Hung-po Chao *is Area Manager of Policy and Risk Analysis & Power Market Design at the Electric Power Research Institute (EPRI), Palo Alto, CA, and is Consulting Professor of Management Science and Engineering at Stanford University. In 1999, Dr. Chao led a study for the Taiwan government assessing Taiwan's electricity liberalization strategy. Currently, he assists NEPOOL and ISO New England on market decision issues. Dr. Chao holds a Ph.D. degree in operations research from Stanford.*

Stephen C. Peck *is Technical Executive of the Environment Division at EPRI. He has served as a member of the economics faculty of the University of California, Berkeley, and a member of the NAS/NRC Committee on Global Energy and Associated Ecological Problems. He holds a Ph.D. in business economics from the University of Chicago.*

Shmuel S. Oren *is Professor of Industrial Engineering and Operations Research at the University of California at Berkeley and a former chairman of that department. He is the Berkeley site director of PSerc, a multi-university Power Systems Engineering Research Center. Dr. Oren has been a consultant on electricity restructuring issues to the Brazilian Electricity Regulatory Agency (ANEEL) and to the Texas Public Utility Commission. He holds a Ph.D. in engineering economic systems from Stanford.*

Robert B. Wilson *holds the McBean Chair of Economics at Stanford Business School. He is an elected member of the National Academy of Sciences and the American Academy of Arts and Sciences. Dr. Wilson advises the California Power Exchange, the California and New England independent system operators, and others on the design of wholesale electricity markets. He holds a D.B.A. degree from Harvard University.*

The authors thank Steven Stoft, Larry Ruff, and William Hogan for helpful comments. *They may be contacted via email at hchao@epri.com.*

Flow-Based Transmission Rights and Congestion Management

Combining the advantages of financial and physical rights, a flow-based transmission reservation approach facilitates liquidity and efficient risk management.

Hung-po Chao, Stephen Peck, Shmuel Oren, and Robert Wilson

Transmission is playing an increasingly vital role in the modern electric power system. The transmission network is critical for system reliability and security by physically connecting geographically dispersed regions; at the same time, it enables economies of scale in generation plants, economies from pooling diverse demands and supplies, and economies from maintenance coordination. As competition is introduced through industry reforms around the world, transmission assumes new strategic importance in supporting market trading between individual buyers and sellers. Despite the widespread experience of electricity restructuring during the past decade, important issues remain unsettled concerning the best way to organize transmission to support reliability management and market trading. Some related issues are documented in a recent Federal Energy Regulatory Commission (FERC) order regarding the formation of regional transmission organizations.[1]

Underlying these issues is the fundamental reality that with the present technology, electricity markets are inherently incomplete, and the real-time dispatch of generation and transmission resources is most effectively managed by a central system operator. This suggests the necessity of a hybrid market architecture with multiple settlements of a sequence of decentralized forward markets and a cen-

  PII S1040-6190(00)00146-9

tralized spot or real-time market. While forward markets generally promote market competition and efficient price discovery (*ex ante*), the spot market and system operations are complementary functions, which are coordinated through the system operator to ensure system reliability and transparent price settlement (*ex post*). Therefore, we consider the scope of decentralized market trading to be limited to the forward markets. Within this framework, alternative market designs can be differentiated by 1) the comparative scopes of forward markets and the spot market, 2) the adjustment mechanism that provides the congestion management interface between these two fundamentally different types of markets and the financial risk management instruments to ensure market liquidity, and 3) the methods for handling the procurement and use of ancillary services. This article, however, shall focus on the first two issues and ignore the issue of ancillary services. In an unbundled market design, as in the California power market, the scope of forward markets is broad, and the adjustment mechanism is critical for the overall market performance. In a consolidated design, as in the Pennsylvania-New Jersey-Maryland (PJM) power pool, forward markets are of secondary importance, and the central pool rules largely obviate an elaborate adjustment mechanism.[2]

Central to the issue is the establishment of a set of tradable transmission rights. Tradable transmission rights facilitate energy trading in forward markets independent of real-time system operations. The definition of transmission rights, however, is complicated by the loop flow problem, which is ubiquitous in interconnected electric power networks. Basically, the problem is that in an electricity network, power flows along parallel paths dictated by physical laws rather than the contract path, creating widespread externalities whose complexity grows with the network size. In the absence of an appropriate mechanism to allocate transmission capacity, individual traders are unlikely to take into consideration the effects of power flows that diverge from the contract path.[3] When these traders do not confront the true opportunity costs of congestion and resistive losses that are imposed on others, the market outcome is infeasible or inefficient. The resulting economic costs from market fragmentation and missing markets can be significant.

Economic costs resulting from market fragmentation and missing markets can be significant.

This article examines ways that a system of flow-based transmission rights can be implemented. For expositional purposes, we follow the approach proposed by Chao and Peck.[4] Essentially, a system of flow-based transmission rights builds on the simple principle that it is desirable to match the scheduled transactions and actual power flows as closely as possible. The system adopts a trading rule that embodies the power transfer distribution factor (PTDF) to translate the physical effects of each energy transaction into requirements of transmission rights and transmission loss coverage. Chao and Peck[5] demonstrate the separation principle that with flow-based transmission rights, spontaneous trading in the separate markets of transmission and energy can achieve efficient resource allocation and price discovery. This theoretical prediction is supported by experimental results conducted in a laboratory setting.[6] As the North American Electric Reliability Council (NERC) is developing a flow-based transmission reservation and scheduling procedure, we believe that flow-based transmission rights will facilitate an efficient market design that is consistent with the new procedures proposed by NERC. In its forward markets for congestion management, the California independent system operator's procedures for allocating interzonal transmission capacity are essentially flow-based.[7] If the ISO's calculation of the flows implied by its scheduling coordinators' (SCs) proposed transactions indicates congestion on the interzonal transmission links, then it uses the adjustment bids SCs submitted with their initial preferred schedules to allocate the available capacity among those valuing transmission most highly, and for each transfer between zones it charges a price that is the marginal cost of the adjustments required. The ISO also auctions transmission rights that ensure priority in scheduling flows on interzonal links and refunds of the prices.

I. Basic Principles of Transmission Rights

Transmission rights are so fundamental to an efficient design of competitive electricity markets that their definition must be an integral part of market rules and should not be designed by private commercial entities afterward. The specification of transmission rights, however, is complicated by externalities due to loop flows. Since the actual power flows in an electrical network observe the physical laws known as Kirchhoff's Laws, the power flow paths generally diverge from the intended delivery paths, known as contract paths. These parallel flows, or loop flows, can cause the apparent costs of running generators to diverge from the real costs, and make it difficult to determine the available transfer capabilities of the transmission system. This leads to misalignment between the private cost and the social cost in electricity transactions and causes a potentially costly dislocation of resources in the power market. The inability to account for externalities due to parallel flows has been responsible for the increased use of transmission loading relief (TLR) procedures in North America to ration transmission services. Increased reliance on these administrative procedures for curtailments adversely impacts system reliability.

A market-based solution to this externality problem is to issue a set of well-defined transmission rights that internalize these effects. A market for these rights enables the external effects associated with a transaction to be incorporated into private purchasing and sales decisions. In essence, a transmission right is a property right that allows its holder to access a portion of the transmission capacity. Generally, a property right consists of three components: 1) the right to receive financial benefits derived from use of the capacity, 2) the right to use the capacity, and 3) the right to exclude others from accessing the capacity. Since the transmission schedule is centrally controlled, transmission rights can be defined as any combination of these three components. There are three possibilities.

Transmission rights are so fundamental to efficient design that their definition must be an integral part of market rules.

The first possibility is based solely on financial benefits, and this is generally known as the *financial-right approach.*[8] Financial rights—also known as passive rights[9]—provide market traders an instrument for constructing financial hedges as part of long-term energy contracts. The second approach combines financial benefits with capacity reservations or scheduling priority and is called the *capacity-reservation approach.* For most purposes, this approach provides adequate assurance of access to the network. The third approach includes all three components and is known as the *physical-right approach.* Paul Joskow has argued that the right to withhold access would contribute to market power without offsetting benefits.[10] In addition, if the system operator is excluded from accessing withheld capacity, this approach may reduce system reliability and security. In practice, however, and in conformity with FERC requirements, the system operator must take complete control of all transmission capacity in real-time, and thus withholding can be disallowed. Therefore, we can safely exclude the third component (exclusion) from the definition of transmission rights.

