

Integrating Financial and Physical Contracting in Electric Power Markets*

Chitru S. Fernando**
Paul R. Kleindorfer***

July, 1996

Abstract

Against the background of continuing restructuring of the U. S. electric power market, this paper considers the integration of financial and physical contracting under various models of market structure and transmission pricing. We begin with a delineation of the objectives which we believe, implicitly or explicitly, underlie the move towards restructuring. These include transparent and efficient markets for both long-term and short-term transactions, dynamic efficiency and innovation, customer-focused operations, and system integrity. We use these objectives to derive a number of important policy implications for the restructured power markets including clear ownership boundaries and regulatory incentives for market participants to operate in a commercial manner, and transparent rules and incentives for efficient contracting and pricing. We point out the implications for decreased competition and increased regulatory transactions cost from proposals which do not satisfy the stated requirements for commercial operations (e.g., recent proposals for nodal pricing of transmission service coupled with highly complex settlement and reconciliation procedures among participants). We then describe a general approach which does satisfy the *prima facie* requirements of market transparency and economic incentives. This approach is based on zonal and ex ante transmission pricing, regulated for-profit transmission service providers (TSPs), and permissive market intermediation. We indicate for this approach, under various models of the Independent System Operator (ISO), how financial and physical contracting could be integrated and how regulation of TSPs could be accomplished. The required contracting includes financial instruments (spots, forwards, futures, and performance contracts) encompassing long-term and short-term energy contracts, asset-use and resource supply contracts, ancillary service contracts, investments in generation and transmission assets, load-management and demand-side management contracts, and contracting for other market-mediated services required for the efficient configuration and operation of the power market. We conclude the paper with a discussion of some open research questions.

1. Introduction

This paper analyzes approaches to the integration of financial and physical contracting in electric power markets. This integration is essential to the efficient restructuring of the electricity supply industry to assure the benefits of competition. On the one hand, restructuring has begun to unbundle the prices and other service attributes associated

* The authors are grateful for constructive comments on an earlier draft by Shimon Awerbuch, Beverly Brereton, Mark Reeder, Sandy Robinson and Donjun Wu.

** A. B. Freeman School of Business, Tulane University, New Orleans, LA.

*** The Wharton School, University of Pennsylvania, Philadelphia, PA.

with the stages (from generation to distribution and billing) of electric power supply, making these more open and transparent to the buyer and seller of electricity and supporting services. On the other hand, these services must be coordinated, rebundled and financed to assure that the various stages of supply operate smoothly and efficiently. Since the earliest discussions of competition¹, the associated joint problems of unbundling and (re-)contracting have been recognized as the centerpiece of the debate on competition in electric power in the United States.² The reason is clear. Although unbundling is central to achieving the benefits of competition, inefficient unbundling (i.e., unbundling which leaves undue recontracting or regulatory problems in its wake) may impede and dissipate all the expected benefits of competition. Several problems are apparent in this regard. First is the issue of assuring system stability and integrity and the associated issue of reliability. Second is structuring an appropriate solution to the stranded cost recovery problem during the transition. Third, is determining the appropriate structure of ownership, control and regulatory governance of transmission services. The central focus of this paper is the last named issue on structuring transmission services to facilitate competition in generation.

Concerning the appropriate structure of transmission service in the unbundled market, the example of natural gas has underlined the clear benefits of open access and transparency in price and service offerings from a common carrier bulk transport/transmission provider (see. e.g., [Doane & Spulber, 1994]). Indeed, these perceived benefits were motivating forces in drafting the requirements of the Energy Power Act (EPAct) of 1992 and the FERC's subsequent actions to implement open access, comparable service and transparent pricing. However, recent proposals for achieving these requirements in transmission service have been highly complex and seem ill-suited for normal commercial activity, let alone as vehicles for promoting transparency and competition in generation and new services, the main sources of benefits from unbundled electric power. Thus, we argue for clear ownership boundaries for transmission service, with transparent and simple pricing structures, and with performance-based regulation on transmission service providers to assure that they face incentives to consider total system operations and efficiency in their long-range and short-range decisions. This leads us to discuss various organizational boundary issues for both transmission asset providers (TAPs) and the system operator(s) (the so-called Independent System Operator or ISO). We argue that regulated, profit-maximizing agents should be given both of these responsibilities, and we discuss various ways in which TAPs and ISOs might contract with one another to assure economic efficiency and breakeven operations.

The leit motif of this paper is that unbundling of the electric power value chain must be followed by contracting and rebundling along the value chain and that efficiency in rebundling will require transparent markets and commercially oriented market participants. As in any other active market, the market for electric services will consist of both the participants on the physical side of the business (providing generation and associated supply-side support, transmission services, and distribution/demand-side management), as well as the financial side (providing brokering and other intermediation such as financial risk management, and generally enhancing the liquidity and efficiency of the markets they support). The key issue we address is how to assure an efficient integration of these two complementary sides of the market.

In the next section, we set out some principles which we believe should guide the design of proposals for restructuring. In section 3, we describe the elements of the unbundled market place, and point to several key issues which we intend to explore. In section 4, we explore the first of these issues, the structure of the Independent System Operator (ISO) and its role in assuring open access, efficient transmission service and in facilitating the market. In section 5, we explore alternative organizational and ownership structures for the ISO and TAPs. In section 6, we consider access and pricing for transmission services. In section 7, we discuss the role of financial instruments and intermediation in the market. Section 8 recapitulates and points to some open research questions.

2. Principles of Restructuring

¹ For a discussion of (re-)contracting issues in this context, see [Joskow and Schmalensee, 1984].

² For a discussion of the underlying forces of change in the U.S., see [Fernando et al. ,1995].

When considering proposals for restructuring, most observers have in mind a set of assumptions (often implicit) on the principles and objectives of restructuring. These generally evoke a vision of an end-state which one might summarize as an efficient, dynamic and competitive market for power. While there is general agreement about this end-state, the factors and conditions which may influence achieving it are often either unstated or left as points of contest in the debate. It is useful for our following argument to summarize these underlying factors and conditions explicitly.³

Efficient Pricing: Pricing should be based on short-run marginal cost (SRMC), with breakeven prices derived through efficient demand charges or markup procedures based on SRMC.

Efficient Long-term Contracting: Both on the supply and demand side, long-term contracting for power and for support services should enable risk management and longer-term asset commitment.

Efficient Spot Market: To promote efficient matching of residual assets and demands for service, net of longer-term commitments, an efficient and transparent spot market should exist.

Incentives for Efficient Investment and Maintenance of Capital Stock: All service providers should have appropriate incentives to invest in capital and human assets, and to maintain them, in support of the market.

Incentives for Cost Minimization in Operations and System Configuration: In the short run, all market participants should face incentives to minimize total costs of system operations and to make available to the system assets which are needed for this purpose.

