reviewed in Section 2, the core design element for all seven U.S. RTO/ISO-managed wholesale power markets is a bid/offer-based DAM/RTM two-settlement system. In all seven RTOs/ISOs, the DAM SCED optimization\textsuperscript{15} conducted during the morning of each day D (conditional on given commitments for generation units) determines scheduled power dispatch set-points (MW) for each committed generation unit at the start of each hour H during the following day D+1.

|\textsuperscript{15} SCED optimizations for U.S. RTO/ISO-managed DAMs and RTM sub-markets are similar, apart from operating-period duration and restrictions on LSE submission of fixed demand bids. For simplicity of exposition, this sub-section focuses solely on DAM SCED optimizations. |
More precisely, apart from ISO New England, these scheduled dispatch set-points determine co-optimized scheduled maintained power levels or maintained power-slope levels (hence scheduled energy deliveries) and scheduled operating reserve (unencumbered generation capacity levels) for each committed dispatchable generation unit for each hour H of day D+1. These determinations are subject to system constraints that include nodal and/or zonal reserve requirements for operating reserve with different availability characteristics.

The supply offer that a dispatchable power resource (DPR) submits to a day-D RTO/ISO-managed DAM SCED optimization for each hour H of day D+1 is intended to convey variable cost information about this DPR to the RTO/ISO. As detailed in [35], these hourly supply offers are generally required to take a simple energy-block form: namely, a possibly-zero fixed maintained power level for hour H plus a small upper-limited number N of successive “energy blocks” i = 1, ..., N in the MW-$/MWh plane indicating the supplier’s requested per-unit compensation $/MWh for delivered energy (MWh) at successively-higher offered maintained power levels p_i (MW) for hour H. Starting at maintained power level p = 0 and continuing through some maintained power level p' > 0, the summation $/h of the finitely-many energy-block compensations required by the DPR at p' is considered to be an approximation of the DPR’s non-decreasing variable cost function for hour H, evaluated at p'.

For example, given N = 3 and 0 < p_1 < p_2 < p' ≤ p_3, there are three energy blocks to consider for the calculation of VC(p') ($/h): namely, the three energy blocks corresponding to the three successive power ranges (0, p_1], (p_1, p_2], and (p_2, p']. Let π_i ($/MWh) for i = 1, 2, 3 denote the DPR’s requested per-unit compensation for delivered energy for these three successive maintained-power ranges. Then:

\[ VC(p') = \pi_1 \cdot [p_1 - 0] + \pi_2 \cdot [p_2 - p_1] + \pi_3 \cdot [p' - p_2]. \quad (8) \]

Figure 3 depicts the SCED determination of optimal market-clearing (demand = supply) solutions for per-unit energy price ($/MWh) and energy transactions for a future hour H. These solutions are based on hour-H energy bids submitted in energy-block form by two energy buyers (B1, B2) and hour-H energy offers submitted in energy-block form by three suppliers (S2, S3, S3). As noted in Section 2, the market

---

16 ISO New England conducts Forward Reserve Market (FRM) auctions for 10-minute contingency reserve and 30-minute supplemental reserve separately from DAM energy scheduling; see, for example, [6, Tables 1-2] and [15, Table 1].

17 The types of operating reserve procured on a co-optimized basis with energy in U.S. RTOs/ISOs (apart from ISO New England) include Regulation, Spinning Reserve, and Supplemental Reserve. See [15, Sec. II.A & Appendix].

18 As detailed in [2], ERCOT permits Qualified Scheduling Entities (QSEs) to submit supply offers in a three-part form that allows inclusion of some unit-commitment cost information in addition to variable cost information. For example, a supply offer submitted to a day-D ERCOT DAM by a currently off-line QSE for some hour H during day D+1 typically consists of three parts: Startup Offer ($/start); Minimum-Energy Offer consisting of an energy price ($/MWh) and a Low-Sustained Limit (LSL) power level (MW); and a non-decreasing piecewise linear Energy Offer Curve in the MW-$/MWh plane consisting of a finite collection of linearly-connected power-price points (p, π) whose power levels p commence at the LSL level.
process depicted in Fig. 3 is an energy-market process, despite the appearance of power levels (MW) along the quantity axis. This follows because these power levels denote possible choices for the level of power (MW) to be maintained during hour H. Choosing a maintained power level \( p \) (MW) for hour H is equivalent to choosing an energy amount \( p \cdot 1h \) (MWh) to be delivered for hour H.

### 3.4 Presumption (P3): Energy (MWh) Functions as a Commodity

#### 3.4.1 (P3): Overview of Conceptual Concerns

Presumption (P3) implies that, within the context of a grid-supported U.S. RTO/ISO-managed wholesale power market, energy (MWh) conditional on grid delivery location and delivery time-period is a commodity whose perfectly substitutable units (MWh) should be bought and sold in a spot market at a competitively-determined uniform market price ($/MWh). The fundamental concern regarding presumption (P3) is its serious conceptual inconsistency with Counter-Claims 1–4 in Section 1 that are supported in detail in [36] and throughout the current study.

To explain and critique this conceptual inconsistency in a careful, clear, yet succinct manner, some economic background materials are essential. These background materials are presented in the next sub-section.

#### 3.4.2 (P3): Economic Background Materials

**Basic Economic Definitions:**

1. **(Def1)** An asset is anything in physical or financial form that can function as a store of value over time. An example of a physical asset is a battery, and an example of a financial asset is a stock share.

2. **(Def2)** A market for an asset is a spot market if transacted asset amounts, payments for these transacted asset amounts, and deliveries of these transacted asset amounts all occur at the same location and time (“on the spot”).

3. **(Def3)** A market for an asset is a forward market if transacted asset amounts and payment obligations for these transacted asset amounts are determined in advance of the delivery of these transacted asset amounts.

4. **(Def4)** Market revenue sufficiency holds for a market \( M \) if, for each supplier \( i \) participating in \( M \), the total market-determined revenue earned by supplier \( i \) from participation in \( M \) suffices over time to cover supplier \( i \)’s total avoidable cost incurred from participation in \( M \).

5. **(Def5)** A commodity \( Q \) is an asset with a standard unit of measurement such that, at any given location and time, the units of \( Q \) are perfect substitutes, i.e., the substitution of an available \( Q \)-unit for any other available \( Q \)-unit has no effect on any \( Q \)-trader’s valuation of this \( Q \)-unit.\(^{19}\)

---

\(^{19}\)The requirement that all units of a commodity \( Q \) available for receipt or supply at a given location and time must (by definition) be perfect substitutes has the following two important implications: (i) Any buyer \( j \) that is told they will receive one additional \( Q \)-unit at a given location and time
Law-of-One-Price for Commodity: At any given location and time, each unit $u$ of a commodity $Q$ should sell at the same unit price ($/u$).

Definition of a Competitive Commodity Spot Market (CCSM): A Competitive Commodity Spot Market (CCSM) is a commodity spot market CSM that satisfies the following six assumptions (A1)–(A6) for some commodity $Q$ with a standard unit of measurement $u$:

(A1) The participants in CSM consist of a fixed set of $Q$-buyers $j$ and a fixed set of $Q$-suppliers $i$.

(A2) Each buyer $j$ and supplier $i$ is a price-taker.

(A3) Each buyer $j$ has a non-increasing demand schedule $D_j(\pi)$ that maps non-negative per-unit $Q$-prices $\pi$ ($/u$) into $Q$-demands $q$ measured in $u$. Given any $(\pi', q')$ satisfying $D_j(\pi') = q'$ with $\pi' \geq 0$, the quantity $q'$ denotes the maximum $Q$-amount that buyer $j$ is willing to demand at per-unit $Q$-price $\pi'$. Conversely, given regularity conditions,22 the per-unit $Q$-price $\pi'$ is the marginal benefit ($/u$) that buyer $j$ should sell at the same unit price ($/u$).

Suppose, for example, that buyer $j$ attains from this receipt the same for all $Q$-units; this is a necessary condition for buyer $j$ to have a well-defined “marginal” benefit (utility) function for $Q$ at this given location and time. (ii) Any supplier $i$ that is told they must deliver one additional $Q$-unit at a given location and time is completely indifferent with regard to which precise additional $Q$-unit they deliver because the incremental (“marginal”) cost $\Delta\text{Cost}$, that supplier $i$ incurs for this delivery is the same for all $Q$-units; this is a necessary condition for supplier $i$ to have a well-defined “marginal” cost function for $Q$ at the given location and time.

See [36, Sec. 9.3.4] for a careful presentation and discussion of the standard economic definition of a CCSM, including key related concepts such as net surplus extraction and market efficiency.

A participant in a spot market for a commodity $Q$ is said to be a price-taker if the participant behaves as if his own market transactions have no effect on the determination of the market price for $Q$. One situation that could reasonably induce price-taking behavior by all market participants is if the maximum volume of purchase/sale transactions that each individual market participant could feasibly undertake is extremely small relative to the total volume of market transactions, and collusion among market participants is either not possible or not permitted.

These regularity conditions pertain to the ability to invert buyer $j$'s ordinary demand schedule $D_j(\pi)$ mapping per-unit $Q$-prices $\pi \geq 0$ into $Q$-demands $q \geq 0$ to obtain a well-defined inverse demand schedule mapping $Q$-demands $q \geq 0$ into per-unit $Q$-prices $\pi \geq 0$; see [36, Sec. 9.3.4]. Suppose, for example, that buyer $j$ has a total benefit function $B_j(q)$ mapping $Q$-demands $q \geq 0$ into attained real-valued total benefit levels $B_j(q)$ that is non-decreasing, differentiable, and concave over $q \geq 0$. Buyer $j$'s marginal benefit, evaluated at any $q' \geq 0$, is then the derivative of buyer $j$'s total benefit function $B_j(q)$ with respect to $q$, evaluated at $q = q'$; this mapping of $q' \geq 0$ into a marginal benefit evaluation $\partial B_j(q')/\partial q = MB_j(q') =: \pi' \geq 0$ is buyer $j$'s inverse demand schedule for $Q$. Finally, if buyer $j$'s marginal benefit function $MB_j(q)$ is a strictly decreasing function of $q$ for $q \geq 0$, a common “diminishing marginal return” assumption in economics, it can be inverted over $q \geq 0$ to give buyer $j$'s ordinary demand schedule for $Q$: that is, to give a strictly-decreasing function $D_j(\pi)$ mapping per-unit $Q$-prices $\pi \geq 0$ into $Q$-demands $q' = D_j(\pi) \geq 0$. In this case, by construction, the price $\pi'$ that maps into $q'$ is indeed the marginal benefit of buyer $j$ evaluated at...
that buyer \( j \) obtains from the last ("marginal") \( Q \)-unit that he demands at \( Q \)-demand level \( q' \), hence also buyer \( j \)’s purchase reservation value (maximum acceptable purchase price) for this last demanded \( Q \)-unit at \( q' \).

(A4) Each supplier \( i \) has a non-decreasing supply schedule \( S_i(\pi) \) that maps non-negative per-unit \( Q \)-prices \( \pi \)($/u$) into \( Q \)-supplies \( q \) measured in \( u \). Given any \((\pi', q')\) satisfying \( S_i(\pi') = q' \) with \( \pi' \geq 0 \), the quantity \( q' \) denotes the maximum \( Q \)-amount that supplier \( i \) is willing to sell at per-unit \( Q \)-price \( \pi' \). Conversely, given regularity conditions,\(^{23}\) the per-unit \( Q \)-price \( \pi' \) is the marginal cost ($/u$) that supplier \( i \) incurs for the last ("marginal") \( Q \)-unit that he sells at \( Q \)-supply level \( q' \), hence also supplier \( i \)’s sale reservation value (minimum acceptable sale price) for this last \( Q \)-unit at \( q' \).