The definition of transmission rights depends on how transmission capacity is specified and measured. There are two common ways to specify the transfer capacity of the network. One way is to compute the point-to-point transfer capabilities, and the other is to specify the power-flow-carrying capacity for each link of the network. The point-to-point definition is rooted in what has been

  PII S1040-6190(00)00146-9

commonly known as the contract path approach. However, the transfer capability between any two points in a network changes continuously as the pattern of power flows shifts; therefore, it needs to be updated constantly. In contrast, the capacity of each link or flowgate is determined by physical factors associated with the link (e.g., thermal limit, voltage stability, and dynamic stability) and is generally insensitive to the power flow pattern. Each power transfer requires approximately a constant fraction (known as the power transfer distribution factor) of the capacity of each link in the network.

By combining the options discussed above, transmission rights can be defined in four possible ways: 1) point-to-point financial rights, 2) flow-based financial rights, 3) point-to-point capacity reservations, and 4) flow-based capacity reservations. For example, Hogan introduced a point-to-point financial right approach that is now implemented by the PJM power pool.[11] Richard D. Tabors introduced a form of point-to-point capacity reservations.[12] Chao and Peck introduced flow-based transmission rights and developed a theory that can be applied to financial rights, capacity reservations, or physical rights.[13] In fall 1999, the California Independent System Operator (CAISO) auctioned tradable annual capacity reservations in the form of "firm transmission rights" (FTRs) that included refunds of usage charges and scheduling priority. CAISO's FTRs are flow-based rights defined on the major links in the transmission system; however, the power network in California is aggregated into zones while ignoring interzonal loop flows. Consequently, CAISO defines FTRs only on the interzonal links in its transmission network. Presently, CAISO uses a network model that appears to be "radial" between the zones, but its definition of transmission rights and congestion charges enables looped networks to be handled as well.[14]

Flow is more fundamental than point-to-point as a basis for defining transmission rights.

The current debates about tradable transmission rights have focused on the comparative advantage between the flow-based approach and the point-to-point approach.

II. Advantages of Flow-Based Transmission Rights

Flow-based transmission rights, or flowgate rights for short, have several attractive features. The most important feature is that flowgate rights can be defined independently of the pattern of power flows. The feasible quantity of flowgate rights on each link is not sensitive to the topology of the network or to varying load conditions. The market value of such rights varies, of course, but the physical availability of transfers does not; hence, the quantities of available rights do not need to be frequently re-evaluated with respect to simultaneous feasibility constraints (as in the PJM power pool) and they are, therefore, relatively stable over time. Essentially, each point-to-point right is equivalent to a portfolio of flowgate rights, but a flowgate right cannot be decomposed into point-to-point rights in general. Therefore, flow is more fundamental than point-to-point as a basis for defining transmission rights. In addition, since flowgate rights are more specific to transmission assets, they are more tangible to the owner, providing investment incentives to those who might build transmission.

The flow-based specification of transmission rights is consistent with the existing NERC protocols for transmission load relief (TLR) and with the direction of the current NERC proposals for transition to a flow-based alternative that would align transmission reservations and energy schedules to actual flow and existing transmission loading relief procedures.[15] Both the current and proposed protocols rely on calculation of flows on congested flowgates using distribution factors derived from physical laws. Defining rights directly in terms of link capacity simplifies the process of curtailing transactions, or economic settlements among transactions competing for capacity on congested flowgates. Furthermore, the PTDF calculations that underlie the TLR protocols can also be used to calculate congestion charges and financial payoffs to rights from nodal or zonal prices.

This is also essentially the mechanism that CAISO uses when it manages interzonal congestion on its system. The "flowgates" are defined as the interzonal links in the transmission system, with flow limits and capacity rights defined on these links. The PTDF are implicit in the admittance matrix of the DC power flow approximation used to calculate the flows that each scheduling coordinator's transactions would place on each interzonal interface. CAISO adjusts the SCs' schedules to allocate available capacity to those valuing it most, and sets the congestion charge for use of an interzonal link to the marginal value of capacity on that link. Because of the compromises that led to the California power market's structure, CAISO does not charge for marginal transmission losses. Each generator is responsible for supplying its locational share of the "scaled marginal losses," so marginal losses are not included in the congestion charges nor in the definition of FTRs.[16]

An important feature of flowgate rights is that only congested links require financial settlement. In a congested meshed network even one congested link will cause each pair of zones or nodes to have different prices, and therefore require financial settlement of each point-to-point right. By contrast, when rights are defined in terms of flowgate capacity, only rights corresponding to the congested flowgates are entitled to financial compensation. Typically, the number of flowgates prone to congestion is relatively small compared to the number of nodes; hence, in a flow-based system the number of rights needed to provide each user the ability to hedge

against commercially significant congestion risk and secure scheduling priority is relatively small. Keeping the number of forward instruments small enhances market liquidity and improves efficient trading of risk management contracts. Moreover, once the prices of flowgate rights are known, it is straightforward to derive nodal energy prices from the hub price. But the converse may not be true. The nodal prices alone do not provide sufficient information to determine the flowgate prices. Therefore, flowgate rights offer greater price transparency than point-to-point.

Another important feature of flowgate rights is that they can be issued by the ISO as options and their values are never negative. If the flowgate for which a right has been issued is not congested in the direction corresponding to the right, then the right has no value; however, under no circumstances will its holder be liable for payment. Consequently, flowgate rights entitle holders to shares of the congestion revenue on the corresponding flowgate (if it is congested), but they do not entail any financial liability if its holder does not produce physical flow that matches the right. By contrast, point-to-point rights can result in a financial liability that can only be offset through a physical transaction that matches the right. In other words, a point-to-point right can have a negative value, which is quite prevalent in a meshed network with loop flow. The nonnegativity of flowgate rights stems from the fact that they represent a forward contract entitling (but not obligating) their holder to enjoy a physical characteristic of a transmission line—its ability to carry flow in a given direction. A point-to-point right, on the other hand, can be viewed as a portfolio of "short" and "long" link-based forward contracts needed to support a point-to-point transfer of power as determined by the PTDF. Such a portfolio can have either positive or negative value.

Underwriting point-to-point rights that have negative value poses commercial complications, yet is essential for fully hedged utilization of the network capacity because the number of rights that can be issued between different pairs of nodes are interdependent. In other words, the trading potential between two points may be greatly diminished unless a point-to-point right with negative value is underwritten. On the other hand, the available number of flowgate rights on a link is determined only by the contingency-adjusted flow constraints on that link independently of the rest of the network.

Another advantage of the flow-based approach occurs in the context of transmission congestion relief protocols across multiple control areas, as suggested in Cadwalader *et al.* and Oren and Ross.[17] There is general agreement that a market process for inter-regional transmission load relief is desirable, and indeed NERC is moving in this direction. In such a scheme, a flow-based approach enables a control area operator to account for the economic impact of transactions on other control areas due to loop flows.

To illustrate our points, see the simple example in **Figure 1**. In this network, we assume that nodes 1 and 2 are supply nodes and that node 3 is a demand node, which is selected as the hub. The transmission capacities for lines (1,2), (1,3), and (2,3) are 100 MW, 300 MW, and 220 MW, respectively.[18] For expository purposes, it is assumed that these transmission lines have identical electrical characteristics (i.e., line impedance) and, initially, that transmission losses are negligible. The power flows always follow Kirchhoff's laws, which imply for the present case that for any amount of power transferred from node 1 to node 3, two-thirds of it flows directly through link 1→3, and the remaining one-third flows through links 1→2 and 2→3.[19] Similarly, for power transferred from node 2 to node 3, two-thirds flows on link 2→3 and one-third flows on links 2→1 and 1→3.

Under a point-to-point approach, this setup calls for two types of transmission rights: TR1 from node 1 to node 3 (the hub), and TR2 from node 2 to the hub. (There could also be TRs in the reverse direction but the demand for such TRs in this case is purely speculative.) **Figure 2** illustrates (as the shaded area) all the possible combinations of TR1 and TR2 that would meet the simultaneous feasibility constraints under the assumption of a lossless DC approximation.