Customer-focused Design and Delivery of Services: Where additional value is attached to changes in services (e.g., in billing, in quality, in documentation, in service support of applications, etc.) by any buyer in the electric power supply chain, there should be incentives for sellers of such services to create these value-adding design changes.

Effective, Fair and Efficient Regulation: Where regulation is involved, it should satisfy the usual regulatory performance criteria, including an appropriate regard for minimizing regulatory transactions costs.

Clear and Transferable Property Rights: To assure discipline and information from the capital market and to provide operational meaning to the value of asset and franchise ownership, property rights (including the right to be an ISO or a TSP) should be clear and transferable.

Effective Competition: Whether in generation, between generation, transmission and load management or in service provision (leveraged by intermediation), competition is the main driver of change and benefits. Thus, proposals for individual pieces of unbundling policy (e.g., for transmission access) must be evaluated in terms of their impact on overall competition and not simply as stand-alone proposals.

System Integrity and Stability: In addition to the economic viability of the system, it must also satisfy a host of engineering requirements related to the special nature of electric power requiring instantaneous balancing of supply and demand across the network.

Let us first note some of the tradeoffs and implications implicit in the above principles. First, with regard to pricing, a tradeoff exists between short-run and long-run welfare. In the short-run, maximizing the traditional welfare measure of consumer and producer surplus (possibly subject to breakeven constraints if scale economies are present) gives rise to SRMC-based pricing. On the other hand, longer-term welfare or fairness considerations may require significant departures from SRMC-based pricing, either to recover stranded investments (e.g., via access charges) or to promote market transparency.

³ For a discussion and elaboration of efficiency criteria in the context of regulated industries, see [Crew and Kleindorfer, 1986].

Second, for incentives and for transferable property rights as well as for regulatory reasons, ownership boundaries must be clear. Absent such clarity, the ability to make decisions and to understand the motives of market participants will be impaired. As a case in point, loosely structured Regional Transmission Groups (RTGs) may have significant problems with decisions regarding maintenance and investment decisions in the transmission network unless the RTG itself is imbued with a decision and property rights structure that makes plain what benefits accrue to whom from such decisions. We return to this issue in section 5.

Third, both long- and short-run markets must have the requisite structure and institutional support to assure that they are liquid, transparent and not captured by anyone.⁴ Besides the transparency and ease of access implications of this, we also believe that this implies a relatively permissive approach to intermediaries to promote learning and experimentation and to exhaust gains to trade.

Fourth, the above principles should suggest much more to the reader than simply driving electric power supply toward more cost-reflective pricing. At least as important as this is the change in mind set which accompanies the move from monopoly to competitive markets, a change from internally-driven service provision to market-focused provision, from engineering-focused to customer-focused service delivery, and from homogeneous product offerings to segment-specific products and marketing. The key here is that aggregate welfare is driven both by consumers' willingness-to-pay (which can be expected to increase dramatically if service providers become market-focused) as well as by the total cost of providing a given set of products and services. Thus, there are two sets of conditions appropriate to benchmarking market efficiency:

- (i) The traditional price-cost benchmark that indicates that price should be set to SRMC (which incorporates implicitly the assumption of cost minimization) and capacity set to assure that SRMC and LRMC are equal;
- (ii) The requirement that new services be introduced when the benefits (measured by customer willingness-to-pay, WTP) exceed the cost of such services; with service quality and other service attributes determined, on a market-segment specific basis.

3. Structure of the Unbundled Electric Power Sector and Key Issues

The Nature of Unbundling

Unbundling occurring at two physical levels (see Figure 1): (1) between generation, transmission and distribution; and (2) within generation, between the provision of energy and various other ancillary services. In addition there is a separation of physical products and financial services as we discuss below. Unbundling aims to achieve clear pricing and service separability between the separate elements along the value chain. The benefits of unbundling are to clarify for competitive reasons the cost and value of each of these separate elements. The problem created by unbundling is that these separate elements must be rebundled, via contracting or spot markets, in an on-going fashion to (re-)create from these elements desired services and end outputs.

⁴ The early history of the British experience on this point is instructive; see [Newbery, 1995].

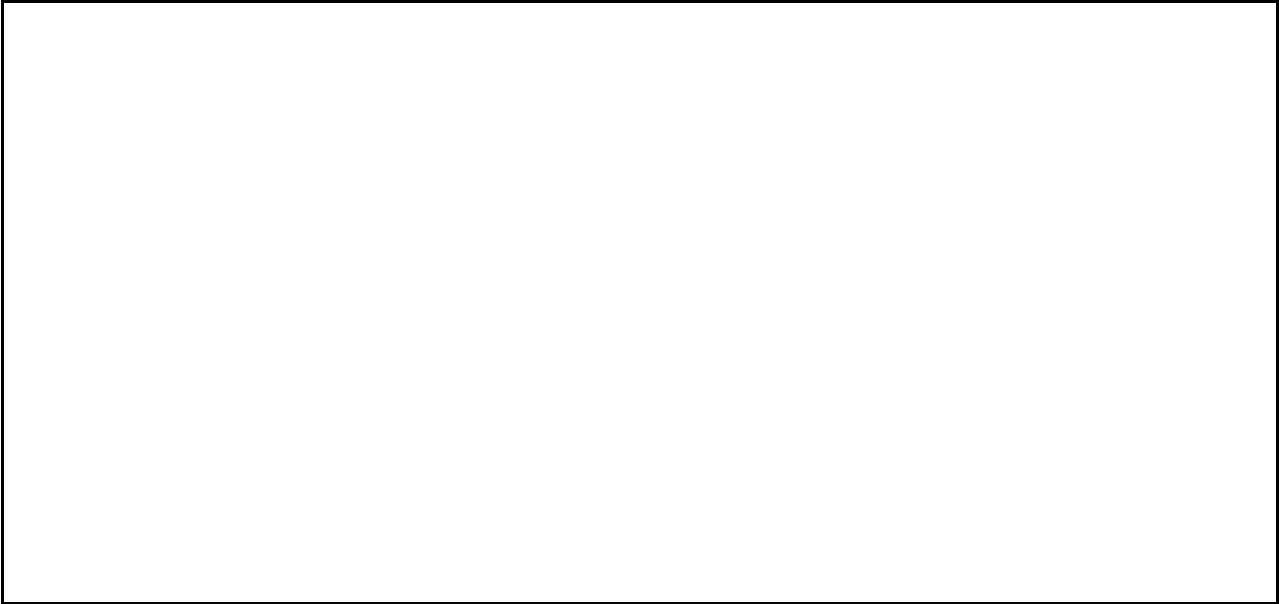


Figure1: Unbundled Electricity Value Chain

Figure 2 reflects the structure of an industry in which generation, transmission and distribution services are unbundled. Note that the break-up is facilitated actively by power market intermediaries who will also provide or arrange network coordination and other support services. From the standpoint of achieving the efficiency gains which are sought through unbundling, separation of generation, transmission and distribution is clearly the primary goal. Unbundling of generation services (spinning reserves) is also important in order to provide the same competitive and transparency benefits. The latter unbundling could take many forms, but will likely involve contracting for such services in spot and longer-term contract markets by the ISO.