(A5) The price-rule for CSM is competitive-equilibrium pricing.\(^{24}\)

(A6) All fixed cost for each supplier \( i \) participating in CSM is sunk cost.\(^{25}\)

The five CCSM-assumptions (A1)–(A5) directly imply that a CCSM has the following important property:

demand level \( q' \). Economists studying competitive markets typically work with ordinary demand schedules mapping prices into quantities because, as seen in assumption (A2), all buyer participants in such markets are assumed to be price-takers. However, in U.S. RTO/ISO-managed wholesale power markets, demand schedules ("demand bids") are typically expressed in inverse form, as mappings from quantities into prices.

\(^{23}\) These regularity conditions pertain to the ability to invert supplier \( i \)'s ordinary supply schedule \( S_i(\pi) \) mapping per-unit \( Q \)-prices \( \pi \) \( \geq 0 \) into \( Q \)-supplies \( q \) \( \geq 0 \) to obtain a well-defined inverse supply schedule mapping \( Q \)-supplies \( q \) \( \geq 0 \) into per-unit \( Q \)-prices \( \pi \) \( \geq 0 \); see [36, Sec. 8.2]. Suppose, for example, that supplier \( i \) has a total cost function \( C_i(q) \) mapping \( Q \)-supplies \( q \) \( \geq 0 \) into incurred real-valued total cost levels \( C_i(q) \) ($\) that is non-decreasing, differentiable, and convex over \( q \) \( \geq 0 \). Supplier \( i \)'s marginal cost, evaluated at any \( q' \) \( > 0 \), is then the derivative of supplier \( i \)'s total cost function \( C_i(q) \) with respect to \( q \) \( \geq 0 \), evaluated at \( q = q' \); this mapping of \( q' \) \( \geq 0 \) into a marginal cost evaluation \( \partial C_i(q')/\partial q =: MC_i(q') =: \pi' \geq 0 \) is supplier \( i \)'s inverse supply schedule for \( Q \). Finally, if supplier \( i \)'s marginal cost function \( MC_i(q) \) is a strictly increasing function of \( q \) for \( q \geq 0 \), a common "increasing marginal cost" assumption in economics, it can be inverted over the range \( q \geq 0 \) to give supplier \( i \)'s ordinary supply schedule for \( Q \) that is, to give a strictly-increasing function \( S_i(\pi) \) mapping per-unit \( Q \)-prices \( \pi' \) \( \geq 0 \) into \( Q \)-supplies \( q' = S_i(\pi') \) \( \geq 0 \). In this case, by construction, the price \( \pi' \) that maps into \( q' \) is indeed the marginal cost of supplier \( i \) evaluated at supply level \( q' \). Economists studying competitive markets typically work with ordinary supply schedules mapping prices into quantities because, as seen in assumption (A2), all supplier participants in such markets are assumed to be price-takers. However, in U.S. RTO/ISO-managed wholesale power markets, supply schedules ("supply offers") are typically expressed in inverse form, as mappings from quantities into prices.

\(^{24}\) A price-quantity pair \((\pi', q')\) for CSM is a competitive equilibrium for CSM if it is an intersection point of the aggregate demand schedule \( q = D(\pi) \) and the aggregate supply schedule \( q = S(\pi) \) plotted in the \((\pi, q)\) plane (with horizontal and vertical flat segments included), where \( D(\pi) := \sum_j D_j(\pi) \) and \( S(\pi) := \sum_i S_i(\pi) \).

\(^{25}\) By the first five assumptions (A1)–(A5), the CCSM takes place at a given location and time for a given set of participants whose demand and supply schedules are automatically submitted to the CCSM and instantly cleared (or not cleared) to determine competitive equilibrium outcomes. Thus, no supplier participating in the CCSM is a decision-maker able to avoid (or not avoid) some cost depending on a decision the supplier makes at this given location and time.
CCSM1: A CCSM is a uniform-price market in the following sense. At any competitive equilibrium $e^* = (\pi^*, q^*)$, the same per-unit $Q$-price $\pi^*$ ($$/u$) is: (a) paid by each inframarginal (cleared) buyer $j$ for each $Q$-unit that buyer $j$ purchases; and (b) received by each inframarginal (cleared) supplier $i$ for each $Q$-unit that supplier $i$ sells.

In addition, the five CCSM-assumptions (A1)–(A5) imply that, at any competitive equilibrium $e^* = (\pi^*, q^*)$, the total (possibly zero) revenue earned by each supplier $i$ suffices to cover the sum (or integral) of the non-negative marginal costs that supplier $i$ incurs at $e^*$. According to the supplier cost definitions given in Section 3.3, this sum (or integral) of marginal costs constitutes the total variable cost for supplier $i$ at $e^*$. Moreover, CCSM-assumption (A6) implies that all fixed cost for supplier $i$ at $e^*$ is sunk cost, i.e., unavoidable fixed cost. Thus, the total avoidable cost of supplier $i$ at $e^*$ coincides with supplier $i$’s total variable cost.

Consequently, the six CCSM-assumptions (A1)–(A6) together imply that a CCSM has a second important property:

CCSM2: Market revenue sufficiency (Def4) holds for a CCSM. That is, at any competitive equilibrium $e^* = (\pi^*, q^*)$, the total revenue earned by each supplier $i$ suffices to cover supplier $i$’s total avoidable cost.

As carefully shown in Tesfatsion [36, Ch. 12], the six CCSM-assumptions (A1)–(A6) also imply that a CCSM has a third important property:

CCSM3: Total net surplus (TNS)$^{27}$ at a competitive equilibrium $e^* = (\pi^*, q^*)$ for a CCSM provides the largest possible total net surplus that buyers and suppliers can extract from the underlying commodity spot market CSM. Thus, market efficiency holds for CSM at $e^*$, in the following sense: At $e^*$ there is no wastage of opportunity to extract additional net surplus from CSM.

Property CCSM3 asserts that the six CCSM-assumptions (A1)–(A6) are sufficient to ensure market efficiency holds for the underlying commodity spot market CSM at any competitive equilibrium. However, these six CCSM-assumptions are not necessary for the market efficiency of the underlying CSM.

For example, it can be shown that market efficiency holds for the underlying CSM if the competitive-equilibrium price-rule (A5) is replaced by any price-rule PR that satisfies the following three price-rule conditions:

---

$^{26}$ See Section 3.3 (especially Footnote 7) and Tesfatsion [36, Ch. 12].

$^{27}$ By definition, total net surplus attained at any point $e = (\pi, q)$ is the sum of total net buyer surplus and total net supplier surplus attained at $e$. Total net buyer surplus at $e$ is the difference between the maximum amount that buyers would have been willing to pay for $q$ and the amount that they actually pay for purchase of $q$ at $e$. Total net supplier surplus at $e$ is the difference between the payment that suppliers actually receive for sale of $q$ at $e$ and the minimum payment that they would have been willing to receive for sale of $q$. 
PRC(a): All \( Q \)-units traded in CSM under price-rule PR also trade at some competitive equilibrium \( e^* = (\pi^*, q^*) \) for CSM.

PRC(b): All \( Q \)-units failing to trade in CSM under price-rule PR also fail to trade at the competitive equilibrium \( e^* = (\pi^*, q^*) \).

PRC(c): Under price-rule PR, the price paid by a buyer \( j \) to purchase a unit of \( Q \) is the same as the price received by the supplier \( i \) who supplies this unit of \( Q \).

An example of a price-rule PR for CSM that satisfies the price-rule conditions PRC(a)–PRC(c), distinct from (A5), is the \( k \)-discriminatory-price rule \( (k \in [0, 1]) \) defined as follows: For any matched buyer-supplier pair for which the buyer’s purchase reservation value \( \pi^b \) and the supplier’s sale reservation value \( \pi^s \) satisfy \( \pi^b \geq \pi^s \), set the strike price for this pair at the weighted-average level \( \pi^k =: k\pi^b + [1 - k]\pi^s \) lying between their reservation values. Thus, the division between buyer and supplier of the net surplus increment \( [\pi^b - \pi^s] \) resulting from their trade is determined by \( k \); however, the total amount of this net surplus increment is not affected by \( k \).

A common intuitive argument given in favor of the use of a competitive uniform-price rule and against the use of a \( k \)-discriminatory-price rule for a CSM is that the uniform-price rule provides more incentive to suppliers (buyers) of a commodity \( Q \) to use the most efficient available technology for extraction of net supplier (buyer) surplus. For example, suppose a supplier switches to a new technology that strictly lowers his marginal cost of production (hence his sale reservation value \( \pi^s \)) for each strictly inframarginal unit of \( Q \) he sells. The uniform-price rule permits this supplier to keep all of his resulting increased net supplier surplus; the \( k \)-discriminatory-price rule with \( k < 1 \) does not. An analogous argument holds for a buyer able to switch to a new technology that permits him to increase his marginal benefit (hence his purchase reservation value \( \pi^b \)) for each strictly inframarginal unit of \( Q \) he buys.

However, under either price-rule, a welfare-maximizing supplier or buyer will not switch to a more efficient technology unless the cost of this switching is less than the expected net surplus gain from future CSM transactions. Yet low or costless technology switching could strongly deter anyone from engaging in costly research and development (R&D) efforts needed to develop more efficient technologies. Clearly,

---

28 Condition PRC(c) essentially holds for U.S. RTO/ISO-managed DAM/RTM SCED optimizations in the absence of transmission grid congestion because SCED-determined bus LMPs then collapse to a single uniform LMP across the grid. Conversely, when the grid is congested (i.e., when at least one transmission-line capacity constraint is active), at least some LMP separation occurs across the buses of the grid. However, since power injected or withdrawn at any one grid location rapidly affects power flow on all directly or indirectly connected transmission lines, power injections and withdrawals do not in fact constitute a collection of bilateral buyer-supplier trades as presumed in the statement of condition PRC(c).

29 The two extremes of the \( k \)-discriminatory-price rule are of special interest: The 0-discriminatory-price rule awards all generated net surplus to buyers because the price received by suppliers is their minimum acceptable sale price \( \pi^s \). Conversely, the 1-discriminatory-price rule awards all generated net surplus to suppliers because the price charged to buyers is their maximum acceptable purchase price \( \pi^b \).
a *dynamic* joint analysis of market and R&D processes is needed to examine with care the long-run efficiency implications of alternative market price-rules.

A second intuitive argument given *in favor of* the use of a competitive uniform-price rule and *against* the use of a $k$-discriminatory-price rule for a CSM concerns incentives for truthful revelation. Under a $k$-discriminatory-price rule with $k < 1$, a self-interested supplier $i$ would have a strategic incentive to report a higher-than-true supply schedule for his strictly inframarginal $Q$-units in order to receive a higher discriminatory price for these units, thus increasing his true net revenue (i.e., his revenue minus his true variable cost) from the sale of these units. This incentive only disappears when the supplier has shifted up his reported supply schedule to a point that the discriminatory price assigned to each of his strictly inframarginal $Q$-units equals the discriminatory price assigned to the marginal (last) sold $Q$-unit.

In summary, a *competitive commodity spot market (CCSM)* has a number of attractive efficiency and optimality properties. However, the following important caution *must* be kept carefully in mind:

**CCSM4: Important Caution.** In order for CCSM-assumptions (A1)–(A6) and derived CCSM-properties (CCSM1)–(CCSM3) to be *conceptually meaningful*, the transacted asset *must be a commodity* as defined by (Def5). That is, the transacted asset *must have a standard unit of measurement such that, at any given location and time, all units of this asset are perfect substitutes.*

To see the need for the cautionary statement CCSM4, suppose the requirement that the transacted asset be a commodity is simply omitted from the definition given above for a competitive commodity spot market (CCSM). By definition (Def2) for a spot market, the resulting competitive spot market must still take place at a given location and time. Moreover, assumptions (A3)–(A5) depend strongly on the concept of a “unit price” for the transacted asset, implying this transacted asset must have a standard unit of measurement. Thus, as seen in definition (Def4), the only way the transacted asset can fail to be a commodity at a given location and time is if its units are not perfect substitutes at this given location and time.