First, we observe that it is impossible to issue quantities of TR1 and TR2 that are independent of each other. In Figure 2, it appears feasible to issue 300 MW of TR1 and 300 MW of TR2, if they are considered separately; however, when these two types of transmission rights are issued simultaneously, the combination violates the capacity of line 2→3. Second, to ensure that the TRs give the holder the right to exercise them (and obtain a payoff) only when the payoff is positive, but not the obligation to use them when the payoff is negative, the number of TRs must be limited to no more than 300 MW for TR1, and similarly for TR2. To see this, suppose that 350 MW of TR1 is issued; then a minimum of 50 MW of TR2 must be exercised to ensure feasibility, and the exercise of TR2 becomes an obligation. This implies that the price of TR2 could be negative. This negative price entails an obligation for the holder of TR2 to pay for counterflow and could increase the cost of implementing a property right system. On the other hand, if prices must be kept nonnegative, the issuance of transmission

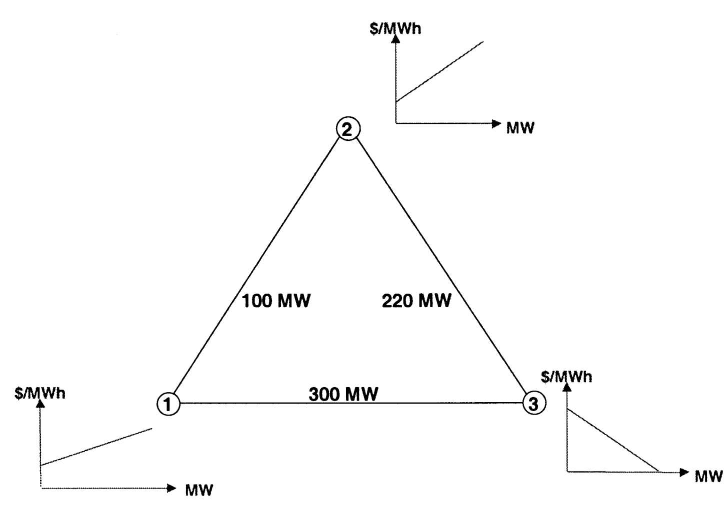


Figure 1: A Simple Three-Node Network

PII S1040-6190(00)00146-9

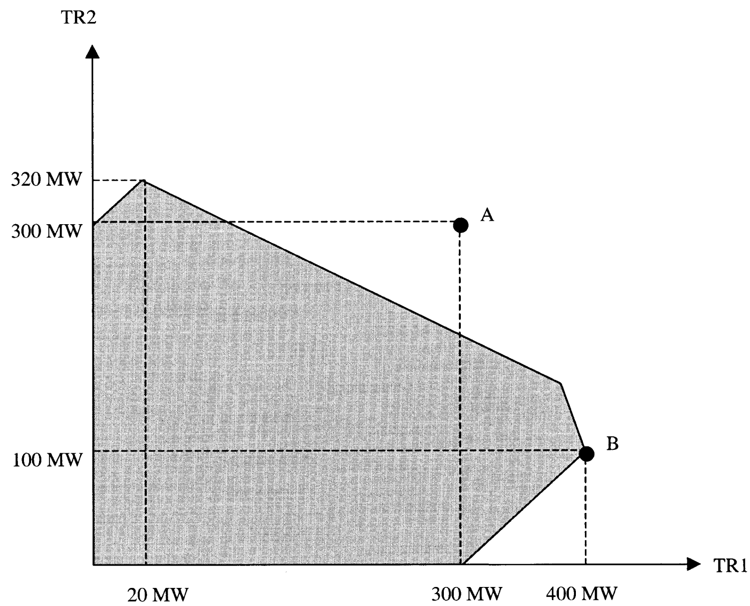


Figure 2: Simultaneous Feasibility Condition for Point-to-Point Transmission Rights of Network in Figure 1

rights must be limited, and the transactions cannot be fully hedged. This suggets that , a point-to-point transmission-rights system is inherently limited and cannot fully support a decentralized market design. A flow-based transmission rights system does not have such limitations.

However, in a longer-term context in which the grid configuration and the PTDFs are likely to change due to new investments in transmission, point-to-point rights for firm transmission between locations in a looped grid raise more fundamental issues. The basic consideration is that a transmission right defined point-to-point includes insurance against changes in the grid configuration and resulting changes in the PTDFs. That is, since it is defined *independently* of the grid topology, such a transmission right includes what might be called "PTDF insurance." If PTDF-insured transmission rights were issued for a duration of several years in an era when the grid topology is likely to change significantly, the cost to other market participants or to the ISO of fulfilling the obligations inherent in this insurance could be very large, and might have a substantial impact on the ISO's uplift charges in later years. Even if those investing in new generation might benefit substantially from multi-year PTDF-insured transmission rights, it seems clear that these benefits, costs, and risks must be assessed carefully in the context of a general plan for how the (potentially higher) revenues from auctioning PTDF-insured transmission rights would be allocated among market participants. With a flow-based approach, however, there may be economic incentives for private commercial entities to develop point-to-point hedging contracts that include PTDF insurance. This is an unresolved issue requiring further research.

III. A Competitive Electricity Market Design

This section describes in non-technical language the basic design of a wholesale market for electricity with a flow-based transmission rights system. First, one should be aware that each market design is in essence a creative contrivance, entailing intricate interactions of economic incentives and technical requirements. No matter how compelling the arguments may be, a market design should be carefully evaluated and tested in light of overall design objectives before implementation. There are at least five important factors worth serious consideration:

- System reliability
- Market efficiency
- Congestion management
- Market power mitigation
- Investment incentives

One of the immediate objectives of market design is to ensure system reliability. In the short term, keeping the lights on is commonly viewed as the most important indication of a successful restructuring. In the long run, market competition is expected to improve reliability through financial incentives. Overall, improvement in market efficiency is the primary objective of market design. However, with a poor implementation plan, electricity markets may lack liquidity during the transition, yielding erratic price swings; or become fragmented, resulting in restricted trading opportunities. Congestion management is critical to ensure efficient utilization of transmission capacity, thereby improving liquidity and reducing fragmentation. Electricity markets are particularly prone to local market power. With a sound design, the market can avoid traps of "load pockets," or remove "corners" that enable manipulative use of market power. Finally, an important goal of competition is to obtain accurate price signals and investment incentives to guide transmission expansion.

Several significant features that distinguish the electricity industry from others contribute to the structure of market institutions:

- Nonstorability
- Intertemporal and random variability of demand
- Necessity of balance in an interconnected transmission network
- Direct connections to customers
- Capital intensity and economies of scale

The demand for electricity varies continuously and unpredictably from hour to hour, day to day, and season to season. Practically, however, electricity cannot be stored. In an interconnected network, the demand and supply of electricity must be balanced instantaneously over time at every point of the network to maintain frequency, voltage, and system stability, and to avoid power outages. Unlike other types of networks, electricity in an AC electric transmission network flows in directions determined by physical laws rather than by contracts, and therefore is more difficult to control. A local variation of demand or supply of electricity affects power flows throughout the interconnected network. An equipment failure in one part of the network can cause the entire system to collapse. Efficiently meeting a new demand may involve coordinated adjustments of generators located far from the source of the demand. Therefore, the transmission system does not provide a simple physical connection between electricity generation and consumption facilities; it also involves active coordination of generating units dispersed throughout the network to meet variations in demand and supply.

Due to these technical characteristics, electricity markets are inherently incomplete. The centralized organization afforded under the traditional structure of vertical integration facilitated the essential task of coordinating efficient system operation and balancing the supply and demand of electricity continuously in response to changing system conditions. The challenge now is to design a market organization that can accomplish system coordination in transmission without compromising the opportunities for market competition in energy. This article focuses on the basic market architecture consisting of a set of decentralized forward markets that are separate from a centralized spot or real-time market conducted by a system operator, whose main responsibility is to manage the system operations and reserves to maintain reliability.[20]

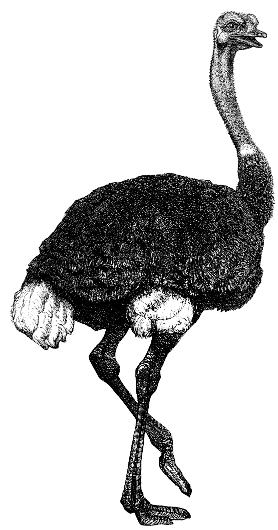
A. Forward Markets

Among the myriad of possible organizational forms for forward markets, the ones most common in commodities markets are *bilateral exchanges.* Those organized as "rings" or "pits" depend on oral outcry of bids and asks (usually by brokers acting for traders), whereas others use computerized bulletin boards to post offers. Those that depend on market makers to establish prices are conducted by specialists who clear orders from a book, or by dealers who post bid and ask prices. Market makers are usual where it is important to sustain intertemporal continuity of prices and reduce volatility, and typically they trade for their own accounts and maintain inventories. Market makers in the energy industries often play an important role reconciling differences among short- and long-term contracts, and more generally providing a variety of contract forms and auxiliary services, but rarely can one take an exposed position or accumulate inventories.

  PII S1040-6190(00)00146-9

Compared to the other organizational forms, the most salient distinction of bilateral markets is the continual process of trading, with contract terms and prices unique to each transaction. The experimental and empirical evidence indicates that, in general, bilateral markets are not less competitive or efficient than exchanges or pools, but often they are less transparent and prices are more widely differentiated. Among those with market makers, further distinctions are "product differentiation" represented by the variety of contracts and terms tailored to individual customers, and maintenance of some degree of price continuity.