Figure 2: The Unbundled System

The Role of Intermediaries in an Unbundled Industry

In the old vertically integrated structure of the electric utility industry, there was little scope for intermediation, since all transactions along the value chain were internalized within a single company. However, the trends toward emergence of full-fledged intermediation have been evident for some time, paralleling the trends toward greater competition. Power pooling and exchange arrangements across groups of vertically integrated utilities have been a first step in this direction. Furthermore, facilitated by these power pools and wholesale access, transactions across utility boundaries have expanded rapidly, accompanied by the emergence of NUGs as significant sources of generation. Some of these transactions have been intermediated by power marketers and brokers. In the new industry structure, the role of intermediation should expand rapidly. This is consistent with our view, discussed further in section 7, that intermediation is the "lubricant" of competitive markets.

The Workings of an Unbundled Electric Power Industry

Figure 3 provides a snapshot of the physical functions provided by the electric power system and the financial decisions and instruments which complement and parallel the physical. We structure the physical system *functions*

and the financial market *decisions/contracts* as they occur in 4 time frames, Long-Term, Medium-Term, Short-Term and Real-Time.

Long-term Functions and Decisions

Physical: Technology planning and acquisition, human resource planning and development, to build and operate assets to support generation, transmission and distribution (GTD).

Financial: Secure required capital, technology and human resources to accomplish the physical functions.

Medium-term Functions and Decisions

Physical: Schedule and implement system maintenance of GTD assets.

Financial: Forward contracts and bilateral agreements are negotiated for power delivery and contracts for load management, for transmission constraint payments, and for delivery of ancillary generations support are determined.

Short-term Functions and Decisions

Physical: Forecast and schedule near-term power demand. Unit commitment decisions and other set-up decisions to enable economic dispatch are made.

Financial: Execution of medium-term contracts (e.g., forwards); spot markets and economic dispatch provide clearing mechanisms for residual supply and demand.

Real-time Functions and Decisions

Physical: Network coordination occurs to assure system reliability, security and stability. This coordination and balancing occur through spinning reserves and load management, with Automatic Generation Control (AGC) and ancillary generation support providing frequency and voltage support.

Financial: Execution of medium- and short-term contracts for interruptible loads, VAR contracts and other support services.

Figure 3: Electric System Time Line: Market and Physical

In terms of organizational boundaries, the natural demarcation is as shown in Figure 3 between the organization(s) controlling long- and medium-term transactions, and those occurring in the short-run or in real-time. The latter transactions are the purview of system operations and organizationally will be the responsibility of the Independent System Operator (the ISO). The longer-term functions and decisions are the responsibility of Generation, Distribution and Transmission Asset Providers (we refer the last-named as TAPs). Concerning transmission service and network coordination, the key is the organization and ownership boundaries of the ISO and the TAPs. We discuss this in the next Section in more detail, but it should be clear right away that two general possibilities exist: either the ISO and the TAPs are brought under the control of one (presumably for-profit, regulated) company, or the ISO and the TAPs remain under separate ownership and control. The former instance is seen in the structure of the UK and New Zealand power markets,⁵ in which a single entity owns and controls both transmission assets as well as the ISO. The latter is the model which is being pursued in several ISO proposals under the Regional Transmission Group concept in the US.⁶

In the transition to the unbundled electric power industry, the short-term functions and decisions (those that occur in the time frame of a day down to an hour or possibly to 15 minutes) require the greatest evolution from today's utility operations. Development of a spot market for electricity is the major change in the short term domain. How much change does this actually entail? Looking to other commodities (e.g., natural gas), spot markets develop both

⁵ For a description of the U.K. transmission infrastructure, see [Newbery, 1995]. For a description of the New Zealand infrastructure, see [Ring and Read, 1996].

⁶ For example, *Electricity Daily*, February 13, 1996, describes the launch of a "Super-ISO" in which six midwestern investor-owned utilities have announced their agreement to lease their assets to an ISO organization which would then control all transmission assets as system operator, and would pay for the use of these assets under long-term contracts with the respective asset owners.

rapidly and efficiently. The functions of the pool operator (i.e., the Independent System Operator or ISO) will include responsibility for least cost dispatch for the voluntary pool, together with the real-time functions of reliability, system security and stability for all transactions. The nomination and/or posting of transactions will occur *ex ante* such that the physical and financial transactions can be verified *ex post* and any over or under delivery/receipt can be identified and dealt with in the balancing costs.

The pool functions and associated financial instruments are well understood by now and include market clearing dispatch and settlement procedures.⁷ The physical functions of balancing and coordination can be maintained through a new/modified set of market instruments that can be exercised by the ISO. As an independent, performance-based, regulated entity, the ISO purchases contracts for reserves -- call contracts -- with specific performance characteristics based on expected needs that provide for MW and MWh. These contracts will include megawatt as well as megawatt resources. Contracts would be called to cover unplanned outages and increased demands. The cost of operation of this aspect of system operations would be covered through *ex ante* contracts with *ex post* verification -- plus a management fee -- to the responsible participants. Within a prespecified range, hourly costs could be traded off between participants before the actual transactions came due. The result of these *ex post* trades is the creation of a secondary market in capacity and/or energy directly analogous to the market that has emerged in natural gas with a longer clearing time.

The second function to be fulfilled by the ISO, even closer to real-time, is that of maintaining system frequency. In today's system, frequency is maintained by Automatic Generation Control devices that are installed and operating on most generating units. While the physics are more complex, these devices are best thought of as monitors that automatically sense deviations away from nominal 60 Hz frequency. When frequency is low, additional primary energy (steam) is introduced into the unit thus providing more rotating energy in the system. When frequency is high the reverse is true. Both the AGC device and its operation have a cost to the generating unit owner/operator. This function is readily provided through long-term contracting between the unit owner and the ISO, which would provide contractual incentives for system efficiency.⁸

Two final functions must be fulfilled by the ISO for stability and security to be maintained. The first is the requirement for VAR support (seconds to minutes) and the second the need to respond to rapid changes in system configuration that will induce transience -- i.e. manage transience such that the system automatically returns to acceptable operating conditions rather than becoming unstable (in a time frame of cycles to seconds). VAR support today is provided by generators capable of "lagging or leading" in phase angle of generation, through capacitor banks or through static VAR compensation or so called "FACTS" devices, (Flexible AC Transmission Systems). This capability provides the trade-off between real and reactive (VAR) generation at any unit. As with AGC, this function has a capacity cost -- the capability -- and an operating cost that needs to be contracted for, usually as a call contract.⁹ VAR contracts will be long term with performance based on monitored unit output -- parallel to that employed today.