*However, the standard economic definitions given in assumptions (A3) and (A4) for ordinary demand and supply schedules are not conceptually coherent for an asset $A$ whose units are not perfect substitutes.* At any given location and time, how are the units of $A$ to be ordered along the quantity axis for a buyer’s ordinary demand schedule or along the quantity axis for a supplier’s ordinary supply schedule if these units are not perfect substitutes?

To cast this conceptual issue in more concrete terms, consider an experimenter who wishes to use a large sealed fruit-bag containing a mixture of HoneyCrisp Apples and Dole Mandarin Oranges to elicit an ordinary “fruit demand schedule” from a human fruit-buyer $B$ at a given experimental location and time. In accordance with (A3), the experimenter hands $B$ an ordered list of successively higher fruit-unit prices $\pi$ expressed as dollars per fruit-piece ($$/fp$) and asks $B$ to report the maximum fruit-amount $q = D(\pi)$ (measured in number of fruit-pieces) that $B$ would be willing to buy at each listed fruit-unit price $\pi ($$$/fp$). At the end of the experiment, one of the listed fruit-unit prices $\pi^*$ will be randomly selected, the fruit-bag will be unsealed,
and \( B \) will be required to pay the fruit-unit price \( \pi^* \) for \( q^* = D(\pi^*) \) fruit-pieces that the experimenter draws randomly from the unsealed fruit-bag.

Unfortunately for the experimenter, suppose \( B \) does not view a HoneyCrisp Apple as a perfect substitute for a Dole Mandarin Orange; that is, suppose the specific apple-versus-orange attribute of a fruit-piece matters to \( B \)? In this case, the value that \( B \) assigns to any fruit-amount \( q \) (measured in number of fruit-pieces) depends on the specific apple-orange composition of \( q \). Consequently, the maximum fruit-amount \( q \) that \( B \) is willing to buy at each listed fruit-unit price \( \pi \) (\$/fp) depends on how \( B \) resolves his uncertainty regarding the composition of fruit-pieces in the sealed fruit-bag, hence his uncertainty regarding the specific apple-orange composition of the fruit-amount \( q \) that the experimenter would randomly draw from the unsealed fruit-bag if \( B \) reports that \( q = D(\pi) \) is the maximum fruit-amount that \( B \) would be willing to buy at the listed fruit-unit price \( \pi \).

The bottom line is that an ordinary fruit demand schedule cannot be constructed for a fruit-buyer \( B \) in accordance with the standard economic definition (A3) if \( B \) does not view fruit-units as perfect substitutes. Analogous arguments can be used to show that an ordinary fruit supply schedule cannot be constructed for a fruit-supplier \( S \) in accordance with the standard economic definition (A4) if \( S \) does not view fruit-units as perfect substitutes.

What about inverse demand and supply schedules for fruit, defined in accordance with the standard economic definitions given in footnotes 22 and 23? Here the need for the cautionary statement CCSM4 is even simpler to see.

- Suppose fruit-units (pieces of fruit) are not perfect substitutes for fruit buyer \( B \). How can \( B \) express his maximum acceptable purchase price – equivalently, his attained marginal benefit – for a “next” purchased fruit-unit, given \( B \) has already purchased a fruit-amount \( q \) (measured in fruit-units), without knowing: (i) which specific fruit-unit, apple or orange, is to be this “next” purchased fruit-unit; and (ii) what is the specific apple-orange composition of \( B \)’s already-purchased fruit-amount \( q \)?
- Suppose fruit-units are not perfect substitutes for a fruit supplier \( S \). How can \( S \) express his minimum acceptable sale price – equivalently, his incurred marginal cost – for a “next” sold fruit-unit, given \( S \) has already sold a fruit-amount \( q \) (measured in fruit-units), without knowing: (i) which specific fruit-unit, apple or orange, is to be this “next” sold fruit-unit; and (ii) what is the specific apple-orange composition of \( S \)’s already-sold fruit-amount \( q \)?

3.4.3 (P3): First Fundamental Conceptual Concern

This sub-section argues that, in contradiction to presumption (P3), energy (MWh) does not function as a commodity within the context of current U.S. RTO/ISO-managed wholesale power markets. More precisely, a participant in such a market typically does not view a MWh of energy to be a perfect substitute for any other MWh of energy, conditional on delivery location and delivery time-period.

To the contrary, power producers and power customers typically care about the dynamic attributes of the power-paths they use to inject/withdraw power at their
grid locations during successive operating periods. For example, power producers dispatched to inject power at a grid location \( b \) during an operating period \( T \) might reasonably care about equipment wear-and-tear cost incurred due to the fast ramping required to follow received dispatch set-points. And power customers electrically connected to \( b \) might reasonably care about the degree of flexibility they have to meet their diverse power requirements during \( T \) by just-in-time determined power withdrawals at \( b \). Finally, RTOs/ISOs also typically care about power-paths, not energy deliveries. Given any future operating period \( T \), an RTO/ISO needs availability of dispatchable power-paths with diverse dynamic attributes for possible just-in-time RTO/ISO dispatch during \( T \) to ensure continual net-load balancing.

Thus, if a producer, customer, or RTO/ISO were asked to assign a monetary value to a specific amount of energy \( E^* \) (MWh) to be delivered at a specific grid location \( b \) during a specific future operating period \( T \), typically they would not be able to do so without knowing the power-path to be used for this delivery. For example, what value would each of these entities assign to each of the following four power-path options \( p = \{ p(t) \mid t \in T \} \) for the injection of power (MW) at a specific grid location \( b \) during a specific future 24-hour operating period \( T \), where each power-path option accumulates to the same amount \( E^* = 12 \) (MWh) of delivered energy:

**Power-Path (a):** Power (MW) is injected at \( b \) at level \( p = 24 \) during the first half-hour of \( T \) and at level \( p = 0 \) during the remainder of \( T \).

**Power-Path (b):** Power (MW) is injected at \( b \) at level \( p = 0.5 \) throughout \( T \).

**Power-Path (c):** Power (MW) is injected at \( b \) at level \( p = 1 \) during every other hour of \( T \), with \( p = 0 \) during the remaining hours of \( T \).

**Power-Path (d):** Power (MW) is injected at \( b \) during \( T \) in a flexible manner that is entirely up to the entity, apart from the requirement that the resulting total energy delivery must equal \( E^* = 12 \) (MWh).

Note that the flexible power-path option (d) would presumably be assigned a higher value than the rigid options (a) through (c). Indeed, option (d) would presumably be assigned at least as high a value as any other power-path option for power injection/withdrawal at \( b \) during \( T \) subject only to the delivered-energy requirement \( E^* = 12 \) (MWh) since option (d) encompasses all such options.

The clear implication of these examples is that energy (MWh) conditional on delivery location and delivery time-period does not function as a commodity within the context of current U.S. RTO/ISO-managed wholesale power markets. In consequence, recalling the important caution CCSM4 carefully justified at the end of Section 3.4.2, the following four basic economic constructions are conceptually ill-defined for such markets, conditional on energy delivery location and energy delivery time-period: namely, energy demand schedule; energy supply schedule; energy marginal cost function; and energy marginal benefit function.

All attempts to justify the DAM/RTM two-settlement system (based on LMP pricing) at the core of current U.S. RTO/ISO-managed wholesale power markets by means of efficiency and optimality arguments established for competitive commodity spot markets (based on marginal-cost pricing) are thus conceptually unsupportable. However, a great many products successfully transacted in the real world are...
not commodities. Indeed, product innovation occurs through continual striving to produce differentiated products permitting at least temporary market advantages.

The next section stresses that grid-supported centrally-managed wholesale power markets are necessarily forward markets. The key to the conceptually-consistent design of such markets is to consider with care how real-world forward markets transact products whose units are not viewed as perfect substitutes by market participants. As will be seen in Section 5, the short answer is appropriate contract design.

3.4.4 (P3): Second Fundamental Conceptual Concern

A second fundamental conceptual concern regarding presumption (P3) is that a market process $M(T)$ conducted within a grid-supported RTO/ISO-managed wholesale power market for a future operating period $T$ cannot be a spot market due to the speed of real-time grid operations. That is, $M(T)$ cannot coincide with $T$.

Indeed, the DAM/RTM two-settlement system at the core of each U.S. RTO/ISO-managed wholesale power market is a collection of grid-supported forward markets $M(T)$ with positive-duration look-ahead horizons $LAH(T)$; see Fig. 6. The scheduled generation-unit commitments and dispatch set-points for hour $H$ of day $D+1$, determined in the day-D DAM, are subject to change in supplementary unit-commitment processes and in RTM sub-markets held between the close of the day-D DAM and the start of hour $H$ on day $D+1$.

Nevertheless, the settlements for these scheduled next-day unit commitments, generation levels, and operating reserve levels are determined at the end of day $D$ as if they were actual spot-market transactions carried out on day $D$; see Fig. 2. This pay-for-performance in advance of actual performance typically results in time-inconsistent settlements, i.e., settlements determined and assigned to resources on day $D$ for unit commitments, energy levels, and operating reserve levels scheduled for day $D+1$ that are subsequently adjusted by OOM and RTM LMP payments due to discrepancies that arise between scheduled and actual outcomes on day $D+1$.

3.5 Presumption (P4): Spot Markets Ensure Fixed-Cost Coverage

Presumption (P4) implies that the revenues attained by suppliers participating in a Competitive Commodity Spot Market (CCSM) ensure coverage of their fixed cost.

Presumption (P4) is false for a CCSM. Assumption (A6) given in Section 3.4.2 for any CCSM implies that all suppliers participating in a CCSM have zero avoidable fixed cost; that is, all of their fixed cost is sunk cost, not avoidable cost. However, there is nothing in the CCSM-assumptions (A1) through (A6) that guarantees suppliers will be able to cover any or all of their sunk costs.

---

30 These settlements include out-of-market (OOM) make-whole payments for partial reimbursement of avoidable fixed costs (e.g., start-up costs) for committed generation units plus DAM SCED-determined nodal (dual-variable based) price payments intended to cover the variable costs for scheduled energy and reserve provision.

31 A multi-stage optimization problem that jointly determines an optimal solution $(s_0^*, s_1^*, \ldots, s_N^*)$ for successive time-periods $(s_0, s_1, \ldots, s_N)$ is said to be time-inconsistent if re-optimization undertaken at the beginning of some later time-period $s_n$ with $0 < n \leq N$ results in an optimal solution for $(s_n, \ldots, s_N)$ that deviates from $(s_0^*, s_1^*, \ldots, s_N^*)$. See [36, Sec. 10.2].
For example, by derived property CCSM3 in Section 3.4.2, note that market efficiency holds at any competitive equilibrium point $e^* = (\pi^*, q^*)$ for a CCSM in the sense that total net surplus at $e^*$ is as large as possible for the underlying commodity spot market CSM. However, referring to the precise definition of market efficiency given in Footnote 27, it is seen that market efficiency does not guarantee that the revenues received by participating suppliers cover any of their sunk costs.

This is a special case of a broader economic fact: There is no economic efficiency justification for instituting market rules that ensure suppliers are reimbursed for sunk costs. As stressed by the “Sunk Cost is Sunk” dictum in Section 3.3.2, sunk costs are already incurred costs (using up of resources) that suppliers cannot avoid by current or future decisions; hence, sunk costs should have no effect on these decisions.