Exchanges represent another important organizational form that complements bilateral markets for forward trades. A *power exchange* is a central market that establishes a uniform clearing price for standardized contractual commitments. It offers several advantages and also brings some disadvantages compared to bilateral markets. The uniform clearing price has some minor potential to realize the last iota of the gains from trade, but often the motives are more practical.[21] For a critical commodity like electricity, there is also a perceived advantage in establishing an official exchange with minimal transaction costs, open access for all traders, transparency that enables regulatory and public scrutiny, and countervailing power against private market makers with influence sufficient to extract some portion of the potential rents. The disadvantages lie in reliance on restrictive contract forms and inflexible procedural rules. Further, if the governance structure is inadequate, there is potential to impose restrictive procedures that are more convenient for administrators than traders. In addition, most exchanges rely on private bilateral markets for auxiliary services such as financial contracts to hedge prices. Exchanges for con-



tracts with terms longer than day-ahead are usually thin and illiquid, so they may be confined mostly to short-forward transactions with some supplemental provisions for bilateral trading of longer forward contracts.

A primary consideration to ensure the independent operation of forward markets is a proper definition of transmission rights. These rights define not only the spatial dimension of the commodity traded but also the transaction terms. To assure full independence of forward markets, these transmission rights should initially be distributed through annual auctions so that they can be traded along with energy in forward markets.

B. Flowgate Rights

Flowgate rights (FR) grant the holder a capacity reservation or scheduling priority for using specific transmission links or flowgates. The numbers of these rights required to complete each transaction are determined by a trading rule based on the power transfer distribution factors derived from Kirchhoff's laws. Essentially, these transmission rights are defined in such a way that they track the predicted physical flows of power on transmission links or flowgates. The total number of rights issued for each flowgate is determined by the flowgate capacity. Since the flowgate capacity does not depend on the entire network configuration, the number of rights issued can be relatively stable over long periods of time, even though their market value and the number of rights needed to fulfill a transaction could vary with the flow pattern and the network configuration.

The holder of a transmission right can use the right by scheduling power transactions; otherwise, the scheduling priority expires and the right reverts to the system operator, as required by FERC rules. Therefore, these transmission rights differ from physical rights in that they cannot be withheld to prevent others from accessing the unused transmission capacity. In practice, we may assume that by design, the terms (at least, the scheduling priority) of all unused transmission rights expire when the forward markets

  PII S1040-6190(00)00146-9

close and the spot market begins. In other words, the system operator can take full advantage of the unused network capacity for real-time dispatch. In this way, there are no financial incentives for traders to withhold transmission rights, nor to exercise market power using transmission rights in forward markets.

The trading rule can be viewed as a codification (albeit an approximate one) of the impacts of an electric power transaction on the physical flows throughout the network. The codification need not be precise for the market mechanism to capture most of the potential benefits of market competition. The difference between the estimated power flows based on the rule and the actual flows can be reconciled in the spot market.

As mentioned above, these flowgate rights can be interpreted as physical rights that cannot be withheld, or equivalently as financial hedges with assigned scheduling priority. The first interpretation is closer in spirit to the traditional definition of tradable property rights, which emphasize price determination through market trading. This definition is compatible with a system operator whose role is limited to prescribing trading rules that specify the number of rights that are needed to support any bilateral energy trade and to monitoring compliance with these trading rules. The latter definition is compatible with the current operational regime where real-time congestion is relieved through central dispatch of out-of-merit generators or energy adjustment bids, whereas the transmission rights define a financial entitlement to sharing congestion revenues. Under a system of flow-based rights, energy transactions can be hedged against congestion risk by acquiring a portfolio of flow-based rights that replicates the flow induced by the transaction according to the power transfer distribution factors. In reality, only rights on

congestion-prone flowgates will be actively traded and traders will hedge most of the congestion risk by acquiring rights on the congestion-prone flowgates that carry the bulk of their flows.

It is important to note that transmission rights provide the holders both a financial hedge and scheduling priority for power transfer across the network. A power contract with full coverage of transmission rights for all the links required by the trading rule is protected by its scheduling priority on every transmission link or flowgate; thus it constitutes a firm contract. In practice, only a relatively small number of links are likely to become congested. A power contract with a partial coverage of transmission rights on these links is practically firm, but there is always a risk that the uncovered links might become congested in the spot market.

The scheduling priority bundled with the financial right takes effect when real-time congestion cannot be fully resolved via energy adjustment bids due to insufficient bids, or when ties and schedules must be curtailed through a rationing scheme. In such situations a financial right with scheduling priority has the force of a physical right. However, since scheduling priority can only be claimed by scheduled transactions, withholding of capacity by owners of such priority is impossible.

Combining the advantages of both financial and physical rights, the flow-based transmission reservation approach described above facilitates liquidity and efficient risk management. Like physical rights, the flowgate rights are tradable under stable market rules, and therefore provide incentives for innovative contracting. Like financial rights, flowgate rights give the system operator full scheduling flexibility in real-time operations, while enabling traders in forward markets for energy to secure physical delivery without the threat of transmission withholding. Incidentally, this feature removes the market power problem for transmission in forward markets while providing system operator additional flexibility in real-time dispatch.

B. The Spot Market

The spot market is structurally a residual market settled after all the forward transactions of energy and transmission rights are closed. The system operator is obligated to implement the transactions concluded in forward markets and then manage spot trades along with adjustment bids and reserves that are needed for unexpected congestion or imbalance conditions. All trades in the spot market are settled *ex post* at a uniform market-clearing price. In calculating the spot prices for energy and transmission rights, the system operator performs centralized optimization of power flows and schedules that takes account of capacity availability, minimum generation requirements, ramping rates, transmission constraints and losses, as well as contingency requirements.[22] The spot prices are obtained as the "shadow prices" on a system balance constraint in an optimization program, which takes as input information from 1) the offers and bids from spot traders, 2) the adjustment bids, and 3) the reserves price schedules.[23] An apparent advantage of centralized system operation is the capability to closely coordinate activities in these three complementary areas.

The spot market provides an important adjustment mechanism linking the decentralized forward market and the centralized real-time operation. The real-time dispatch may deviate from the contracts traded in the forward market for at least two reasons. First, the trading rule may be based on an approximation of the physical system. Second, uncertain events (e.g., loss of generation units or transmission lines, or higher demand than anticipated) may occur after the forward market is closed. The adjustment mechanism can be viewed as mutual insurance that ensures reliability as well as efficiency. The system operator acts as the adjuster, settling payments among the market participants.

The basic objective is to ensure the integrity of forward market trading and real-time system operation. Ideally, priority insurance offered through franchise auction would be an adjustment mechanism that offers the system operator both the flexibility and incentives to enhance system reliability and efficiency.[24]

In the spot market, all unused transmission rights, even if honored financially, lose the privilege of scheduling priority.[25] The financial settlements for these rights are based on the system operator's security-constrained optimal dispatch computation. Instead of the settlement being based on nodal price differences as with point-to-point rights, the flow-based rights are settled according to the computed "shadow prices" on the congested links, which are mathematically related to the "nodal prices." These shadow prices define the marginal values to the system of an incremental unit of capacity on the corresponding interface (the shadow price on an uncongested interface is always zero); hence, the shadow price determines the congestion charge to transactions creating flow on that interface (in proportion to the flow), the payment per MW of transmission reservation, as well as the payment to out-of-merit generators producing counterflow on the congested flowgates. From a practical point of view, it may be more convenient to set congestion charges to transmission users and payment to generators for counterflow on a point-to-point basis while using the link shadow prices for the flow-based rights settlements. Such an approach has some important advantages and it is the basis for the way California settles charges for congestion.[26]

The above construction can accommodate a broad range of market design options in which the degree of decentralization is largely determined by the span of the forward markets for energy and transmission rights relative to that of the spot market. An unbundled market design will be conducive to large and active forward markets, while a consolidated market design is likely to be dominated by the spot market with limited use of tradable transmission rights.

IV. A Numerical Example

To illustrate how the basic principles work, we consider a simple numerical example. We assume a market structure with two separate institutional entities: a system operator (SO) and a power exchange (PX). Basically, the SO manages the central dispatch and spot market, while the PX facilitates the forward markets. In the centrally coordinated dispatch process, the SO continually balances electricity supply and demand requirements by scheduling generation to meet electricity demand. In real-time, all electricity must be traded through the spot market. Using the submitted offers and bids, the SO calculates a market clearing spot price to match supply and demand for each hour period during the day.

Energy and transmission rights are traded in forward markets. The transmission rights are initially distributed through annual auctions. SO may maintain flowgate right reserves to meet contingency conditions and to avoid infeasibility in local markets, which may lead to cornering and market power.