⁷ For the UK, see [Newbery, 1995]. See also [Einhorn and Siddiqi, 1996] for a description of pool operations and settlement procedures in other countries.

⁸ Starting in October of 1994 the National Grid Company of the UK advertised in the *London Financial Times* (October 13, 1994) for "Frequency Control Services" and "Reserve and Constraint Services" in advertisements headlined "Have you got the power to make money?" Their bid is to purchase on either the supply or the demand side services that will respond rapidly to frequency change or services that can respond to needs for system reserves or constraints. Both services were called to bid by December 2, 1994.

⁹ It should be noted that contracting for VARs was one of the earliest modifications introduced into the UK Pooling system.

The final issue in the real-time domain is how the services of system operation and network coordination provided by the ISO would be paid for. As discussed under reserves above, some services required by the system are directly attributable to individual participants in the system. This is specifically true of both shortfalls in supply or excesses in demand relative to contracted levels. ISO can attribute and bill for these services given known contracted capabilities. The balancing and bookkeeping can occur *ex post* as part of an established accounting routine as occurs with the “uplift” function of the UK Poolings and Settlements¹⁰. The other functions to be fulfilled in real-time by the ISO are systems based and not attributable. These functions need be paid for, in essence, by a performance-based contract between the users of the system (end consumers) and the ISO. As a regulated entity, the ISO will perform as close to a competitive entity as possible if its earnings are a function of the difference between a price cap and its costs. This drives its costs of operating the system to a minimum for provision of a predefined and regulated level of service.

The above sketch of how the unbundled electric power industry might function suggest three critical issues which will need to be resolved. In some sense, these all revolve around the area of the Transmission Service Provider (TSP) and the Independent System Operator (ISO). More specifically, the issues we explore below in detail are these:

1. What are the possible structures and roles of the of the ISO? Will the ISO simply be a market facilitator which controls the Network (the real-time functions above) and the voluntary pool(s), while contracting for assets and support services with other market participants? Or will the ISO be a commercial entity with assets (e.g., wires or generation plant) of its own?
2. How will transmission access, pricing, investment, contracting for services and regulation be accomplished for each of the feasible alternatives identified in (1)?
3. What should be the role of financial instruments and intermediation in the market?

4. Role and Structure of the ISO

This section briefly discusses several approaches to organizing and regulating the ISO and Transmission Service Providers (TSPs) and their relationship to facilitating long-term markets (between Gencos and Discos) and short-term markets between all participants via the Pool. This is one of the central questions which will drive the efficiency and operation of the reorganized industry. A variety of ISO models are technically possible, differentiated in broad terms by the following (inter-linked) factors:

1. involvement of the ISO in the energy market;
2. the scope of business activities undertaken by the ISO, including the extent of support functions bundled within the ISO;
3. ownership and/or control of assets by the ISO.

We will first describe three benchmark ISO models -- CoorCo, GridCo and PoolCo -- which are broadly differentiated along these lines, and discuss their potential for meeting the criteria set out above. Thereafter, we consider hybrid versions which combine features of the above models.

CoorCo -- Coordination Service

¹⁰ Note that the *ex post* balancing function in the UK does not differentiate between sources of problems and spreads these costs evenly to all consumers.

The CoorCo model of an ISO has some similarities to the various coordination operations that have existed in the U.S. to interconnect transmission lines and pool generation belonging to vertically integrated electric utilities in various parts of the country. Historically, these arrangements were driven mainly by system security considerations, with very little commercial activity taking place across utilities relative to transactions internal to their vertically integrated structures.

An example of a CoorCo-type scheme is the Regional Transmission Group (RTG). RTGs have been proposed to coordinate the transmission resources of different utilities, taking on the coordination of operations, planning and investments, and dispatch and settlement duties on behalf of its members. A prominent example is the Western Regional Transmission Association (WRTA) in the United States which encompasses a large segment of the Western U.S.

Figure 4 illustrates the concept of the CoorCo model. In a strict CoorCo-type ISO model, energy transactions occur only through bilateral transactions between generators and consumers. The ISO is informed of the power flows that would be associated with these bilateral contracts, so that it could make necessary arrangements to accommodate these power flows while maintaining system reliability. There is no "financial" pooling in the energy market -- all settlements are undertaken bilaterally between buyers and sellers -- and physical flows relating to particular transactions are assumed to follow "contract paths".¹¹ Buyers and sellers would arrange with the ISO for meeting the cost of losses and system support services that are associated with individual bilateral transactions. Emergency conditions will, on occasion, occur through severe weather or other extremes. Under these conditions, the CoorCo could call on individual generators to supply emergency power or energy.

In this bilateral model of energy transactions, individual generators and consumers will receive and pay different prices. Competition will occur through buyers seeking out least-cost sellers leveraging upon the transmission network. The ISO plays no market-making role nor implements economic dispatch to facilitate competition and least cost generation usage.

A CoorCo-ISO is not required to own any generation or transmission assets, only to have control over the operation of sufficient assets to carry out its coordination function properly. The extent of control required by the CoorCo would clearly be system-specific, but is likely to include a significant portion of the transmission network together with generation plant that are necessary for back-up reserves, frequency, VAR support, etc. Owners of transmission assets would sell the right to the use of the capacity of these assets to power marketers and principals striking power contracts. They could also sell transmission capacity to the CoorCo who would acquire this capacity for the purpose of fulfilling its responsibilities. Buyers and sellers would be free to transact in transmission capacity. These transactions could occur on both a firm or non-firm basis, in both the primary market for transmission capacity sales and in secondary markets. "Firm" in this context implies that owners of firm capacity would have first priority in its use.

Transmission investment is undertaken by third-party transmission providers or by the CoorCo itself. The CoorCo can assist in system planning and by being a central information source for potential investors, thereby ensuring that the most profitable (and hence value-creating) investment opportunities are identified, while simultaneously avoiding excess capacity build-up. For such a decentralized scheme of investment to be effective, investors in transmission would need to receive the full value provided by their investment through an appropriate scheme of pricing (see below).

¹¹ Even though actual flows will follow the laws of physics and may be quite different.

A key question underlying the CoorCo model is how its operations will be financed and regulated, and what impact this would have on CoorCo's incentives to fulfill the desired objectives. It has been suggested by some that the CoorCo fits the mold of a non-profit or even public enterprise since (a) the value created by CoorCo is largely reflected in quality and reliability terms, both of which are already very high especially in industrialized countries; (b) due to the significant externalities associated with electricity service, it would be difficult to price-differentiate CoorCo's services based on value; and (c) CoorCo may be able to perform its service without significant asset ownership. In the RTG model of the U.S., the ISO is owned by member utilities but operates independently on a non-profit basis. One potential approach is to hold CoorCo to a performance standard and set its revenue based on a "cost-plus" approach. From an efficiency standpoint, this will result in excess conservatism and lack of attention to costs. In particular, CoorCo would not have any incentive to apply pressure on third-party suppliers of transmission services to provide their services at least-cost.