Indeed, the only way a supplier $i$ can ensure coverage of some or all of his sunk cost through a market participation is if he has some type of structural or strategic market advantage relative to other actual or potential suppliers that reduces or eliminates the ability of these other suppliers to compete for supplier $i$’s customers if supplier $i$ attempts to charge these customers for sunk costs. Examples of situations giving rise to supplier market advantage include:

- **Regulatory Protection (Entry Barrier):** A supplier might have patent protection for his product that prevents other suppliers from producing this same product.
- **Product Differentiation:** A supplier’s product might have a special attribute (e.g., sale location, flavor based on secret recipe), highly valued by buyers, that differentiates it from all other products currently being supplied in the market and that is hard (or impossible) for other suppliers to copy;
- **Supply-Capacity Constraints:** A supplier might be a monopolist (sole supplier) with respect to “residual demand” customers that other capacity-constrained suppliers are unable to service, enabling him to include “extra” charges in the product price he sets for these customers for coverage of his sunk cost.

Conversely, buyers participating in a “perfectly contestable” market have no incentive to compensate participant suppliers for sunk costs. Roughly defined, a market is said to be perfectly contestable if any participating supplier charging a product price that strictly exceeds his avoidable cost of production can be successfully challenged and replaced by an existing or newly entering rival supplier able to charge a lower price for the same product. See Baumol et al. [3].

What about market revenue coverage of supplier avoidable fixed cost? As seen in Section 3.4.2, suppliers participating in spot markets have no avoidable fixed cost. In contrast, avoidable fixed cost essentially always arises for suppliers participating in forward market settings because avoidable fixed cost includes opportunity cost, i.e., earnings foregone by not committing assets to an alternative next-best use.

As discussed more carefully in [31], structural market advantage refers to an instituted feature of a market that systematically favors some market participants over others. In contrast, strategic market advantage is an opportunity available to a market participant to influence market outcomes in their favor in an officially unintended manner through some behavioral means. The standard economic term for these types of market advantages is “market power.” However, the use of “market power” in studies of electric power markets could cause confusion.
As discussed in Section 3.4.4, market processes conducted within the context of RTO/ISO-managed wholesale power markets operating over high-voltage AC transmission grids are necessarily *forward* markets, not spot markets, due to the speed of real-time grid operations. Various types of avoidable fixed cost that arise for suppliers participating in U.S. RTO/ISO-managed wholesale power markets are listed in Appendix A.4. This list includes: capital investment cost; transaction cost; opportunity cost, and unit commitment cost. Specific examples are given for each type of avoidable fixed cost. It is difficult to conceive how these avoidable fixed costs could be expressed in an empirically credible manner as functions solely of delivered energy amounts. Yet, this is what would be needed in order for suppliers participating in RTO/ISO-managed DAMs/RTMs to be assured coverage of their avoidable fixed costs solely through some form of extended LMP pricing mechanism.

Fortunately, this is not necessary. Forward markets instituted in other industries routinely rely on two-part pricing contracts to permit suppliers to receive appropriate separate compensation for avoidable fixed cost and variable cost. Section 5 of this study proposes an alternative Linked Swing-Contract Market Design for RTO/ISO-managed wholesale power markets that demonstrates how two-part pricing contracts could advantageously be introduced in these markets as well.

### 4 Legacy Core Design: Roadblock for Grid Decarbonization

#### 4.1 Overview

U.S. RTO/ISO-managed wholesale power markets are large complex organizations. From an external vantage point, the continued reliance of these markets on the legacy core DAM/RTM two-settlement system design reviewed in Section 2 appears to be greatly hindering these markets from transitioning smoothly to decarbonized grid operations. This section briefly discusses several external indicators in support of this concern; a more detailed discussion of these indicators is given in [36].

#### 4.2 Proliferation of Participation Models

The continued focus of U.S. RTO/ISO-managed Day-Ahead Markets (DAMs) and Real-Time Markets (RTMs) on energy as the key transacted product, as reflected in legacy economic presumption (P1), appears to be resulting in a proliferation of *participation models* functioning as artificial market entry barriers.33

More precisely, to participate in these DAMs/RTMs, a power resource must be classified in accordance with a designated taxonomy of participation models, each with its own eligibility rules and performance requirements. At the top of this taxonomy are two categorizations: “Energy (MWh)” and “Operating Reserve (MW).” The latter category consists of various forms of unencumbered generation capacity (MW) distinguished by availability characteristics; see Section 3.3.3.

---

33 For example, see the complicated eligibility requirements that power resources must satisfy to participate in MISO DAM/RTM processes as variously defined types of “qualified resources,” covered in MISO’s 278-page Energy and Operating Reserve Markets manual [28].
However, this entire DAM/RTM taxonomy is conceptually problematic because, as discussed with care in [36, Sec. 3.2.1], “Energy” and “Operating Reserve” are not conceptually well-defined as independent participation categories for such markets. For example, consider an RTO/ISO-managed DAM SCUC/SCED optimization conducted on day D in order to co-optimize scheduled energy deliveries and scheduled operating reserve for day D+1. As usual, suppose the forecasted net fixed load-profile at a grid bus \( b \) (or for a grid zone \( z \)) for day D+1 functions as a centroid (i.e., mid-point anchor) for a nodal (or zonal) uncertainty set expressing RTO/ISO-specified operating reserve requirements for day D+1 at \( b \) (or for \( z \)). Then:

- *energy* levels determined in the day-D DAM SCUC/SCED optimization for scheduled dispatched delivery at bus \( b \) (or within zone \( z \)) for day D+1 consist of a collection of RTO/ISO-dispatchable power-paths deemed capable of covering the centroid of the uncertainty set for bus \( b \) (or zone \( z \)) during day D+1.

- *operating reserve* levels determined in the day-D DAM SCUC/SCED optimization for scheduled availability at bus \( b \) (or zone \( z \)) during day D+1 consist of a collection of RTO/ISO-dispatchable power-paths deemed capable of covering the remainder of the uncertainty set for bus \( b \) (or zone \( z \)) during day D+1.

Thus, there is no fundamental conceptual distinction between optimal scheduled energy dispatch levels and optimal scheduled operating reserve levels for day D+1. Both are collections of RTO/ISO-dispatchable power-paths for day D+1 whose purpose is to ensure the balancing of uncertain net fixed load during day D+1.

### 4.3 Proliferation of Flexibility Products

A related concern is the proliferation of flexibility products as supplemental support for real-time net-load balancing.

U.S. RTO/ISO-managed DAMs/RTMs have traditionally relied on operating reserve products taking the form of unencumbered generation capacity distinguished by speed of availability. At any given time, unencumbered generation capacity is the power production capability of generation units that are currently without scheduled dispatch obligations. Thus, operating reserve and scheduled generation dispatch are joint products.\(^{34}\)

Three types of operating reserve procured on a co-optimized basis with energy in U.S. RTO/ISO-managed DAMs (apart from ISO New England) are Regulation, Spinning Reserve, and Supplemental Reserve [15, Sec. II.A & Appendix]. Next-day hourly prices for each operating reserve type, derived as dual variable solutions for hourly reserve requirement constraints incorporated into the system constraints for SCED optimizations, take the form of foregone energy-price (LMP) payments.

---

\(^{34}\) In economics, two or more products are said to be joint products if their production results jointly from the use of common inputs in a production process. Operating reserve and scheduled generation dispatch are rival joint products, meaning that – for any given level of inputs (here generation capacity) – an increase in one of the joint products requires a decrease in the other. The use of a single fuel source to co-generate electricity and steam is an example of a non-rival joint-production process well known to engineers.
As noted in Section 1, net loads for U.S. RTO/ISO-managed wholesale power markets are expected to become increasingly uncertain and volatile as these markets transit to increased reliance on IPRs and more active demand-side participation. Indeed, net loads for CAISO and other U.S. RTOs/ISOs are already exhibiting more frequent and dramatic down/up ramping swings. Consequently, the current heavy reliance on unencumbered generation capacity (MW) for operating reserve is becoming increasingly risky.

Industry and academic researchers are thus exploring the possible introduction of new types of “flexibility products” to facilitate the balancing of more uncertain and volatile net loads. For example, FERC has approved proposed ramp flexibility products for CAISO, MISO, and SPP; see [15, p. 12].

A major concern regarding these developments is that the newly approved RAMP (MW/min) products, together with already instituted CAP (MW) and ENERGY (MWh) products, are not independently produced products that can be separately transacted at separately determined prices in a conceptually consistent manner. To the contrary, they are the correlated attributes of individual power-paths, hence joint products. See footnote 34.

The conceptually-problematic treatment of RAMP, CAP, and ENERGY as independently produced and priced products presumes the value of a power-path can be appropriately measured by means of separate prices assigned to its attributes, treated as independent products. In actuality, the attributes of a power-path \( p(T) = \{ p(t) \mid t \in T \} \) for an operating period \( T \) – such as power-delivery start-time \( t^* \in T \), power capacity (MW) profile, ramp-rate (MW/min) profile, power-factor (kW/kVA) profile, power-delivery duration \( \Delta t \), and total delivered energy (MWh) – are correlated jointly-produced attributes. A change in any one attribute of a power-path can necessitate changes in its other attributes.

In economics, hedonic pricing is the pricing of a product on the basis of prices separately assigned to its intrinsic physical attributes as well as to its external circumstances. In some situations it might be desirable to use a hedonic-price approximation for the variable cost \( \phi_m(p_m(T)) \) ($) that a dispatchable power resource \( m \) would incur for RTO/ISO-dispatched delivery during \( T \) of each of its offered power-paths \( p_m(T) \).

For example, the variable cost \( \phi_m(p_m(T)) \) might be approximated as a linear combination of contractually-agreed metric functions that separately assign costs ($) for post-T verified power capacity (MW) profile, ramp-rate (MW/min) profile, and total delivered energy (MWh), as follows:

\[
\phi_m(p_m(T)) \approx C_{\text{CAP}}(p_m(T)) + C_{\text{RAMP}}(p_m(T)) + C_{\text{ENERGY}}(p_m(T)).
\] (9)

However, it would be conceptually incorrect and highly problematic in practice to ignore that the CAP, RAMP, and ENERGY “products” whose costs are evaluated in (9) are in fact highly-correlated attributes of a single underlying power-path \( p_m(T) \).


4.4 Proliferation of Out-of-Market Make-Whole Payments

U.S. RTO/ISO business practice manuals provide detailed descriptions of business operations for stakeholders and other interested parties. The manuals that focus specifically on the trade and settlement of energy and operating reserve have become extremely complex over the years. Much of this growing complexity has arisen from the need to explain various types of out-of-market (OOM) payments that RTOs/ISOs have instituted for their supplier participants as supplements to their market-determined revenues in an attempt to ensure coverage of their incurred costs.

The names and definitions of these OOM payments are not standardized across the seven U.S. RTOs/ISOs, and their relationship to the OOM payments discussed in FERC Orders is not entirely clear. An essential aspect in need of clarification is the distinction between “uplift” OOM payments and “make-whole” OOM payments.

For example, FERC Order No. 844 [12, Sec. I.2] on “‘uplift transparency,” released in 2018, directs RTOs/ISOs to provide a more transparent monthly reporting of “uplift payments” characterized as follows:

“RTO/ISO markets can be affected by a number of operational challenges such as unplanned transmission and generation outages and the need to maintain adequate voltage throughout the system. Limitations in the ability of the market software to incorporate all reliability considerations can at times result in prices that fail to reflect some of these challenges. In such situations, certain resources needed to reliably serve load may not economically clear the market and RTOs/ISOs must take out-of-market actions (i.e., operator-initiated commitments) to ensure system needs are met. These actions give rise to uplift costs. ... Uplift payments reflect the portion of the cost of reliably serving load that is not included in market prices.” [pp. 4-7]

Thus, FERC Order 844 characterizes uplift payments as OOM reimbursements to power resources for undertaking RTO/ISO-requested OOM actions deemed necessary to maintain grid reliability.