We consider a simple six-node transmission network without transmission losses. As shown in **Figure 3**, the network is divided into two interconnected zones: North and South. The northern zone, which consists of nodes 1, 2, and 3, and the southern zone, which consists of nodes 4, 5, and 6, are connected through two transmission interties 1–6 and 2–5. We assume that nodes 1, 2, and 4 are supply nodes, and nodes 3, 5, and 6 are a demand nodes. The physical transmission capacities for interties 1–6 and 2–5 are 200 MW and 250 MW, respectively.

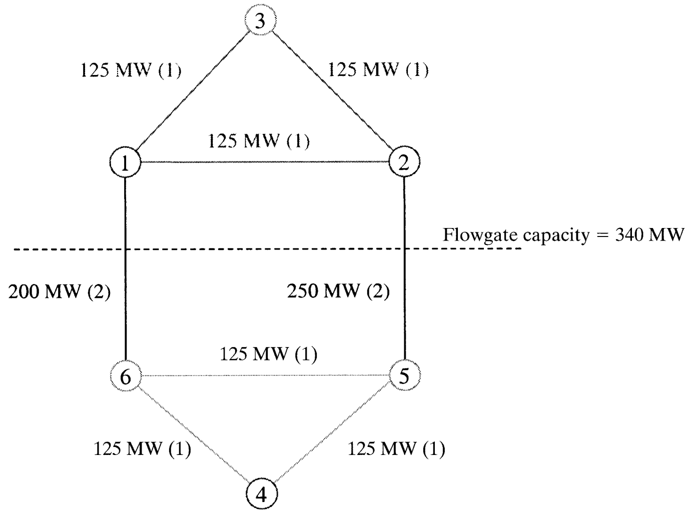


Figure 3: A Six Node Electric Network (Numbers in Parentheses Represent the Impedance of the Line)

System security necessitates additional constraints on power flows. We assume that based on a contingency analysis, the total power flows must not overload these lines in the critical event of losing 110 MW of line capacity. To meet this security constraint, a flowgate is shown in Figure 3. Then the security constraint can be characterized as a nomogram with a maximum flow constraint of 340 MW on this flowgate. In other words, as long as the total flow on these two lines does not exceed 340 MW, the system can survive the contingency of losing 110 MW line capacity.

A set of transmission rights can be defined to track the power flows in each direction of every line or flowgate. Therefore, there is a maximum of 18 flowgate rights that the SO may demand to execute transactions. Holding one of these rights implies the scheduling priority for transferring power in a specific direction on a congested link or flowgate. An important feature of power systems is that power flows in opposite directions on a line will offset each other. This bi-directional nature of the power system is recognized in our definition of transmission rights. We designate node 6 as the hub. The power transfer distribution factors can then be estimated, and some are illustrated in **Table 1**.

A. Forward Markets for Energy and Transmission Rights

As mentioned above, there are many possible organizational forms for forward markets. To illustrate how the process may practically work, we consider a situation where the trading of energy and transmission rights takes place in an exchange using simultaneous

  PII S1040-6190(00)00146-9

Table 1: Power Transfer Distribution Factor of Network Shown in Figure 3

Selected Link/ Flowgate	Power Injected at Node 1	Power Injected at Node 2	Power Injected at Node 3	Power Injected at Node 4	Power Injected at Node 5
1→6	0.625	0.5	0.5625	0.0625	0.1250
2→5	0.375	0.5	0.4375	−0.0625	−0.125
1→3	0.125	−0.1667	−0.5208	−0.0208	−0.0417
N→S	1	1	1	0	0

multiple-round auctions,[27] though continuous bilateral trading through brokers is allowed. The flowgate rights are initially supplied through annual auctions. These rights can be traded freely in forward markets. In addition, transactions that produce counterflows can also supply flowgate rights, though these transactions need to be covered by other flowgate rights. To ensure reliability, the system operator may purchase flowgate rights for transmission reserves.

At the beginning of the auction, the buyers and sellers tend to behave conservatively—submitting low purchase bids and offering high selling prices.[28] This is equivalent to the buyers understating their marginal benefits, while the sellers overstate their marginal costs. **Table 2** shows the specific assumptions about the demand and supply functions and the initial bids of the price and the electricity output (q in MW).

Suppose that according to the system operator's forecast, only the flowgate N→S is likely to be congested. Therefore, the system operator will require only flowgate rights N→S for all the trades from north to south in order to set the schedule for the next day. Note that the total supply of flowgate rights is not fixed, because energy traders can produce flowgate rights for N→S by generating counterflows from south to north.[29] The initial offers for the transmission rights are summarized by a linear function $(5/48)q$.

An optimal power flow model is employed to compute the quantities of power generation and demand that are committed each round using the energy and transmission bids as input data. The market clearing prices at different nodes represent the revealed marginal cost and willingness-to-pay at those nodes. For example, the market clearing price at node 1 is \$20/MWh + $0.05q$. **Table 3** shows the results of the first round of auction assuming that the marginal price for transmission rights is \$25/MWh. According to the modeling results, the generators at node 1 will commit 170 MW for a price of \$28.5/MWh [20 + (0.05)(170) = 28.5] and the buyers at node 6 (the hub) will take 165 MW at a price of \$53.5/MWh [70 − (0.1)(165) = 53.5]. The buyers at node 3 take nothing for the moment because their maximum bid, \$27.5/MWh, is higher than \$28.5/MWh. Table 3 also shows the physical flows of the 240 MW sale across the flowgate N→S, according to the DC-flow approximation of Kirchhoff's Laws. Based on these results, the network is not actually congested because the flows on each line are less than the corresponding line's capacity.

In auctions, bidders rarely volunteer private information. In this example, we examine the implications of these interim results based on the true demand and supply

Table 2: Assumed Market Demand and Supply Functions for the Example

Node	Function Type	Actual Function	Initial Bid
1	Marginal cost of supply	$10 + 0.05q$	$20 + 0.05q$
2	Marginal cost of supply	$15 + 0.05q$	$25 + 0.05q$
3	Inverse demand function	$37.5 - 0.05q$	$27.5 - 0.05q$
4	Marginal cost of supply	$42.5 + 0.025q$	$52.5 + 0.025q$
5	Inverse demand function	$75 - 0.1q$	$65 - 0.1q$
6	Inverse demand function	$80 - 0.1q$	$70 - 0.1q$

  PII S1040-6190(00)00146-9

Table 3: Results of the Initial Round of Auction for the Example

Location	Price ($/MWh)	Quantity (MW)
Node		
1	28.5	170
2	28.5	70
3	28.5	0
4	53.5	40
5	53.5	−115
6	53.5	−165
Link/Flowgate		
1→6		129.4
2→5		110.6
N→S	25	240

functions that are given. The potentially profitable output hidden at node 1 is calculated by examining the final-round offering at node 1. At the price of $28.5/MWh, the potentially profitable output there is 370 MW [10 + (0.05)(370) = 28.5]. This figure is 200 MW higher than the accepted offers in round 1. Similarly, the hidden demand at node 6 is 265 MW, which is 100 MW higher than the accepted bids in the initial round. Hence, a lower selling price than that offered in round 1 would lead to the sale of more power. In effect, this high price in the initial round puts pressure on the suppliers to lower the offered price and the demanders to raise their bids. At the same time, there is apparent pressure for the owners of flowgate rights to lower their asking prices, since there are still 100 MW of unsold flowgate rights.

The simultaneous bidding process of energy and transmission rights tends to converge to competitive equilibrium, until profitable trading opportunities are exhausted.[30] By encourage serious bids, well-designed market activity rules can expedite efficient price discovery.[31] At equilibrium, there are two zonal prices, and no profitable arbitrage opportunities remain. **Table 4** shows the equilibrium result in which flowgate N→S is congested with a price of $22/MWh. Selling electricity from the north to the south yields $48.5/MWh in revenue but costs $26.5/MWh for the electricity at node 1 plus $22/MWh for the transmission charge.

Considering transmission losses complicates the process. The above trading rule can be refined so that market trading incorporates both transmission congestion cost and transmission losses.[32] Transmission line losses depend on the square of the power transmitted, and thus the marginal loss increases with the power flow on the line. In this case, each new electricity transfer could affect the distribution of transmission losses throughout the entire network, and thus affect the losses other users of the transmission system sustain. To achieve economic efficiency, and thus the appropriate dispatch of generation and use of electricity, traders must pay for marginal transmission losses as well as congestion. Since marginal transmission losses are generally approximately twice as high as average transmission losses, a new type of economic rent is created, and an allocation rule is needed. In this case, the trading rule is augmented by specifying the compensation for average power losses and allocation of economic rent associated with the transmission losses.