The multiplier effects of a CoorCo which has no strong economic incentive to hold down its own costs could increase costs quite substantially for other industry participants. This could occur through increased costs of losses and system support borne by downstream users, due to "risk averse" system operation. Resolving this problem using a price cap or other scheme of incentive regulation could create new problems, especially if the ISO is a non-profit organization. The basis selected for setting the price cap (such as Rate of Return) could potentially cause new incentive problems, including conflicts of interest between CoorCo and third-party transmission suppliers, generators and customers, thereby jeopardizing its independence and objectivity. On balance, there appear to be very significant incentive and control problems with the CoorCo model of the ISO.

GridCo -- Integrated Transmission Grid

The GridCo model expands the role of the ISO by enlarging the scope of its shorter term activities relative to CoorCo and also by taking on the longer term functions associated with the efficient operation, planning and investment of the transmission grid. The purpose of integrating these functions within the ISO is to achieve the economic efficiencies associated with the operation, planning and investment of the transmission grid that are less explicit in the CoorCo model.

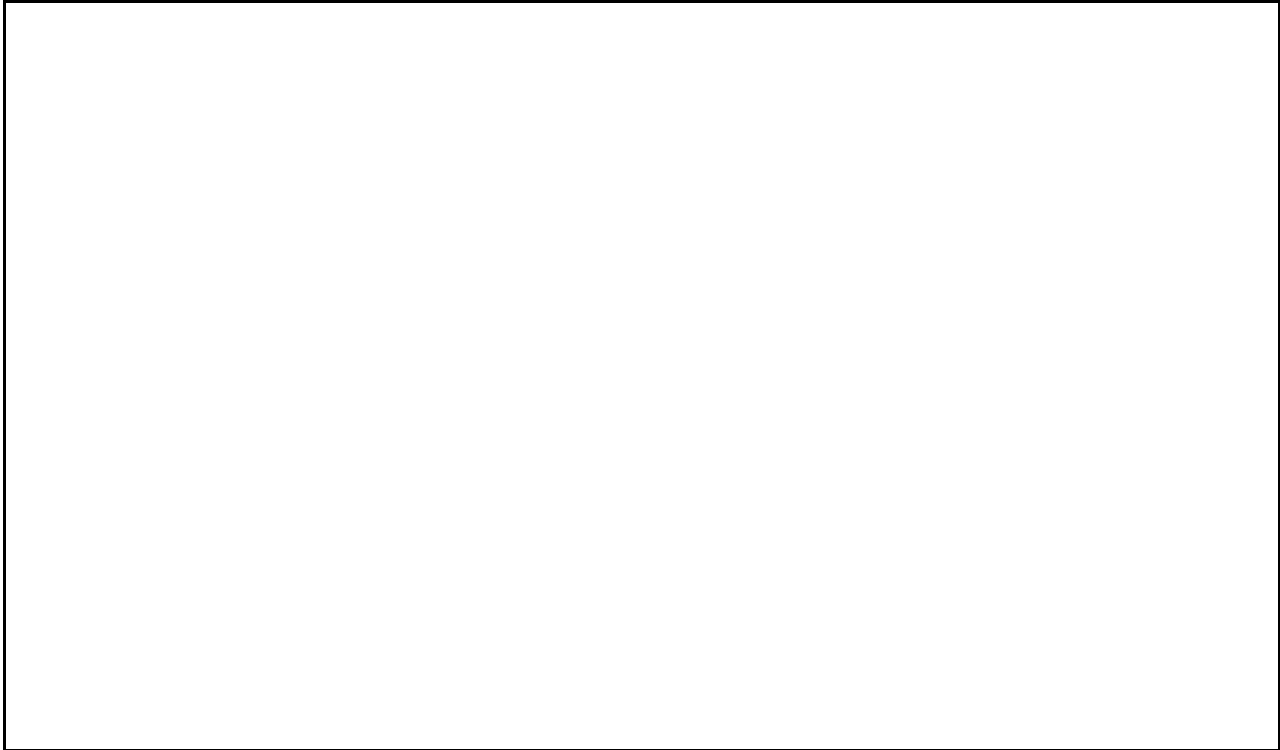
Whereas the CoorCo model was based on the concept of competing providers of transmission services, in the GridCo model all transmission services are brought under the umbrella of a single transmission provider -- GridCo. This explicitly reduces the reliance on competition to provide efficient transmission operation, although the GridCo should seek to outsource as many of its service requirements as possible. Unlike the "multi-provider" CoorCo model, all power flows associated with bilateral power contracts would use GridCo's lines and transmission services.

Figure 5 illustrates the concept of the GridCo model. As in the CoorCo model, energy transactions will occur through bilateral transactions between generators and consumers. The transmission services associated with these transactions would need to be arranged through the GridCo, including access to the grid and its use. The GridCo would be responsible for operating the system at specified levels of reliability in least cost fashion.

As in the previous CoorCo model, the GridCo ISO plays no direct market-making role nor does it implement economic dispatch. However, by providing a reliable and efficient grid system, it would facilitate competition in the generation market.

A GridCo will own or contract for the use of (e.g. through leasing or long-term usage contracts) the transmission assets in the network. This would include both wires and associated infrastructure as well as system support services. The system would be "single provider" in the sense that all transmission services would be provided by the ISO. Unlike in the CoorCo model, there would be no market in transmission capacity, primary or secondary. The GridCo would be the sole (regulated monopoly) seller of transmission capacity. This would not preclude the differentiation of transmission service on the basis of firm and non-firm, nor differentiating the pricing of transmission based on space and time (see our discussion below on transmission pricing).

As in CoorCo, transmission investment can be undertaken by third-party transmission investors or by the GridCo itself. In either case, GridCo will act as the clearinghouse for new transmission projects. With an appropriately designed incentive scheme (see below), it would be in GridCo's interest to seek out the cheapest possible solutions to the system's transmission needs.



In contrast with the CoorCo model, GridCo can be set the clear economic objective of minimizing total short run and long run costs of transmission, since all these costs (including system losses) are internalized within GridCo. GridCo would also be held to a quality standard. Revenues to the GridCo would accrue from transmission charges levied on system users. These transmission charges would be designed to recover the capital and operating costs of the system together with an appropriate profit scheme which is designed to sustain incentives for continued cost minimization. A well-designed price cap scheme would provide such incentives.

A significant advantage of GridCo versus the CoorCo model is its reduced complexity. Although a monopoly provider, GridCo will depend on outsourcing for as many services as possible so that the benefits of competition will still accrue to system users without the costs of regulating diverse transmission providers. GridCo would find it in its interest to create competition in the provision of various transmission services, including constrained generation, interruptible load, line maintenance, voltage support, etc. While hold-up problems may be difficult to avoid in the short run (such as plants charging exorbitant amounts for constrained running, or maintenance contractors marking up their prices), GridCo will move actively to eliminate such situations.