In contrast, RTOs/ISOs have instituted various types of OOM make-whole payments for suppliers participating in DAMs/RTMs whose market-determined revenues fail to provide coverage for certain types of avoidable fixed cost. Examples include ERCOT’s payments for accumulated power usage (energy) required for start-up, and for the maintenance of an on-line state at a minimum possible maintained power-injection level; see Footnote 18.

A key concern regarding existing OOM make-whole payments is that they do not ensure supplier revenue sufficiency. For example, the OOM make-whole payments that ERCOT awards to suppliers for start-up and for the maintenance of an on-line state by no means provide full coverage for all supplier avoidable fixed costs; see the avoidable fixed cost examples listed in Appendix A.4.

On the other hand, various OOM make-whole payment methods proposed to ensure better coverage of supplier costs have tended to blur the operationally-critical distinctions among sunk cost, avoidable fixed cost, and variable cost discussed in Section 3.3. For example, the Notice Of Proposed Rule-making (NOPR) released by FERC in 2016 [9], and subsequently withdrawn by FERC in 2017 [10], would have required unit-commitment costs for fast-start resources to be incorporated into the energy and operating reserve prices determined in co-optimized DAMs/RTMs. As
reviewed by Hartman [18, pp. 6-7], the original NOPR release encouraged commentators to suggest that unit-commitment costs for other types of generation should be incorporated into these prices as well. Moreover, researchers are continuing to explore extended-LMP methods for broader-based incorporation of avoidable fixed costs into DAM/RTM energy prices, i.e., into LMPs ($/MWh).

A key conceptual argument against the incorporation of supplier avoidable fixed costs into DAM/RTM energy prices is as follows. These avoidable fixed costs are insurance costs, not production costs. That is, they are the costs that suppliers incur to be able to fulfill their contractual commitments to provide availability of power-paths for possible RTO/ISO-dispatch during future operating periods. This assured availability, in and of itself, provides a critically important service: namely, reduction of volumetric grid risk for these future operating periods. A supplier should be compensated for providing risk-reduction service (dispatchable power-path availability) for a future operating period T whether or not the RTO/ISO subsequently chooses to dispatch this supplier for actual power-path delivery during T.

However, incorporation of supplier avoidable fixed costs solely into DAM/RTM energy prices for a future operating period T would prevent these suppliers from receiving compensation for period-T risk-reduction services per se. Rather, suppliers would only receive compensation for period-T services if they were dispatched for energy deliveries during T.

Consider the following analogous situation. Suppose a fire-insurance company FIC is interested in providing risk-reduction products to households for some future period T. These risk-reduction products are fire-insurance contracts sold to households in advance of T that promise to provide make-whole house repairs in case of a period-T house-fire. Define company FIC’s insurance pool to be the subset of households that purchase a fire-insurance contract from FIC in advance of T. Suppose company FIC is not permitted to require each household in its insurance pool to pay a common “premium payment” ($) whether or not the household experiences house-fire damage during T.

To stay in business, company FIC would then be forced to require each household in its insurance pool to pay the full avoidable cost of any house-fire repair that FIC provides to this household during period T. Thus, why limit house repairs to burned houses? The would-be “fire insurance” company FIC is thus incentivized to function as an ordinary “home repair” company with no provision of risk-reduction services.

Finally, various conceptual and practical arguments can be given against the proposed incorporation of supplier avoidable fixed cost into any energy or operating reserve price determined for a co-optimized DAM/RTM. The incorporation of a positive avoidable fixed cost $c_0$ ($) incurred in advance of an operating period T

---

35 Suppose: (i) the number $N$ of households in FIC’s insurance pool is large; (ii) each household in this insurance pool has the same small independent probability $\beta \in (0, 1)$ of experiencing a house-fire during T; and (iii) each household in this insurance pool would have the same house-repair cost $HRC$ ($) in case of a house-fire. Then, by the Law of Large Numbers, the FIC can “almost surely” guarantee full coverage of its actual total house-fire repair cost during T by: (a) requiring each household in its insurance pool to pay a relatively small premium $\beta \times HRC$ in advance of T; and (b) offering each household in its insurance pool free make-whole house-fire repair during T.
into the period-T price \( \pi_A(T) \) determined for a product \( A \) that would be procured in total period-T amount \( a(T) \) at price \( \pi_A(T) \) requires changing the recorded book-price \( \pi'_A(T) \) to a price level \( \pi'_A(T) = \pi_A(T) + \Delta \pi_A(T) \) such that

\[
\pi'_A(T) \times a(T) = c^o + \pi_A(T) \times a(T),
\]

hence:

\[
\Delta \pi_A(T) \times a(T) = c^o.
\]

However, as detailed in Section 3.3.2, an avoidable fixed cost is defined to be a cost that: (i) arises from a decision to commit now to undertaking some designated type of action at a future time; and (ii) does not depend on the specific form of the undertaken action. Thus, there is no guarantee that an avoidable fixed cost \( c^o \) incurred in advance of some future operating period \( T \) due to a commitment to be available for possible RTO/ISO-instructed delivery of a currently uncertain amount of product \( A \) during \( T \) can be expressed as a stable function of the actual delivery \( a(T) \) of product \( A \) during \( T \). Consequently, the recorded book-prices \( \pi'_A(T) \) for each operating period \( T \) could be highly unstable, implying there is no practical advantage gained relative to making simple direct OOM make-whole payments \( c^o \).

Second, if the book-price \( \pi'_A(T) \) were instead implemented as the period-T price of product \( A \) in place of \( \pi_A(T) \), the transacted \( A \)-amount would presumably change from \( a(T) \) to \( a'(T) = a(T) + \Delta a(T) \) for some non-zero increment \( \Delta a(T) \). That is, any attempt to endogenize the needed “price distortion” by incorporating it into a co-optimized DAM/RTM SCUC/SCED formulation in advance of the determination of an optimal solution would typically result in an inefficient distortion of the resulting quantity solution.

Third, as stressed in Section 3.4.3, energy and operating reserve do not function as commodities within the context of a co-optimized DAM/RTM. Hence, the uniform “unit prices” derived in these market processes for energy and operating reserve, conditional on a given delivery location and delivery time-period, are conceptually-problematic constructs with no efficiency or optimality justification. The incorporation of an avoidable fixed cost into such a presumed uniform market price would thus simply result in a double-layered price distortion.

### 4.5 Growing Revenue Insufficiency Concerns

Market revenue sufficiency for a market \( M \) means that the market-determined revenue earned by each supplier \( i \) participating in \( M \) suffices over time to cover the avoidable cost incurred by supplier \( i \) from this participation; see definition (Def4) in Section 3.4.2. The proliferation of OOM make-whole payments reported in Section 4.4 indicates that RTO/ISO-managed DAMs/RTMs are not currently ensuring revenue sufficiency for their participant suppliers.

For reasons presented in previous sections of this study, and justified in greater detail in [36], DAM/RTM revenue insufficiency is arising in all seven U.S. RTO/ISO-managed wholesale power markets due to fundamental conceptually-problematic economic presumptions embedded in their legacy core DAM/RTM two-settlement system designs: namely, presumptions (P1)–(P4).
4.6 Dangerous Market-Design Conundrum

Participation models, flexibility products, OOM make-whole payments, and revenue insufficiency concerns all appear to be proliferating in current U.S. RTO/ISO-managed wholesale power markets. This proliferation suggests these markets could be caught up in a dangerous market-design conundrum, characterized as follows:

Ptomeaic Epicycle Conundrum for Market Design (“Onion Problem”):

- Fundamental conceptual inconsistencies in the core design principles instituted for a market result in operational difficulties.
- These operational difficulties are addressed by introducing a new layer of rules (an “epicycle”) around the initial core design principles, which results in further operational difficulties.
- Rule-layer (epicycle) accretion then continues to occur because – ignoring the “Sunk Cost is Sunk” Dictum in Section 3.3.2 – correction of the fundamental conceptual inconsistencies in the core design principles is persistently deemed to be too costly to correct.

5 Conceptually-Consistent Alternative Market Design

5.1 Linked Swing-Contract Market Design: Overview

Previous sections of this study support the contention that the four conceptually-problematic economic presumptions (P1)–(P4) reflected in the legacy DAM/RTM two-settlement system design at the core of current U.S. RTO/ISO-managed wholesale power market operations are hindering the smooth transition of these markets to decarbonized grid operations.

This section briefly reviews the key features of an alternative design for grid-supported centrally-managed wholesale power markets, referred to as the Linked Swing-Contract Market Design [36], that appears better suited for the support of decarbonized grid operations. This alternative design is based on the following four conceptually-consistent economic principles [EP1]–[EP4], presented in preliminary form as Counter-Claims 1–4 in Section 1:

[EP1]: Energy (MWh) is not a commodity for grid-supported RTO/ISO-managed wholesale power markets, conditional on grid delivery location \( b \) and delivery period \( T \), because the cost-benefit valuations that a participant assigns to any given energy delivery \( E^* \) (MWh) at \( b \) for \( T \) typically depends on the dynamic attributes of the power-path that delivers \( E^* \) at \( b \) during \( T \).

[EP2]: Any market process carried out within a grid-supported RTO/ISO-managed wholesale power market must necessarily be a forward market \( M(T) \) for a future operating period \( T \) due to the speed of real-time grid operations.

36 Details regarding the development and testing of the Linked Swing-Contract Market Design can be found in the following studies:[1, 19, 23, 24, 25, 27, 34, 36, 37].
[EP3]: Dispatchable Power Resources (DPRs) participating in a forward market M(T) for a future operating period T within a grid-supported RTO/ISO-managed wholesale power market can provide two distinct types of service:

**Physically-Covered Insurance:** The guaranteed availability of flexible power-path production capabilities for possible RTO/ISO dispatch during the future operating period T to protect against volumetric grid risk;

**Real-Time Power-Path Delivery:** The actual delivery of power-paths in response to RTO/ISO dispatch signals received during operating period T to meet just-in-time customer power demands and grid reliability requirements.

[EP4]: Each DPR in [EP3] can use a two-part pricing swing-contract to submit to M(T) a coupled offer of physically-covered insurance for T and real-time power-path delivery during T that ensures revenue sufficiency for the DPR, as follows:

- The swing contract can include a DPR-specified offer price to compensate the DPR prior to T for any avoidable fixed cost the DPR must incur to guarantee its offered physically-covered insurance for T;
- The swing contract can include a DPR-specified performance payment method to compensate the DPR after T for any variable cost the DPR must incur for verified power-path delivery during T in response to RTO/ISO dispatch signals.

### 5.2 Key Design Features

As carefully explained and illustrated in [36, Chs. 10-11], the Linked Swing-Contract Market Design postulates a collection $\mathcal{M} = \{M(T) | T \in \mathcal{T}\}$ of sets $M(T) = \{M'(T), M''(T), \ldots\}$ of RTO/ISO-managed forward reserve markets $M'(T)$, $M''(T)$, ... for each operating period T in a designated operating-period set $\mathcal{T}$. The forward reserve markets $M'(T)$, $M''(T)$, ... in $\mathcal{M}(T)$ for a given future operating period T are differentiated by the durations of their look-ahead horizons $\text{LAH}'(T)$, $\text{LAH}''(T)$, ..., which can range from years to minutes; cf. Fig. 6.

A power-path for an operating period T is a sequence of power injections and/or power withdrawals that take place at a single grid location during T. Reserve offers submitted by Dispatchable Power Resources (DPRs) to each RTO/ISO-managed forward reserve market M(T) for a future operating period T are the offered guaranteed availability of power-path production capabilities for possible RTO/ISO dispatch during T. Reserve bids submitted by Load-Serving Entities (LSEs) to M(T) are requests for power-path deliveries during T. The RTO/ISO clears reserve offers/bids submitted to M(T) to maximize the expected net benefit (benefit minus avoidable cost) of M(T) participants, given current state conditions (including market linkages), and subject to system constraints that include nodal net-load balancing constraints for T.