Table 4: Results of the Final Round of Auction for the Example

Location	Price ($/MWh)	Quantity (MW)
Node		
1	26.5	330
2	26.5	230
3	26.5	−220
4	48.5	240
5	48.5	−265
6	48.5	−315
Link/Flowgate		
1→6		179.4
2→5		160.6
N→S	22	340

B. The Spot Market

A highly desirable feature of the spot market is to provide a transparent price determination and settlement process that supports flowgates rights in forward markets. Suppose that in the spot market, the system operator adopts the locational marginal pricing rules to process incremental or decremental bids for changes in electricity supply and demand. We illustrate how this works with a scenario in which the actual load at node 3 unexpectedly rises to 240 MW, 20 MW higher than the scheduled level. We assume that the demand cannot respond to price signals in the short run due to limitations in communication capability. Therefore, an adjustment is needed to maintain the system in balance. The adjustment bids can be drawn from a forward reserve auction or a spot market auction. For simplicity, we assume that the adjustment bids are mainly drawn from spot market bids, and that these bids are the same as the true supply functions in Table 2.[33]

Table 5 displays the results of an optimal redispatch for the scenario. We assume that the trading of flowgate rights on link 1→3 in the forward market yielded a zero price, but as a result of the unexpected demand increase, the real-time flowgate price rises to $10/MWh. Note that the generation at node 1, the low-cost supply, is reduced by 19.3 MW, while the generation at node 2, the expensive supply, is increased from 39.3 MW. This counterintuitive result is caused by the congestion on link 1→3. The reason is that each MWh of generation at node 1 would increase power flows on link 1→3 by 2/3 MWh, while generation at node 2 creates only one-half the power flow on link 1→3 as generation at node 1. Therefore, the optimal way to meet the higher load at node 3 is to produce more from the expensive source and less from the low-cost supply. As a result, the prices become dispersed in the northern zone. The spot price at node 1 drops to $25.5/MWh, the price at node 2 rises to $28.5/MWh, and the price at node 3 rises to $32/MWh, which includes the congestion charge for link 1→3.[34] The payments for individual increments and decrements are calculated on the basis of these spot prices. Under this scenario, the net adjustment payment is zero, as shown in Table 5.

Table 5: The Result of Spot Market Adjustment in the Example

	Forward Market		Spot Market		Adjustment	
	Supply/Demand (MW)	Price ($/MWh)	Supply/Demand (MW)	Price ($/MWh)	Increment/Decrement (MW)	Payment ($/hour)
Node						
1	330	26.5	310.7	25.5	−19.3	−493
2	230	26.5	269.3	28.5	+39.3	1118
3	−220	26.5	−240	32	−20	−640
4	240	48.5	240	48.5	0	0
5	−265	48.5	−265	48.5	0	0
6	−315	48.5	−315	48.5	0	0
Link/Flowgate						
1→6	179.4		184.5			
2→5	160.6		155.5			
1→3	123.5	0	125	10	1.5	15
N→S	340	22	340	21.5	0	0
Total					0	0

Similar adjustments can be made for other uncertain events, such as derating of line capacity and changes in power transfer distribution factors, as long as it is predetermined who will assume the financial consequences, the rights holder or the SO.[35] In general, however, such adjustments need not be revenue-neutral.

V. Discussion

Some objections have been raised against a flow-based approach. In a generally positive article in *The Electricity Journal*, Steven Stoft was skeptical that this approach could be implemented in a decentralized manner.[36] To support his skepticism, he adduced two arguments.

First, Stoft explained that the tradable transmission right approach is complex. When one trader seeks to buy the transmission rights to support a trade, Stoft stated that the trader must buy some rights on every line in the network. Stoft suggests that this trading requirement seems much more burdensome than the competing approach of nodal pricing developed by Hogan.[37] Stoft cited Chao and Peck: "a potential disadvantage with the above market mechanism is its apparent complexity, because there could be a large number of transmission capacity rights in a real network, and thus the information-processing cost might be too high, rendering the approach impractical."[38]

Second, and more important to

  PII S1040-6190(00)00146-9

Stoft, was the issue of market power. If any one person owns all of a transmission line, Stoft explained, that owner possesses a monopoly over all users of the grid. Thus every bilateral trader that seeks to make one trade will be faced with the need to buy thousands of transmission rights, with each one purchased from a monopolist.

To mitigate the market power issue, Stoft suggested two options. First, universal divestiture could be required so that no transmission owner possesses more than 20 percent of any line. Stoft noted, however, that residual problems might remain because the demand for minor lines would be inelastic, and even diffuse owners might increase prices. Alternatively, Stoft suggested that line owners be required to sell the entire capacity of each line. Ultimately unconvinced by his own counterarguments, Stoft concluded, "a decentralized market may not be the only goal worth pursuing."

A. A Practical Implementation of a Flow-Based Approach

To address the above reservations, we now describe how a flow-based approach might be implemented practically. There would be five key steps:

- All flowgate rights will be initially distributed through annual auctions, like the one conducted in California.
- Once distributed, these rights can be freely traded in secondary markets, through bilateral trading or auctions, at privately organized markets.[39]
- Each day, the system operator will announce the flowgates that are likely to be congested the next day.
- The system operator will collect schedules from energy traders using the announced flowgate rights to set scheduling priority.
- The system operator will clear the spot market of energy, in conjunction with reserves and adjustments, calculating energy and flowgate prices.

In the annual auction, a major portion of the rights should be issued as regular flowgate rights with scheduling priority, but a small fraction could be issued as pure financial rights to provide the system operator scheduling flexibility to mitigate contingencies. All flowgate rights would be offered for auction. To facilitate the offering of point-to-point transmission services, a portion of the transmission rights could be made available in the bundled form of transmission system strips, or fixed fractions of the entire transmission system, while the remainder would be auctioned on a line-by-line basis. In this way, successful transmission system strip bidders could use some fraction of every line to provide point-to-point service immediately.

The bilateral market would operate for several months until the day before dispatch. This market would enable energy traders to enter into bilateral contracts for future delivery and to be assured of a known cost of transmission services. In this market, energy traders would be able to seek point-to-point transmission services at a competitively determined price from any of the transmission system strip holders or may, at their option, buy rights from each of the transmission rights holders. Complexity would be minimized for the energy traders, since transmission system strip holders would offer point-to-point transmission services. In addition, the existence of many transmission system strip holders would foster competition in transmission rights.

If an active market develops for the individual flowgate rights, the flowgate prices will become apparent; otherwise, they will be implicit in the point-to-point prices. Since individual transmission rights can be combined into point-to-point services, arbitrage will make the prices for point-to-point services consistent with the prices of individual transmission rights.

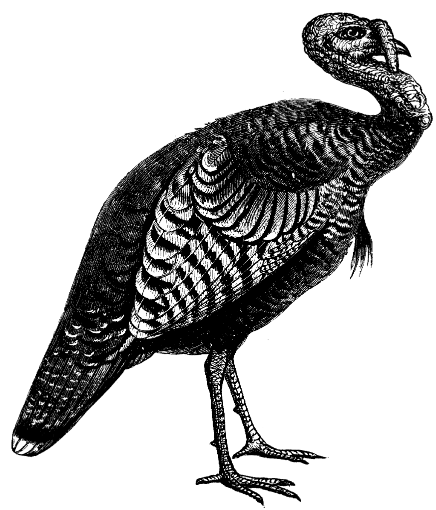
A power exchange can conduct the auction from one day before dispatch until one hour before dispatch. In this market, flowgate rights that have not been committed can enter the auction along with energy bids. The flowgate prices are determined by an algorithm that matches the node-specific bids in a way to maximize the social surplus.

As a result of arbitrage, the transmission prices from the bilateral market should be an unbiased estimator of the transmission prices determined in the power exchange market. For the same reason, these prices should also be unbiased estimators of the spot prices determined in the adjustment market. Since the uncommitted transmission capacity will be freely available to the system operator in real-time dispatch, arbitrage among different markets will ensure that any attempts to withhold transmission rights for the purpose of raising prices will be completely ineffective.

Although there is potentially a large number of flowgate rights, the system operation can be simplified further by using a fixed but small (say, up to 10) floating flowgate rights to set scheduling priority. Each day, the system operator can announce in advance a number of flowgates that are most likely to be actually congested. Although congestion may shift from one set of flowgates to another, the total number of congested flowgates is typically small. For flowgates that experience congestion unexpectedly, the system operator can price them in the real-time market *ex post* and settle the adjustments properly.