The GridCo concept lends itself naturally to systems where transmission assets have been previously owned and operated by a single entity, as in England and Wales. In other situations, as in the U.S. where transmission assets in regional pools have multiple owner/operators, the transition to a single operator has proved to be more difficult, because of the complexities of valuing assets, pricing transmission services to the previous owners and revenue sharing. However, the GridCo concept holds better promise than CoorCo for meeting the objectives of economic efficiency coupled with reliable service that are being sought through the ISO.

PoolCo -- Pooling of Energy and Transmission

While there are several variations of the PoolCo concept¹², the core idea of a PoolCo is that of a service which would buy power at generator nodes and sell it at consumer nodes. The PoolCo would be an independent entity which would control the operation of the transmission network within its region and dispatch all generation for energy or system support. Generators would sell power into the pool and consumers would purchase power from the pool at prices that periodically (e.g. each half hour) "clear the market". Market participants would also bear the cost of transmission, which would cause price differentiation by location. The PoolCo would schedule and dispatch generation according to the merit-order (where possible) established by the contract prices (bids, costs, or previously agreed on some other basis) of the generators. It would also own or contract for system support capacity necessary for cost minimization and preserving system reliability.

In this ISO model (see Figure 6), the PoolCo would have two basic responsibilities:

- a) act as a market clearinghouse, using the transmission facilities at its disposal to link generation and load, thereby "making" the market and preserving its integrity.
- b) preserve reliability of service to market participants.

Applied strictly, the PoolCo concept requires mandatory pooling of generation resources, with voluntary (or residual) pools being a variant thereof in which GenCos announce (e.g., on a day ahead basis) which units are to be pool dispatchable. Mandatory pooling is similar to the way the British power pool currently operates and voluntary pooling is similar to the way the Norwegian power pool currently operates.¹³ Pooling of both energy and transmission resources in this way is intended to assure equal treatment for all spot market participants in system access, pricing, and revenue allocation. All physical flows will enter the spot market, although users may enter into financial contracts ("swap contracts" or "contracts for differences") bilaterally to fix their payments and receipts associated with specific transactions. For the PoolCo to carry out its responsibilities, the only requirement is that it has control over sufficient generation and transmission assets to preserve a competitive market and system reliability. This control can be obtained through ownership of the assets by PoolCo or by contracting for their use.

¹² See, e.g., the early work of [Hogan, 1992] and the recent special issue on transmission pricing of the *Journal of Regulatory Economics*, July, 1996.

¹³ The British system is described in [Newbery, 1995]. The Norwegian system is described in [Westre, 1996].

As in GridCo, the PoolCo would be the sole (regulated monopoly) seller of transmission capacity. This would not preclude the differentiation of transmission service on the basis of firm and non-firm (as in interruptible service), nor differentiating the pricing of transmission based on space and time. As in the previous models, transmission investment can be undertaken by third-party transmission investors or by PoolCo. As before in the case of GridCo, PoolCo will have an incentive to assure transmission investments at least cost. It will facilitate this by long-term planning and publicizing investment needs in the network.

PoolCo goes a step further beyond GridCo by effectively mandating economic dispatch to be carried out by the ISO. This reduces the burden on system users to competitively seek out opportunities for cost reduction. Apart from this, PoolCo's cost and revenue structure, and potential regulatory options, would be very similar to GridCo; since the merit order is quite transparent, PoolCo's economic objective boils down to minimizing the costs of transmission.

As in the case of GridCo, the PoolCo concept has met some opposition in systems which were previously multi-owned and operated, and from those who believe that creation of new monopoly structures is antithetical to the current unbundling initiatives which are aimed at increasing competition. In particular, PoolCo cuts out several intermediation functions which are vital for promoting competition in an unbundled industry. We discuss below some hybrid proposals which have been put forward to overcome such objections.

Hybrid Models: Voluntary or Flexible Pools

An idea which has emerged from experience in several countries (especially Norway and Argentina) and actively supported by several utilities in the California restructuring debate in the U.S. is the concept of a voluntary or *flexible* pool. Under this arrangement, system users (both buyers and sellers) have the choice of either accessing the spot market (created by pooling a

segment of generation and load in the system) or transacting bilaterally bypassing the pool altogether. This arrangement is attractive relative to mandated pooling since:

- a. it does not preclude the free choice of market participants;
- b. it does not inhibit the development of value-creating business opportunities; and
- c. it minimizes potential inefficiencies associated with pool rules, letting these evolve over time through experience.

Arrangements similar to flexible pooling exist in markets for all commodities, since these markets consist of both spot and forward contracting arrangements. Some have argued that a flexible pooling scheme is identical to mandated pooling where in the latter case market participants can enter into bilateral side contracts priced off the spot market to fix long-term prices. This remains an open issue but some differences are very clear. In particular, the characteristics of the spot market itself (liquidity, price volatility, etc.) are likely to be quite different in the two cases, since in the latter case spot market participation occurs only by self-selection.

Flexible pooling can co-exist in principle with any of the above models of the ISO. The ISO would provide economic dispatch services to those market participants who opt for it. Given the recontracting and incentive problems with CoorCo noted above, however, we will only consider hybrids of the PoolCo and GridCo in what follows. In the Flexible PoolCo, which is similar to the evolving UK system, most energy is traded through the Pool, with some self-dispatch and intra-zonal bilateral contracting allowed. In the Flexible GridCo, which is similar to the evolving Norwegian system, most energy is traded through bilaterals with residual trades being accomplished through a voluntary pool. In either of these cases, we assume that the ISO is set up as the System Operator responsible for real-time system operations and for short-run operations required to assure timely information on the nature of bilateral transactions is available to assure efficient scheduling and dispatch. We now consider the organization and regulation of this form of ISO and its relationship to Transmission Asset Providers (TAPs).

5. Efficient Organization and Regulation of Transmission

Scope and Organization of Transmission Service

The discussion here is concerned with defining the appropriate scope of transmission service and with the principles underlying the recovery of revenue requirements for transmission assets. Our discussion applies to both single owner (e.g. TransCo) and multi-owner (e.g.) RTG arrangements.