---

37 As discussed and illustrated in [36, Chs. 10-11], “market linkages” refers to the specific linkages established among the markets in $\mathcal{M}$ by the RTO/ISO’s successive contract-clearing and dispatch decisions. The RTO/ISO keeps tracks of these market linkages by carrying forward on its books an adaptively updated record of its cleared reserve bids/offers and its dispatch decisions.
The participants in each RTO/ISO-managed forward reserve market $M(T)$ consist of Dispatchable Power Resources (DPRs), Load-Serving Entities (LSEs), and Intermittent Power Resources (IPRs). DPRs are grid-connected RTO/ISO-dispatchable power resources that include relatively large-scale customers, generators, storage devices, and hybrid entities (e.g., renewable power facilities fully firmed by storage). LSEs are aggregators (intermediaries) that service the power requirements of customers populating grid-connected lower-voltage distribution networks. IPRs are grid-connected power resources whose grid injections and/or withdrawals of power are not fully controllable by RTO/ISO dispatch.

DPRs offer reserve into the RTO/ISO-managed forward reserve markets $M(T)$ in $M$ by means of swing contracts; hence, the markets $M(T)$ are referred to as swing-contract (SC) markets. A swing-contract $SC_m(T)$ offered by a DPR $m$ into an SC market $M(T)$ for a future operating period $T$ is a physically-covered two-part pricing insurance contract expressible in the following standardized form:

$$SC_m(T) = \left( \alpha_m(T), T_{m}^{ex}(T), PP_{m}(T), \phi_m(T) \right)$$ (12)

The first component $\alpha_m(T)$ in (12) is the offer price that $m$ specifies for $SC_m(T)$. A non-negative offer price designates the amount that must be paid to $m$ either directly or in amortized payment-schedule form if the RTO/ISO clears $SC_m(T)$ for $T$. The offer price $\alpha_m(T)$ permits $m$ to ensure recovery ex ante (i.e., in advance of $T$) for all avoidable fixed cost that $m$ must incur to fulfill the terms of $SC_m(T)$, should $m$ decide to submit $SC_m(T)$ to $M(T)$ and should the RTO/ISO then choose to clear $SC_m(T)$ for $T$.

The second component $T_{m}^{ex}$ in (12) is the set of possible exercise times $t_{m}^{ex}$ that $m$ designates for $SC_m(T)$ between the close of $M(T)$ and the start-time for the future operating period $T$. As discussed more carefully in [36, Sec. 4.3], the number and positioning of these exercise times determine whether $SC_m(T)$ is a firm contract or some type of option contract (e.g., European, American, Bermudan, ...).

---

38 As in current U.S. RTO/ISO-managed wholesale power markets, purely financial traders (i.e., traders with no physical generation or customer service obligations) could be virtual participants in market $M(T)$, i.e., participants whose quantity positions for the future operating period $T$ have to be liquidated (reduced to zero) prior to $T$. However, the efficacy of permitting virtual participants in grid-supported centrally-managed wholesale power markets remains highly controversial; see [5]. The critical risks for such markets are systemic (system-wide) risks: namely, volumetric (volume imbalance) grid risk and systemic financial risk. Traders can protect themselves against uncertainty arising at their particular grid locations through privately undertaken hedging actions as long as this uncertainty is not strongly correlated across the grid.

39 These customers include pure power and/or ancillary service producers, pure power consumers, and hybrid prosumers able to function as producers or consumers depending on their economic incentives and/or local state conditions.

40 Swing-contract offer prices are not restricted in sign. For example, a DPR $m$ with a positive minimum sustainable power-injection level could specify an offer price $\alpha_m < 0$ in an attempt to ensure the RTO/ISO clears his swing contract $SC_m$. This negative offer price commits the DPR to pay the amount $-\alpha_m$ should $SC_m$ be cleared.

41 See Section 3.3 for a careful definition and discussion of avoidable fixed cost, and see Appendix A.4 for a listing of different types of avoidable fixed cost that real-world DPRs can incur.
The third component $\mathbb{P}_m(T)$ in (12) denotes the set of power-paths $p_m(T) = \{p_m(t) \mid t \in T\}$ that $m$ is offering for possible RTO/ISO-dispatched delivery at $m$’s grid point-of-connection $b_m$ during $T$. Ideally, the power-path set $\mathbb{P}_m(T)$ should be a “digital twin” expressing the full physical power-path production capabilities of $m$ during $T$.

In practice, $\mathbb{P}_m(T)$ will typically be some form of approximation for this digital twin that conveys the key physical attributes of $m$’s offered power-paths. The attributes of an offered power-path must include its grid delivery location $b_m$. Additionally, attributes can include its static features, such as delivered amount of energy (MWh), as well as its dynamic features such as down-time/up-time profile; power capacity (MW) profile; ramp-rate (MW/min) profile; and power-factor (kW/kVA) profile. The forms and ranges of these power-path attributes across the offered power-paths in the power-path set $\mathbb{P}_m(T)$ determine the degree of swing (flexibility) in $m$’s offered reserve.

The final component $\phi_m(T)$ in (12) is a performance payment method designating the ex-post (i.e., after $T$) compensation that $m$ requires for any variable cost that $m$ incurs for verified period-$T$ delivery of a power-path in $\mathbb{P}_m(T)$ in accordance with RTO/ISO dispatch instructions. Thus, $\phi_m(T)$ permits $m$ to ensure full recovery of his variable cost for $T$, subject to delivery verification, which encourages close following of RTO/ISO dispatch instructions. For example, as indicated in Appendix A.4, this variable cost could include payments for inputs (fuel, labor, ...) needed for power-path production, charges for transmission services, and wear-and-tear depreciation of physical equipment.

The performance payment method $\phi_m(T)$ that a DPR $m$ includes in a swing contract (12) can take a wide variety of forms. Ideally, however, this performance payment method should be expressible in standardized metrics that permit the RTO/ISO and DPRs: (i) to agree ex ante on the precise nature of each DPR’s offered risk-reduction service (dispatchable power-path availability) for $T$; and (ii) to verify ex post the extent to which any RTO/ISO dispatch instructions conveyed to DPRs for period-$T$ performance (power-path delivery) have accurately been followed.

The reserve bids submitted by LSEs into a swing-contract market $M(T)$ for a future operating period $T$ on behalf of their managed customers are demands for power-path deliveries at designated grid locations during $T$. These reserve bids are prepared

---

42 The DPR $m$ is responsible for ensuring the feasibility of the period-$T$ RTO/ISO-dispatched delivery of each power-path $p_m(T)$ in $\mathbb{P}_m(T)$ that is consistent with run-time RTO/ISO-induced linkage requirements at the start of $T$. For example, if $m$ has received and followed RTO/ISO dispatch instructions for the operating period immediately prior to $T$, the RTO/ISO must condition any choice of a power-path in $\mathbb{P}_m(T)$ for period-$T$ dispatch on the dispatch-induced state of $m$ at the start of $T$. See [36, Ch. 10] for a detailed discussion of linkage implementation for the Linked Swing-Contract Market Design.

43 If the closure of the operating period $T$ is compact, and the closure of $\mathbb{P}_m(T)$ over the closure of $T$ consists of a suitably dense subset of the set of all continuous power-paths over the closure of $T$, the power-paths in $\mathbb{P}_m(T)$ can be jointly expressed (in approximate form) by a finite collection of attribute range-limit constraints: e.g., min/max down-time/up-time limits; min/max power capacity limits; min/max ramp-rate limits; and min/max power-factor limits. See [36, Chs. 5,7,16] for concrete illustrations.
differentiated by: (i) the extent to which they are subject to RTO/ISO-dispatch control by means of price signals; and (ii) the extent to which they permit the RTO/ISO to determine accurate forecasts for the aggregate power-paths delivered to customers at different bus locations during T.

For example, an LSE reserve bid can take the form of a fixed (non-dispatched must-service) demand for the delivery of a specific power-path \( p(b, T) \) at a specific grid bus \( b \) during T. Such reserve bids are not subject to RTO/ISO-dispatch control. Moreover, the power-path \( p(b, T) \) could simply be the LSE’s mean forecast for its customers’ aggregate power-path at \( b \) during T, hence subject to large forecast errors. At the other extreme, an LSE reserve bid can take the form of a collection of price-dispatchable demand schedules at a particular grid bus \( b \) for a collection of sub-periods forming a partition of T. In principle, such reserve bids are fully subject to RTO/ISO-dispatch control via price signals communicated to the customers through the LSE as intermediary. Moreover, assuming customers are appropriately penalized for deviations from price-signaled power levels, the RTO/ISO should be able to forecast these power levels with high accuracy.

Finally, the power injections of an IPR unfirmed by storage that take place at some specific grid bus \( b \) during T are considered to be a fixed (non-dispatched must service) power-path delivery during T that must be forecasted by the RTO/ISO. This forecast is directly entered into the system constraints for the contract-clearing optimization that the RTO/ISO conducts for \( M(T) \): specifically, it is entered into the bus-\( b \) power balance constraints for period T.

The complete form of the contract clearing optimization problem conducted by the RTO/ISO for each swing-contract market \( M(T) \) is carefully explained and analytically illustrated in [36, Chs. 6-9]. To preserve its status as an independent fiduciary, the RTO/ISO allocates all net reserve procurement costs and transmission service costs incurred for \( M(T) \) operations back to \( M(T) \) participants based on their relative contributions to these costs. The specific rules used by the RTO/ISO for these cost allocations are carefully presented and motivated in [36, Ch. 6.7].

5.3 Comparison with U.S. RTO/ISO-Managed Markets

This section provides high-level comparisons of the basic design features and the market-clearing optimization formulations for current U.S. DAMs and swing-contract (SC) DAMs under the simplifying assumption that all customer demand takes the form of fixed (non-dispatched must-service) load.\(^{44}\)

Figure 7 indicates the main similarities and differences in basic design features between current and swing-contract (SC) day-ahead markets (DAMs).

The main similarities are as follows: (i) Both DAM designs are RTO/ISO managed; (ii) both DAM designs have the same types of market participants; and (iii) both DAM designs are subject to standard types of system constraints.

\(^{44}\) Detailed comparisons between the design of current U.S. RTO/ISO-managed wholesale power markets and the Linked Swing-Contract Market Design are provided in [36, Chs. 2-3,12-15].
The main differences, listed more carefully below, involve product definition, contract forms, and settlement rules, not real-time operations. As demonstrated in [36, Ch. 16], these market design differences can be introduced gradually into current U.S. RTO/ISO-managed DAMs.

- **SC DAMs** are physically-covered forward reserve markets, where reserve consists of next-day RTO/ISO-dispatchable power-paths;
- **Reserve offers** are two-part pricing swing contracts, in firm or option form, that offer both availability and dispatched delivery of dispatchable power-paths;
- SC DAM settlements consist entirely of offer-price payments (or amortized payment obligations) to submitters of cleared reserve offers as compensation for volumetric risk-reduction provided by power-path availability;
- No performance settlements occur in advance of actual verified performance;
- Dispatchable power resources can ensure their revenue sufficiency (i.e., market revenue ≥ avoidable cost) by appropriate specification of the offer prices and performance payment methods they include in their reserve offers; no resort to OOM make-whole payments is needed.