B. Transmission Rights in Practice

Some systems rely on distinctions among the priorities assigned to various contractual entitlements to transmission. In 1999, the California ISO auctioned annual contracts (which took effect in February 2000) for firm transmission rights for day-ahead interzonal capacity on that portion assigned to the ISO by the investor-owned utilities. These rights amount to about 50 percent of the ISO's average capacity available daily, and nearly 100 percent of a conservative estimate of the capacity available annually (that is, for the entire



year), net of contracts existing when the market was established in 1998 and expiring in later years.

These firm transmission rights differ from the "fixed transmission rights" issued by PJM, which are purely financial contracts entitling each owner to a refund of the difference in nodal prices on a specific point-to-point path. PJM's rights are allocated by an optimization in which bids for point-to-point transmission are used to simulate energy flows. Secondary markets for point-to-point financial rights are too thin to be viable, so PJM offers a monthly re-configuration process as a substitute. Financial rights suffice to meet FERC directives requiring each system operator to provide a means for customers to ensure "price certainty" if private markets for hedges against transmission prices are insufficient, as invariably they have been.

California issues firm transmission rights for each direction for each interface between zones, including import/export interfaces. These rights can be assigned or traded in secondary markets. The market designs described last year by Gribik and by Gribik, Kritikson, and Shirmohammadi, go beyond trading of the FTRs defined by the ISO.[40] In these designs, FTRs compete with sales by the municipal utilities of idle capacity to which they are entitled by existing transmission contracts (ETCs). Like PJM's hedges, firm transmission rights include financial components because they provide refunds of interzonal usage charges, although these refunds are nil in the specified direction if congestion is in the reverse direction (i.e., no credit is given for counterflows).

The analysis by Gribik in 1999 shows that the availability of tradable firm transmission rights enables a private market to consolidate energy and transmission in a way analogous to centralized markets like PJM. In this design, an optimization is used to maximize the gains from trade among suppliers and demanders, subject to spatial constraints on transmission and possibly also intertemporal constraints on ramp rates. The difference is that the transmission capacities are the ones implied by the supplies of ETCs and firm

  PII S1040-6190(00)00146-9

transmission rights offered in the market; that is, the private market allocates transmission *rights* rather than transmission *per se.* The resulting zonal prices for energy obtained at 7 AM in the day-ahead market are like nodal prices, since each is the sum of a global energy price and an imputed credit or charge for associated transmission rights, depending on whether it is an importing or exporting zone, as measured by rights. These transactions are firm in terms of both prices and quantities because each is protected by ETCs and/or firm transmission rights. Each customer can offer additional adjustment bids in the ISO's market, but these transactions are not firm until the close of the ISO's market at 1 PM. One private market might assume a dominant role, perhaps the PX in California, because other market makers can conduct their net trades through this central market to ensure that the full value of transmission rights and counterflows are realized; but arbitrage among private markets might suffice.

It is recognized that efficient allocation can be attained through separate market trading of energy and transmission rights.[41] Gribik, Kritikson, and Shirmohammadi[42] suggested that it is possible to conduct the private market for energy first, and then an auxiliary market for transmission rights afterward, thus preserving the existing format of exchanges like the PX currently designed to trade only energy. They modified the market design in Gribik[43] so that the forward market for energy closes with a uniform price for energy and an allocation of quantities transacted at that price by suppliers and demanders. Shortly after, each customer submits a bid offering the maximum price it is willing to pay to lock in its transaction in the energy market. This offer is interpreted as the price bid for protecting the transaction via ETCs or firm transmission rights acquired from those who offer rights in this market. The result is essentially the

same as a consolidated market: if the energy price is $20 and a supplier in one zone is matched with a demander in an adjacent zone, each paying $5 for locking in its energy transaction, then the net price is $15 for the supplier and $25 for the demander, exactly as if the zonal prices were $15 and $25 in the exporting and importing zones, corresponding to the clearing price of $10 for firm transmission rights from one zone to the other. The prices are determined from an optimization that allocates the supply of submitted rights to energy transactions so as to maximize the gains from trade between the two sides of the market, just as in a consolidated market for energy and transmission rights. The ISO's procedures determine whether suppliers and demanders must be matched for administrative purposes into pairs assigned to each right; if not, then it suffices to submit rights to the ISO in sufficient amounts to protect the aggregate interzonal flows. The California ISO requires precise matching of the particular injections and extractions associated with each right, but this is motivated mainly by the convenience of using existing software to apply the default adjustment bids of $9,000 and $4,000 per MWh to each ETC and firm transmission rights.

The transformation from a nodal or zonal system into a flow-based system is relatively straightforward. It can be accomplished by using the power transfer distribution factors as the exchange rates to translate transmission rights from one system to another without significantly affecting the existing market processes and institutions. After translating into a flow-based system, the main difference between the nodal- and zonal-based rights lies in the different numbers displayed in the PTDF matrix. Immediately, this will obviate the need for bid reconfiguration in PJM and rezoning in California (for intrazonal congestion problems).

C. Information Technology Lowers Transaction Costs

We believe that recent rapid technological advances should have settled the issue of complex-

  PII S1040-6190(00)00146-9

ity. In the past few years, rapid developments of advanced metering, two-way communications, and Internet-based information technologies have clearly set the trend for lowering market transaction costs. For instance, a system such as the Open Access Same-Time Information System (OASIS) will enable participants to complete market transactions continuously. Emerging private firms that offer brokerage services for bilateral markets have been urging FERC to adopt a decentralized approach based on flowgate rights.[44] Evidently, the private sector has already exhibited interest in transforming the electricity business into e-commerce.

As for other potential difficulties, a well-known folk theorem in economics suggests that in the absence of market failures (i.e., externalities in the present case), "whatever a central agency can do, a market can do better." A corollary to this theorem is that once the main cause of market failure is fixed, innovations in the marketplace can be relied upon for efficient self-organization. For instance, transmission brokers might emerge to simplify the trading of transmission capacity rights by aggregating them into transmission capacity contracts in ways that resemble the more pragmatic concept of transmission capacity reservation contracts stipulated by FERC in 1996,[45] or transmission congestion contracts defined in Harvey *et al.*[46] Aggregators might begin to bundle electricity and transmission capacity contracts over different time periods and contingencies into simple contracts that can be traded among common consumers and suppliers. For instance, as Vickrey[47] suggested in the case of air flight tickets, forward markets could be established for reservations of transmission capacity rights at various points in time before the actual dispatch. Obviously, such undertakings involve varying degrees of risk, and various risk management contracts will emerge. Nevertheless, we believe that at a minimum, a simple contract will be created that enables an electricity supplier to agree to deliver to a consumer a known quantity of power at a fixed price over a long period of time. After all, a competitive market is known to be surprisingly innovative in self-organizing for informational efficiency, and spontaneous exchange of information among self-interested individuals is the essence of Hayekian markets.

As for other potential difficulties, a well-known folk theory offers guidance.

  PII S1040-6190(00)00146-9

VI. Conclusion

The externalities due to loop flows in a transmission network represent a critical issue that must be resolved before competition can be successfully introduced into the electric power industry for long-term economic benefits. The main insight is that a system of flow-based transmission rights enables market-based congestion management for efficient energy and transmission markets. Further, once a system of tradable flowgate rights is established, the control of the transmission system is shifted from line owners to the market, in which the transmission charges are determined competitively without excessive complexity or monopoly power abuses. ■

Endnotes:

1. Federal Energy Regulatory Commission, Regional Transmission Organizations, Final Order No. 2000, RM99-2-000, Washington, DC, Dec. 15, 1999.

2. Centralized market organizations often emphasize short-run operational efficiency while maintaining system reliability and security, but neglect the demand-side participation essential for long-term economic benefits.

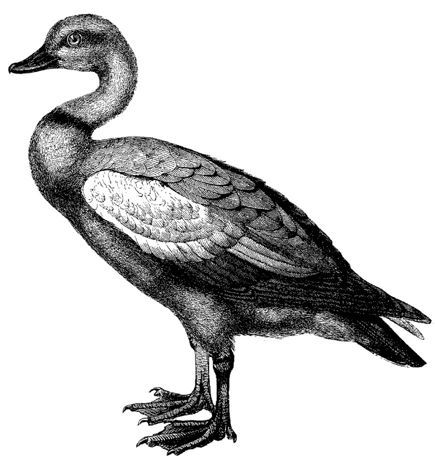
3. A contract path from node *i* to node *j* refers to the transfer of a certain amount of power from node *i* to node *j* along a path that is specified administratively without considering the actual flow of power.

4. Hung-po Chao and Stephen C. Peck, *A Market Mechanism for Electric Power Transmission*, J. REG. ECON., July 1996, at 25–59; Hung-po Chao and Stephen C. Peck, *An Institutional Design for an Electricity Contract Market with Central Dispatch*, ENERGY J., Jan. 1997, at 85–110; and Hung-po Chao and Stephen C. Peck, *Reliability Management in Competitive Electricity Markets*, J. REG. ECON., Sept. 1998, at 89–200.