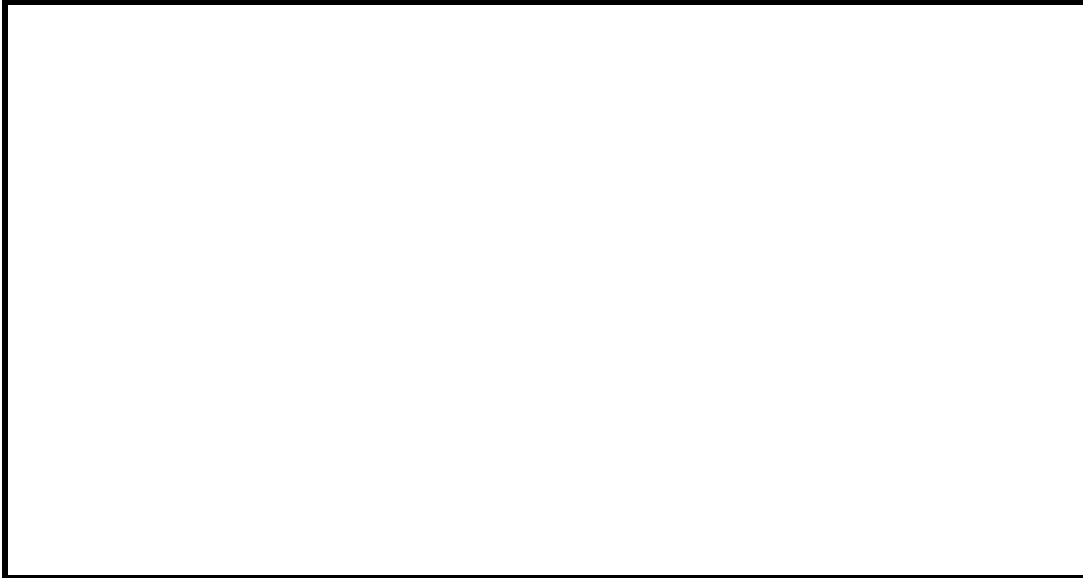
Figure 7 illustrates the components of the transmission service. At a primary level, generators and loads will gain access to the market through a connection to the transmission grid, and their supply and demand gives rise to the electricity marketplace. Bringing generators and wholesale customers could be characterized as the QUANTITY or ENERGY side of the transmission service. The other side of the transmission service is the QUALITY or SYSTEM SUPPORT side, which is concerned with ensuring security of supply, and voltage and frequency standards.

As set out in the framework shown in Figure 7, the quality side of the transmission service would include the procurement of Out-of-Merit (OOM) generation services for constraint control and ancillary services from generators (and other suppliers of these services). The provider of transmission service may also acquire the right to interrupt loads or In-Merit (IM) generation through interruptible service contracts.

The other key aspect of the quality side of the transmission service is the security or insurance value of the network. The point here is that all participants in the energy market acquire through their transmission grid connection a valuable option to generate or consume electricity. This option is made valuable by the additional investments (e.g.

reserve lines) and operational decisions (e.g. scheduling generation reserve) undertaken by the transmission provider. Hence, the transmission grid is both a medium for transportation/trading, as well as a security network.

It is essential that all these elements on the quality side of the transmission service be internalized within the transmission provider in order for this service to be planned and operated efficiently



As noted in section 2, the key to successful unbundling is the ability to rebundle without undue transactions costs. In particular, given the importance of centralized operations in accomplishing its real-time functions, it is clear that the ISO must be located within the organizational boundaries of a single economic entity. This leads to one classification of possible ownership structures for transmission: (a) either the same entity which houses the ISO owns and operates other transmission assets, or (b) this entity consists only of the ISO and does not own these assets but leases/contracts for these from other transmission asset owners. Using comparative institutional economics¹⁴, it is not possible to rule either of these approaches out as *prima facie* inefficient. Approach (a), which sets up a single company, the TransCo, would give rise to the usual problems of providing regulatory incentives through performance-based regulation to assure that the TransCo, a regulated monopolist, undertook its responsibilities in a manner which promoted system-wide efficiency. Approach (b), the ISO+TAPs, would yield clearer information on the value of transmission assets and services (the former provided by TAPs and the latter by the ISO), but would lead to transactions costs between the ISO and the TAPs in contracting for and maintaining the transmission assets. A hybrid approach might create a single organizational entity, the TransCo, but require it to have two separate divisions, TransCo-Wires and TransCo-ISO, to create transparency in cashflows and value-added resulting from the asset management and system operation functions of the TransCo. Let us consider these options in more detail.

In the single, unified TransCo option, a regulated monopolist would be given responsibility for universal transmission service. To assure clarity in its motives and some incentives for X-efficiency, this TransCo would have to be a for-profit, regulated monopoly. As noted above, it could be required to keep separate books on its ISO and its TSP operations. The TransCo would then face various forms of profit, price and quality regulation, as discussed below. Revenues for the TransCo would come from:

- a. Monopoly or reserved services, such as those associated with running the Pool and system operations.

¹⁴ For an introduction to institutional assessment procedures, see [Crew and Kleindorfer, 1986, Chapter 7].

- b. Contestable services, such as connecting new loads or generators to the system, which could be provided by a number of third parties.

Ideally the price and/or revenue for contestable services would not be regulated, but would be determined by an open market in these services. For services of type (a), prices and revenues would be derived from two traditional elements of transmission pricing (see section 6 below for more detail): capacity charges which would depend on the total capacity of generators connected to the grid, and energy charges which would depend on the energy carried by the transmission system. The total of these two charges would cover (for reserved services) asset costs, system operation costs, congestion costs and losses.

Under the ISO+TAPs option, asset providers and transmission service providers would be separated. Here the ISO must deal with the added complication of negotiating with independent asset owners (the TAPs) for continuing use, enhancement and maintenance of their assets. If, as envisioned in several recent RTG proposals, the ISO itself were set up by these TAPs, then additional problems of assuring uniform and fair treatment for all comers (including the TAPs) through a committee decision-making process involving all the TAPs presents additional opportunities for transactions costs and organizational inertia. Presumably, the same guidelines on reserved and contestable services would hold for the ISO+TAPs approach as for the TransCo approach. However, if the ISO is owned by the TAPs, additional monitoring and oversight will no doubt be called for to assure that the ISO fulfills its market facilitation role in an objective fashion.

Regulation

Appropriate regulatory scenarios will depend on which of the organizational alternatives sketched earlier is chosen. In the event that an asset-thin ISO is set up with no “wires” ownership, the key problem will be to provide incentives to the resulting ISO to properly contract for use of assets, since the cost of such use would be largely outside of the ISO’s control. In the event of a TransCo (with, say, an asset-holding division TransCo-Wires and a transmission service division TransCo-ISO, a key regulatory issue will be to assure that the TransCo faces the proper incentives to avoid inefficient strategies such as asset-padding.

Figure 8 captures the revenue and cost flows associated with transmission service. The transmission charge will be levied on loads (directly in the case of network service, indirectly in the case of point-to-point bilateral contracts), and will cover the cost of both the quantity and quality sides of the transmission service. The transmission provider may also charge both generators and loads for connection to the system (which would reimburse the provider the cost of the connection).

The transmission provider will provide service by building adequate capacity (through **investment**) and by operating the system reliably and efficiently. In some cases, it may be more efficient for the transmission provider to meet capacity needs by paying generators to operate **Out-of-Merit** or by paying loads for **interruption**. In addition, the provider would also be required to meet the cost of **system losses** and to pay generators for ancillary services such as reserves, frequency control, etc.