Figure 8 provides high-level comparisons of the market-clearing optimization formulations for current U.S. DAMs and SC DAMs. The main similarity is that each optimization can be formulated as a Mixed Integer Linear Programming (MILP) problem. The main differences are:

- The only binary-valued RTO/ISO decision variables for SC DAM optimizations are yes/no contract-clearing indicators;
- The SC DAM objective function permits explicit full inclusion of supplier avoidable fixed cost ("offer cost") and supplier variable cost ("performance cost");

### Table: Basic Market Design Features

<table>
<thead>
<tr>
<th></th>
<th><strong>Current DAM</strong></th>
<th><strong>SC DAM</strong></th>
</tr>
</thead>
</table>
| **Similarities**    | • Conducted day-ahead to plan for next-day operations  
                      • RTO/ISO-managed  
                      • Participants include LSEs, DPRs, & IPRs  
                      • Same system constraint types: e.g., power balance, line capacity limits, reserve requirements, resource attributes. | **Optimal contract clearing** |
| **Optimization form** | SCUC & SCED                                                                     | Optimal contract clearing |
| **Settlement**      | Locational marginal prices                                                       | Swing contracts are two-part pricing contracts |
| **Market payments** | Payment for next-day energy before actual energy delivery                       | Payment for resource availability now & resource performance ex post |
| **OOM payments**    | Make-whole payments                                                              | No make-whole payments |
| **Info released to participants** | Unit commitments, LMPs, & next-day dispatch schedule | Which swing contracts have been cleared |

**Fig. 7 Basic market design features: Current U.S. RTO/ISO-managed DAM vs. SC DAM.**
**Fig. 8** Market-clearing optimization formulations: Current U.S. RTO/ISO-managed DAM SCUC/SCED optimizations vs. SC DAM optimization.

- The SC DAM objective function includes *imbalance cost* indicating any positive slack-variable solutions ("canaries in the coal mine") for the slack variables included in the system power-balance constraints.
- In SC DAM optimizations, the only settlements are offer-price payments (or amortized offer-price payment obligations) for cleared swing-contract reserve offers for future operating periods; no performance payments are made in advance of verified real-time performance.

## 6 Conclusion

This study presents and critiques four conceptually-problematic economic presumptions (P1)–(P4) reflected in the DAM/RTM two-settlement system design, the legacy core design for all seven current U.S. RTO/ISO-managed wholesale power markets. Proliferating participation models, flexibility products, out-of-market payments, and revenue insufficiency concerns indicate presumptions (P1)–(P4) could be hindering the smooth transition of these markets to decarbonized grid operations.

The study also briefly reviews the key features of an alternative design for grid-supported centrally-managed wholesale power markets: namely, the Linked Swing-Contract Market Design developed and tested in a series of studies over the past ten years; see [1, 19, 24, 25, 27, 34, 36, 37]. This alternative design is based on the four

---

45 See [36, Chs. 7,16]. Future implementations of the Linked Swing-Contract Market Design currently under development will also include slack variables ("canaries in the coal mine") for additional types of system constraints, such as nodal and zonal reserve requirement constraints.

Major extensions of the Linked Swing-Contract Market Design currently under development include a fundamental change of variables from active and reactive power \((p, q)\) to voltage and current \((v, i)\) to facilitate a more comprehensive and empirically-compelling representation of participant net benefits and grid reliability constraints (including voltage constraints); see [7]. This extension could permit the fundamental market design concerns considered in the current study and the fundamental system security concerns considered in [4] to be addressed in an integrated conceptually-cohesive manner.

Finally, throughout this study, emphasis has been placed on grid-supported centrally-managed wholesale power markets as flexibility-support mechanisms: that is, as mechanisms enabling just-in-time production and transmission of bulk power to satisfy just-in-time customer power demands. As seen, ensuring the efficiency and reliability of such mechanisms is a complex multi-faceted problem.

A physically-feasible alternative might be to transit to a fully storage-supported world. Flexible producer determination of power production levels and flexible customer determination of power usage levels could be supported entirely by local producer and customer storage devices, linked via a supply chain consisting of retail storage-device stores and/or a network of charge/discharge stations.

However, a key economic concern regarding this fully storage-supported world is the potential for substantial inefficiency (wasted resources). Suppose the average state-of-charge (SOC) that producers and customers maintain for their local storage devices over time is uniformly bounded above zero. Then, from a global vantage point, it would appear as if a positive, possibly-large, and possibly growing inventory of (potential) energy were being carried forward through time, forever unused, instead of contributing to the creation of net benefit.

Consequently, at least at present, striving to redesign current grid-supported centrally-managed wholesale power markets to enable them to operate as reliable efficient flexibility-support mechanisms over lower-carbon grids would seem to be the better option. Moreover, this redesign, itself, should be flexible and open to further adaptation. The ultimate goal, surely, must be robust wholesale power market design for transacting in a deeply-uncertain continually-evolving world.
Appendices: Quick-Reference Glossaries and Guides

A.1 Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Alternating Current</td>
</tr>
<tr>
<td>CAISO</td>
<td>California Independent System Operator</td>
</tr>
<tr>
<td>CFD</td>
<td>Contract-For-Difference</td>
</tr>
<tr>
<td>CCSM</td>
<td>Competitive Commodity Spot Market</td>
</tr>
<tr>
<td>CSM</td>
<td>Commodity Spot Market</td>
</tr>
<tr>
<td>D</td>
<td>Generic symbol for a day</td>
</tr>
<tr>
<td>DAM</td>
<td>Day-Ahead Market</td>
</tr>
<tr>
<td>DC</td>
<td>Direct Current</td>
</tr>
<tr>
<td>DM</td>
<td>Decision-Maker</td>
</tr>
<tr>
<td>DPR</td>
<td>Dispatchable Power Resource</td>
</tr>
<tr>
<td>ERCOT</td>
<td>Electric Reliability Council of Texas</td>
</tr>
<tr>
<td>FERC</td>
<td>Federal Energy Regulatory Commission</td>
</tr>
<tr>
<td>FTR</td>
<td>Financial Transmission Right</td>
</tr>
<tr>
<td>H</td>
<td>Generic symbol for an hour</td>
</tr>
<tr>
<td>IPR</td>
<td>Intermittent Power Resource</td>
</tr>
<tr>
<td>ISO</td>
<td>Independent System Operator</td>
</tr>
<tr>
<td>ISO-NE</td>
<td>Independent System Operator for New England</td>
</tr>
<tr>
<td>kWh</td>
<td>Kilowatt hour</td>
</tr>
<tr>
<td>kVA</td>
<td>Kilovolt-amperes (1000 volt-amperes)</td>
</tr>
<tr>
<td>LMP</td>
<td>Locational Marginal Price</td>
</tr>
<tr>
<td>LSE</td>
<td>Load-Serving Entity</td>
</tr>
<tr>
<td>MILP</td>
<td>Mixed Integer Linear Programming</td>
</tr>
<tr>
<td>MISO</td>
<td>Midcontinent Independent System Operator</td>
</tr>
<tr>
<td>MWh</td>
<td>Megawatt hour</td>
</tr>
<tr>
<td>NOPR</td>
<td>Notice Of Proposed Rule-making (FERC)</td>
</tr>
<tr>
<td>NYISO</td>
<td>Independent System Operator for New York</td>
</tr>
<tr>
<td>OOM</td>
<td>Out-of-Market</td>
</tr>
<tr>
<td>OPF</td>
<td>Optimal Power Flow</td>
</tr>
<tr>
<td>PJM</td>
<td>PJM Interconnection</td>
</tr>
<tr>
<td>QSE</td>
<td>Qualified Scheduling Entity (ERCOT)</td>
</tr>
<tr>
<td>RTM</td>
<td>Real-Time Market</td>
</tr>
<tr>
<td>RTO</td>
<td>Regional Transmission Organization</td>
</tr>
<tr>
<td>SCED</td>
<td>Security-Constrained Economic Dispatch</td>
</tr>
<tr>
<td>SCUC</td>
<td>Security-Constrained Unit Commitment</td>
</tr>
<tr>
<td>SPP</td>
<td>Southwest Power Pool</td>
</tr>
<tr>
<td>TNS</td>
<td>Total Net Surplus</td>
</tr>
</tbody>
</table>
### A.2 Standard Transmission System Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ancillary service</td>
<td>Service that supports system reliability</td>
</tr>
<tr>
<td>Availability cost</td>
<td>Avoidable fixed cost incurred to ensure reserve availability</td>
</tr>
<tr>
<td>Commitment</td>
<td>Scheduling of a dispatchable power resource for possible future RTO/ISO dispatch</td>
</tr>
<tr>
<td>Dispatch</td>
<td>Signaling a grid-connected power resource to inject/withdraw power</td>
</tr>
<tr>
<td>Energy</td>
<td>Abbreviation for electric energy (MWh)</td>
</tr>
<tr>
<td>Fixed power injection</td>
<td>Non-dispatched must-service power injection into a grid</td>
</tr>
<tr>
<td>Fixed load</td>
<td>Non-dispatched must-service power withdrawal from a grid</td>
</tr>
<tr>
<td>Generation</td>
<td>Production of power either for local behind-the-meter use or for grid injection</td>
</tr>
<tr>
<td>Intermittent power</td>
<td>Power injections/withdrawals not fully under RTO/ISO-dispatchable control</td>
</tr>
<tr>
<td>Load</td>
<td>Used as a synonym for grid power withdrawal; technically, a grid device or grid component to which power is delivered</td>
</tr>
<tr>
<td>Locational marginal price</td>
<td>Energy price conditional on delivery location and delivery time-period</td>
</tr>
<tr>
<td>Merit-order dispatch</td>
<td>Dispatch in accordance with net benefit contribution</td>
</tr>
<tr>
<td>Must-service</td>
<td>Power withdrawal (injection) that must be balanced by power injection (withdrawal) under normal grid operating conditions</td>
</tr>
<tr>
<td>Net load</td>
<td>Customer load &amp; inadvertent power loss minus net non-dispatched power injection</td>
</tr>
<tr>
<td>Net fixed load</td>
<td>Customer fixed load minus net fixed power injection</td>
</tr>
<tr>
<td>Net reserve cost</td>
<td>Reserve procurement cost minus reserve revenue receipts</td>
</tr>
<tr>
<td>Non-dispatchable power</td>
<td>Power not under RTO/ISO-dispatchable control</td>
</tr>
<tr>
<td>Operating reserve</td>
<td>Generation capacity (MW) unencumbered by energy delivery obligations</td>
</tr>
<tr>
<td>Performance</td>
<td>Delivery of a good or service in response to RTO/ISO-communicated instructions</td>
</tr>
<tr>
<td>Performance cost</td>
<td>Variable cost incurred for providing delivery of a good or service in response to RTO/ISO-communicated instructions</td>
</tr>
<tr>
<td>Power</td>
<td>Abbreviation for electric power (MW)</td>
</tr>
<tr>
<td>Power absorption</td>
<td>Incremental down/up changes in power withdrawal offered into a power system as an ancillary service</td>
</tr>
<tr>
<td>Power imbalance</td>
<td>Discrepancy between grid power injection &amp; grid power withdrawal/loss</td>
</tr>
<tr>
<td>Power injection</td>
<td>Insertion of power into a grid at an electrical point-of-connection</td>
</tr>
<tr>
<td>Power-path</td>
<td>Sequence of power injections and/or withdrawals at a single grid location during a specified time interval</td>
</tr>
<tr>
<td>Power-path delivery</td>
<td>Power-path delivered during an operating period in response to received RTO/ISO-dispatch instructions</td>
</tr>
<tr>
<td>Power usage</td>
<td>Use of power as an intermediate good to further some end</td>
</tr>
<tr>
<td>Power withdrawal</td>
<td>Extraction of power from a grid at an electrical point-of-connection</td>
</tr>
<tr>
<td>Reserve</td>
<td>Service or product-provision capability that could be used to support grid reliability</td>
</tr>
<tr>
<td>Reserve bid</td>
<td>Contract requesting reserve availability</td>
</tr>
<tr>
<td>Reserve offer</td>
<td>Contract offering reserve availability</td>
</tr>
<tr>
<td>Transmission service cost</td>
<td>Variable cost incurred for grid operation and maintenance</td>
</tr>
</tbody>
</table>
### A.3 Standard Economic Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset</td>
<td>Anything of durable value, whether physical or financial</td>
</tr>
<tr>
<td>Avoidable cost</td>
<td>Cost that can be avoided by <em>not</em> committing to undertake a specified type of action</td>
</tr>
<tr>
<td>Avoidable fixed cost</td>
<td>Avoidable cost not dependent on exact form of action as long as it has specified type of action</td>
</tr>
<tr>
<td>Benefit function</td>
<td>Function measuring the increase in own-welfare attained by a customer from the consumption and/or use of goods and/or services</td>
</tr>
<tr>
<td>Commodity</td>
<td>Asset $Q$ with a standard unit of measurement such that, at any given location and time, $Q$-traders consider each available $Q$-unit to be a perfect substitute for each other available $Q$-unit.</td>
</tr>
<tr>
<td>Competitive market</td>
<td>Commodity market whose buyers and suppliers are price-takers</td>
</tr>
<tr>
<td>Competitive equilibrium</td>
<td>Competitive market price-quantity outcome such that aggregate demand = aggregate supply</td>
</tr>
<tr>
<td>Contract in firm form</td>
<td>Non-contingent contract whose terms are binding on all parties</td>
</tr>
<tr>
<td>Contract in option form</td>
<td>Holder has the right, but not the obligation, to exercise the contract</td>
</tr>
<tr>
<td>Demand schedule (inverse)</td>
<td>A schedule expressing a buyer’s purchase reservation value for each successively demanded unit of a commodity</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Non-wastage of resources</td>
</tr>
<tr>
<td>Fixed cost</td>
<td>Cost not dependent on a specific form of action undertaken</td>
</tr>
<tr>
<td>Forward market</td>
<td>Transacted asset amounts and payment obligations for these transacted asset amounts <em>occur in advance</em> of the delivery of these transacted asset amounts</td>
</tr>
<tr>
<td>Futures market</td>
<td>Forward market for a commodity</td>
</tr>
<tr>
<td>Good</td>
<td>Exchangeable physical item whose consumption (or use) provides benefit to the consumer (or user)</td>
</tr>
<tr>
<td>Hedonic pricing</td>
<td>Pricing of a product by means of prices separately assigned to its intrinsic physical attributes and/or its external circumstances</td>
</tr>
<tr>
<td>Joint products</td>
<td>Products jointly produced from a given set of inputs</td>
</tr>
<tr>
<td>Law of One Price</td>
<td>In the absence of trade frictions (e.g., differences in trade locations, trade times, or trader product information), trader exploitation of arbitrage opportunities will ensure that the unit-price a trader is charged for an available unit of a commodity $Q$ is the same for each available unit of $Q$.</td>
</tr>
<tr>
<td>Net benefit</td>
<td>Benefit minus avoidable cost</td>
</tr>
<tr>
<td>Net buyer surplus</td>
<td>Difference between a buyer’s maximum willingness to pay for a quantity amount $q$ and the buyer’s actual payment for purchase of $q$</td>
</tr>
<tr>
<td>Net supplier surplus</td>
<td>Difference between a supplier’s actual compensation for supplying a quantity amount $q$ and the supplier’s minimum acceptable compensation for supplying $q$</td>
</tr>
<tr>
<td>Opportunity cost</td>
<td>Earnings foregone by not committing assets to an alternative next-best use</td>
</tr>
<tr>
<td>Pareto efficiency</td>
<td>No wastage of opportunity to increase benefit for some at no cost to others by means of a feasible reallocation of resources</td>
</tr>
<tr>
<td>Perfect substitutes</td>
<td>Two items are <em>perfect substitutes</em> for a trader at a given location and time if substitution of either item for the other item has no effect on the trader’s valuation of this item.</td>
</tr>
<tr>
<td>Price-taker</td>
<td>A trader participating in a market for a good or service that behaves as if his own market transactions cannot affect the market price of this good or service.</td>
</tr>
<tr>
<td>Productive efficiency</td>
<td>No physical wastage of production inputs and/or outputs</td>
</tr>
<tr>
<td>Purchase reservation value</td>
<td>Maximum amount a buyer is willing to pay to purchase a quantity amount $q$</td>
</tr>
<tr>
<td>Revenue sufficiency</td>
<td>Supplier revenue is sufficient to cover supplier avoidable cost</td>
</tr>
<tr>
<td>Risk aversion</td>
<td>Unwillingness to participate in a risky undertaking with zero expected payoff</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Sale reservation value</td>
<td>Minimum payment a supplier is willing to accept for sale of a quantity amount $q$</td>
</tr>
<tr>
<td>Service</td>
<td>Action taken by an entity that provides benefit to another entity</td>
</tr>
<tr>
<td>Spot market</td>
<td>Transacted asset amounts, payments for these transacted asset amounts, and deliveries of these transacted asset amounts <em>all occur at the same location and time</em> (&quot;on the spot&quot;).</td>
</tr>
<tr>
<td>Sunk cost</td>
<td>Unavoidable fixed cost</td>
</tr>
<tr>
<td>Supply schedule (inverse)</td>
<td>Schedule expressing a supplier’s sale reservation value for each successively supplied unit of a commodity</td>
</tr>
<tr>
<td>Transaction cost</td>
<td>Avoidable fixed cost incurred to organize a production process</td>
</tr>
<tr>
<td>Variable cost</td>
<td>Avoidable cost that depends on a specific undertaken action (e.g., production level)</td>
</tr>
<tr>
<td>Volumetric risk</td>
<td>Possibility of loss from an unexpected volume imbalance</td>
</tr>
</tbody>
</table>
A.4 Cost Types for Grid-Supported RTO/ISO-Managed Wholesale Power Markets: Empirical Examples