5. Chao and Peck (1996), *supra* note 4.

6. Hung-po Chao and Charles R. Plott, *A Competitive Mechanism for the Distribution of Electric Power Through a Network and the Determination of Network Capacity: Experimental Tests and Performance*, working paper presented at California Public Utility Commission/ORA Seminar, San Francisco, Feb. 7, 2000.

7. Paul R. Gribik, George A. Angelidis, and Ross R. Kovacs, *Transmission Access and Pricing with Multiple Separate Forward Energy Markets*, 14:3 IEEE TRANSACTIONS ON POWER SYSTEMS, 1999, at 865–76. Harry Singh, Shangyou Hao, and Alex Papalexopoulos, *Transmission Congestion Management in Competitive Electricity Markets*, 13:2 IEEE TRANSACTIONS ON POWER SYSTEMS, May 1998.



8. William Hogan, *Contract Networks for Electric Power Transmission*, J. REG. ECON., Dec. 1992, at 211–42.

9. Shmuel Oren, *Economic Inefficiency of Passive Transmission Rights in Congested Electricity Systems with Competitive Generation*, ENERGY J., Jan. 1997, at 63–84.

10. Paul Joskow and Jean Tirole, *Transmission Rights and Market Power on Electric Power on Electric Power Networks I: Financial Rights* and *Transmission Rights and Market Power on Electric Power on Electric Power Networks II: Physical Rights*, mimeos, Massachusetts Institute of Technology, Cambridge, MA, and Institut d'Économie Industrielle (IDEI), Toulouse, France, Jan. 1999.

11. Hogan, *supra* note 8.

12. Richard D. Tabors, *A Market-Based Proposal for Transmission Pricing*, ELEC. J., Nov. 1996, at 61–67.

13. Chao and Peck (1996), *supra* note 4.

14. Paul R. Gribik and Dariush Shirmohammadi (Perot Systems Corp.), and James Kritikson (California Power Exchange), private communication, Aug. 1999.

15. Energy transfers among the control areas of the interconnection are traditionally scheduled using the contract path method. The limitations of this method are well known. In the past, when the volume of transactions was relatively small and the transactions were mostly between adjacent control areas, parallel flows were tolerable. As a result of deregulation, however, the energy transfers increase in both volume and geographical span, with the result that parallel flows begin to exert serious operational and economic impacts. See NERC, *Aligning Transmission Reservations and Energy Schedules to Actual Flows: Proposal to Mitigate Current Problems Associated with Transmission Reservation and Scheduling Practices In the Eastern Interconnection*, prepared by NERC Transaction Reservation and Scheduling Self Directed Work Team, Nov. 1998.

16. Scaled marginal losses are used to avoid the surplus that results from application of full marginal losses. That is, just as pricing transmission at the margin produces a surplus (congestion rent) that is paid to a transmission rights holder, losses charged or supplied at the marginal rate of loss would produce a surplus (marginal losses are approximately twice average/total losses). The "scaled marginal loss" approach provides an approximate locational price signal because the forward estimate of marginal rate of loss is scaled down equally at all locations to produce a loss factor that is intended to collect total system losses while maintaining a different loss factor at each location.

17. Michael B. Cadwalader, Scott M. Harvey, William W. Hogan, and Susan L. Pope, *Coordinating Congestion Relief across Multiple Regions*, Harvard Electricity Policy Group working paper (Oct. 1999); and Shmuel S. Oren and Andrew M. Ross, *Economic Congestion Relief across Multiple Regions Requires Tradable Physical Flowgate Rights*, working paper, Mar. 21, 2000.

18. In this article, prices are expressed in $/MWh, the power flows in MW, and the electricity outputs in MWh.

19. These fractions are known as the power transfer distribution factors. Roughly speaking, power flow is inversely proportional to the impedance along the path. Since the path along link $1 \rightarrow 3$ is half as long as the path $1 \rightarrow 2$ and $2 \rightarrow 3$, it has half the impedance. Therefore, the ratio of the power flows between them is 2 to 1.

20. Chao and Peck (1997), *supra* note 4.

21. For instance, in California the Power Exchange's price is used to settle some grandfathered contracts, and affects payments for recovery of stranded costs. Requiring the incumbent utilities to trade through the PX also makes it easier to monitor market power. In Britain first, and in Alberta still, hedging contracts used to mitigate the incentives of incumbents with substantial market power are based on the exchange price.

22. To hedge against the volatility of the spot price, buyers and sellers of electricity may enter into bilateral hedging contracts at a mutually agreed price. These contracts are financial instruments outside of the system operator's purview and do not affect the real-time dispatch by the system operator. Two common types of hedge contracts are the swap contract (two-way hedge) and option contract (one-way hedge).

23. In California, the adjustment bids may be obtained from the traders in the forward markets for congestion management, while the reserves price schedules may be obtained through separate forward auctions, as described in Hung-po Chao and Robert Wilson, *Multi-Dimensional Procurement Auctions for Power Reserves: Incentive-Compatible Evaluation and Settlement Rules*, mimeo, Electric Power Research Institute, Palo Alto, CA, and Stanford University, Stanford, CA, 2000.

24. Chao and Peck (1997), *supra* note 4, and Robert Wilson, *Implementation of Priority Insurance in Power Exchange Markets*, ENERGY J., Jan. 1997, at 111–23.

25. In an optimally dispatched system the spot value of the two types of rights are mathematically linked via the matrix of PTDF that specifies the fractions of power injected at any particular bus and taken out at a reference bus that flows through each flow gate. Specifically, the value of a point-to-point right from any given bus to the reference bus is the sum of the values of the congested flowgate rights multiplied by the corresponding PTDF. The value of a point-to-point right between any two arbitrary buses is given by the difference of the point-to-point right from the corresponding buses to the reference bus.

26. Gribik, Angelidis, and Kovacs, *supra* note 7.

27. An example is the Federal Communications Commission's auction for selling spectrum licenses.

28. Chao and Plott (2000), *supra* note 6.

29. However, if the counterflow is scheduled but not delivered, the seller could be subject to a severe penalty because of the impact on reliability.

30. Chao and Peck (1996), *supra* note 4.

31. Robert Wilson, *Activity Rules for the Power Exchange*, Report to California Trust for Power Industry Restructuring, Mar. 14, 1997.

32. Chao and Peck (1996), *supra* note 4.

33. In NordPool, there is no difference between the energy bids and the adjustment bids. In California, however, the traders can submit adjustment bids different from the energy bids. See Chao and Wilson (1999), *supra* note 23, for a description of how a reserve auction could provide a supply curve for adjustments.

34. These nodal prices can be easily calculated from the hub price and the flowgate prices using the distribution factors; however, given only the nodal prices, there isn't a unique set of flowgate prices.

35. The allocation of financial risks does not affect market efficiency, but should have a direct effect on the results of the initial transmission rights auction.

36. Steven Stoft, *Congestion Pricing with Fewer Prices than Zones*, ELEC. J., May 1998, at 23–31.

37. Hogan, *supra* note 11.

38. Chao and Peck (1996), *supra* note 4.

39. Edward G. Cazalet and Ralph D. Samuelson, *The Power Market: E-Commerce for All Electricity Products*, PUB. UTIL. FORTNIGHTLY, Feb. 1, 2000.

40. Paul R. Gribik, *Enhancing the Power Exchange's Day-Ahead Markets: Binding Results at 7:00 AM Prior to ISO Congestion Management, Trading of Transmission Rights, Developing Physically Feasible Schedules, Block Trading*, mimeo, California Power Exchange and Perot Systems Corporation, June 1999; and Gribik, Kritikson, and Shirmohammadi, *supra* note 14.

41. Chao and Peck (1996), *supra* note 4.

42. Gribik, Kritikson, and Shirmohammadi, *supra* note 14.

43. Gribik, *supra* note 40.

44. Federal Energy Regulatory Commission, *supra* note 1.

45. Federal Energy Regulatory Commission, Capacity Reservation Open Access Transmission Tariffs, Notice of Proposed Rulemaking, RM96-11-000, Washington, DC, Apr. 24, 1996.

46. Scott M. Harvey, William W. Hogan, and Susan L. Pope, *Transmission Capacity Reservations and Transmission Congestion Contracts*, Harvard University, 1997.

47. William Vickrey, *Responsive Pricing of Public Utilities*, 2 BELL J. ECON. & MGMT. SCI., Jan. 1971, at 1,337–46.

  PII S1040-6190(00)00146-9