**Types of Avoidable Fixed Cost:**

1. **Capital Investment Cost.** Land acquisition, building construction; equipment purchases. Financed by *internal financing* (i.e., funds on hand), or by *external financing* taking two possible forms:
   - *Direct Financing:* Sell *newly issued* securities in primary security markets to lenders willing to invest in risky assets (i.e., assets with chance of loss) that also offer a sufficiently high chance of gain;
   - *Indirect Financing:* Obtain loans from financial intermediaries, typically secured by some form of collateral, that then result in amortized streams of payment obligations.

2. **Transaction Cost.** Insurance, building code compliance, licensing fees, employee search. Transaction costs are typically financed by internal financing.

3. **Opportunity Cost.** Expected net earnings from a best possible alternative use of assets, e.g., use of generation units directly (behind the meter) for local purposes.

4. **Unit Commitment Cost.** Start-up, no-load, minimum-run, and/or shut-down cost that are incurred for ensuring the availability of power-paths for possible RTO/ISO dispatched delivery during a future operating period but are not dependent on the specific form (if any) of this delivered power-path.

**Types of Variable Cost:**

1. **Fuel Cost.** Charges for pulverized coal, natural gas, nuclear, petroleum, and/or refuse-derived fuels as inputs to power production.

2. **Labor Cost.** Salaries/wages for: legal/tax advice; advertisement; planning; supervision; trading-desk operations; maintenance; and repair.

3. **Intermediate Good (Supply-Chain) Cost.** Rail/barge/pipeline/truck transport charges for fuel deliveries; replenishment of used-up supplies.

4. **Equipment/Software Rental Cost.** Rental charges for office equipment, cars, and software licenses.

5. **Depreciation of Owned Machinery.** Generation unit wear-and-tear due to start-up, normal, and/or shut-down ramping required to follow RTO/ISO-sigualed dispatch set-points during successive operating periods.

6. **Assessed Charges for Transmission Services.** Transmission grid operation and maintenance (O&M) costs allocated across market participants.

7. **Variable-Cost Offsets from Sales of Valuable Bi-Products.** Revenue offset to variable cost of a product due to joint production, e.g., co-generation of valuable heating services along with power by Combined Heat and Power (CHP) units.

8. **Disposal Cost for Waste Bi-Products.** Cost incurred by power plants (e.g., nuclear) to dispose of solid-waste output resulting from plant operations.
### A.5 Swing-Contract Market Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acronyms &amp; Generics:</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Generic symbol for a day</td>
</tr>
<tr>
<td>DPR</td>
<td>Dispatchable Power Resource</td>
</tr>
<tr>
<td>H</td>
<td>Generic symbol for an hour</td>
</tr>
<tr>
<td>IPR</td>
<td>Intermittent Power Resource</td>
</tr>
<tr>
<td>LAH(T)</td>
<td>Look-ahead horizon between close of M(T) and start of T</td>
</tr>
<tr>
<td>LSE</td>
<td>Load-Serving Entity</td>
</tr>
<tr>
<td>M(T)</td>
<td>Swing-contract market for a future operating period T</td>
</tr>
<tr>
<td>m</td>
<td>Generic symbol for a DPR</td>
</tr>
<tr>
<td>n</td>
<td>Generic symbol for an IPR</td>
</tr>
<tr>
<td>p</td>
<td>Generic symbol for a power level (MW)</td>
</tr>
<tr>
<td>p</td>
<td>Generic symbol for a power-path</td>
</tr>
<tr>
<td>p(T)</td>
<td>Generic symbol for a power-path {p(t)</td>
</tr>
<tr>
<td>r</td>
<td>Generic symbol for a ramp-rate (MW/min)</td>
</tr>
<tr>
<td>SC</td>
<td>Swing contract taking the general form SC = (α, T^ex, PP, φ)</td>
</tr>
<tr>
<td>SC_m(T)</td>
<td>SC submitted by a DPR m to a swing-contract market M(T) for T</td>
</tr>
<tr>
<td>t^ex</td>
<td>Exercise time in an exercise set T^ex</td>
</tr>
<tr>
<td>t^ex_m(T)</td>
<td>Exercise time in an exercise set T^ex_m(T)</td>
</tr>
<tr>
<td>T = [t^s, t^e)</td>
<td>Operating period with start-time t^s and end-time t^e</td>
</tr>
<tr>
<td>α</td>
<td>Offer price ($) for a swing-contract SC</td>
</tr>
<tr>
<td>α_m(T)</td>
<td>Offer price ($) for a swing-contract SC_m(T)</td>
</tr>
<tr>
<td>ΔT</td>
<td>Duration of operating period T, measured in real hourly units (e.g., 0.6h)</td>
</tr>
<tr>
<td>φ</td>
<td>Performance payment method for a swing contract SC that maps PP into payments</td>
</tr>
<tr>
<td>φ_m(T)</td>
<td>Performance payment method for a swing contract SC_m(T) that maps each power-path p ∈ PP_m(T) into a dollar payment ($)</td>
</tr>
</tbody>
</table>

Sets & Subsets:
- B = \{1, \ldots, NB\} | Index set for the buses b of a transmission grid |
- C_j(b) | Collection of customers serviced by load-serving entity j ∈ LS(b) |
- L | Index set for the distinct bus-to-bus line segments ℓ of a transmission grid |
- L_j(b) ⊆ L | Subset of transmission-grid line segments originating at bus b |
- L_b | Subset of transmission-grid line segments ending at bus b |
- LS | Index set for the load-serving entities j participating in a swing-contract market |
- LS(b) ⊆ LS | Subset of load-serving entities in LS that service power customers at bus b |
- M | Index set for DPRs m participating in a swing-contract market |
- M(b) ⊆ M | Subset of DPRs m in M that are electrically connected at bus b |
- NG | Index set for non-dispatchable generation units n participating in a swing-contract market |
- NG(b) ⊆ NG | Subset of n in NG that are electrically connected at bus b |
- PP | Set of dispatchable power-paths p offered by a swing contract SC |
- PP_m(T) | Set of dispatchable power-paths p_m(T) offered by a swing contract SC_m(T) |
- P_m | Set of feasible sustainable power levels p (MW) for m |
- R_m | Set of feasible ramp-rates r (MW/min) for m |
- T^ex | Set of possible exercise times t^ex for a swing-contract SC |
- T^ex_m(T) | Set of possible exercise times t^ex_m(T) for a swing contract SC_m(T) |
References