The key to effective regulation of transmission is to internalize all the costs that are associated with transmission service within the transmission provider. This will create the correct incentives for optimal investment and operation in the transmission grid. In the longer term planning horizon, the provider will optimally trade off investment decisions against the various operational options (such as Out-of-Merit-Dispatch (OOM) payments or interruptible contracts). In the shorter term operational horizon, the transmission provider will pick among the various short term options which are available to achieve least-cost system operation.

Regulating a TransCo's Revenues -- A TransCo's revenue stream could be regulated through cost of service, price caps or various other incentive regulation schemes (see Crew and Kleindorfer [1996]). A pure cost of service scheme is probably not appropriate in a setting where TransCo's cost side is subject to significant uncertainty, especially in the case of constraint control costs. The main point here is that the regulatory scheme should meet two criteria from TransCo's standpoint:

- a) Provide TransCo the correct incentives to invest and operate the transmission system. Thus, a price cap applied on a kWh basis for the energy components of the TransCo's services would be appropriate, and would cause the TransCo to confront the correct incentives for investment and contracting if Transco is regulated to cover all energy costs (losses and congestion costs) of transmission.



- b) Provide TransCo a means of passing through risks that it is not equipped to manage (for example, a substantial change in constraint costs as a result of a change in the relative coal/gas price).

Regulating an ISO+TAPs' Revenues -- The same principles as above apply to the determination of the aggregate revenue requirement. In the case of an unbundled TSP with multiple TAPs, revenue allocation mechanisms should provide asset owners proper signals of the value of their existing assets and the incremental value of various options for expanding transmission capacity. This is not as simple as it might seem, since an allocation mechanism based simply on, say, MWh-miles would miss the insurance or quality value of some assets. Thus, a combination of a fixed capacity rental charge per MW-mile per year (set to cover maintenance expenses and depreciation plus a reasonable return on the asset) with a usage-sensitive energy fee would be required.

The issue of multiple TAPs and a correct valuing of their assets for quantity and quality of service remains an open issue. It points to the key difficulty with the ISO+TAPs model, the level of contractual transactions costs with TAPs and the related issue of control of asset quality by ISO. From the TAPs point of view, there are problems of assuring that their assets are valued correctly in contracts with the ISO and that the assets are properly maintained. To the extent that the TAPs jointly own the ISO, there would also be problems of assuring even-handedness in the provision of transmission service to non-TAP users.

6. Transmission Access, Pricing and Investment

Transmission Access

Open access to the electricity transmission networks that criss-cross the country is an essential prerequisite to the operation of a competitive unbundled market in electric power. While the Energy Policy Act of 1992 initiated the opening of access to transmission through wholesale wheeling, several issues surrounding the price regulation of transmission services remain. These need to be resolved before transmission can become fully established as the cornerstone of a competitive electric power market in the US. As we see it, the key issues in transmission pricing to enable effective competition are the following:

- Transparency of prices (unbundled transmission service)
- Non-discriminatory (between native load and third parties)
- Efficiency - cost reflective

The right of access to a utility's transmission network by a third-party generator or distributor provides value in allowing such entities to sell at a higher price or purchase at a lower price than would otherwise be possible. In an unrestricted marketplace, this value would provide the basis for pricing the service. If the differential between the resulting price and the corresponding cost was excessive, this would normally be eliminated through competition or regulation. Under competition, and assuming that economies of scale are exhausted, prices would be driven down to marginal cost levels. This is a state which regulation would attempt (imperfectly) to emulate. In particular, transmission constraints at certain points in the system will be reflected by higher marginal costs of serving those points. So too will time-of-day differentiation of transmission prices reflect the differing marginal costs of serving particular demand points with transmission services as a function of the pattern of supplies and demands on the system at various points of time. Masking these marginal cost differences by uniform (postage stamp) rates, even if differentiated between firm and non-firm service, will deprive customers of valuable information of the costs they impose on the system with their loads.

In an unbundled industry, we envisage transmission services operating in an increasingly competitive environment. It should first be noted that generation and transmission are themselves substitutes in the sense that the bundled product of non-local generation and transmission can compete with local generation. Thus, with a competitive generation market, the market for transmission services will, if allowed by unbundled pricing, become more competitive over time. In particular, as the large energy price differentials even out throughout the country through competition, the opportunity for transmission service providers to extract monopoly rents will be greatly diminished. For the foreseeable future, however, there will still be a need for regulatory oversight of pricing and access rules for transmission services providers, but competition with generation nonetheless is an important efficiency driver for transmission.

One scenario for emerging competition in transmission services is the following. Transmission companies will sell firm capacity rights at regulated prices, which attempt to mirror location and time-dependent costs. Energy brokers, including generators and transmission companies, would bundle together generation and transmission services on a bilateral basis for wholesale customers.

Such bundled services would provide for pricing and billing arrangements, alternative contract lengths and other features which wholesale customers may find useful.

Following our scenario further, interruptible or non-firm service offerings will also be offered competitively, and this from two sources: first through longer-term contractual agreements by companies which have firm transmission and/or generation capacity which they wish to offer on non-firm terms (e.g., by pooling non-coincident demands in an efficient manner); second through medium-term and short-term spot markets which will act further to price the value of interruptible capacity at various points and various times along the transmission grid. In the resulting competitive market among energy brokers, generators and transmission companies, the combination of long-term bilateral contracting markets and shorter term contracting and spot markets will act interdependently to provide appropriate price-cost-value links between suppliers and customers.

The above scenario requires a cooperative organizational compact or regulatory structure (e.g., a Regional Transmission Group) to determine rated system paths and to act as an information or market coordination point for property rights for these paths. For these reason and to assure continuing stable evolution toward a fully competitive market in transmission services, some form of regulation will be required to provide transmission companies with the incentives for efficient operation and investment. In this regard, dictates of low regulatory transactions costs, high transparency of the pricing structures, and flexibility to compete all argue for a regulatory structure which is performance or price-cap based rather than rate-of-return based. Price-cap regulation provides incentives both for operating cost minimization as well as for growing revenues through development of customer-responsive services.

Pricing of Transmission

Unlike generation, transmission and distribution are still perceived as monopolistic, and subject to regulated schemes of pricing, following traditional models. However, even in these areas, innovative approaches are being sought to maximize efficiency. Transmission pricing and regulation must assure that there are proper incentives for investing in the transmission system and for efficiently utilizing existing transmission assets. Any feasible approach must also be transparent and compatible with an unbundled, competitive market for power.

A basic transmission pricing structure will have a combination of three components:

- | | |
|---------------------------|--|
| Access Charges: | Customer-specific costs of connecting a generator or load to the existing transmission network |
| Demand-based Use Charges: | Paid on a per kW basis per annum |
| Energy-based Use Charges: | Paid on a per kWh basis |

Both the demand-based and energy-based charges may (and should) vary by season, by time-of-day and by location. Firm transmission pricing should be structured to assure short-term efficiency and long-term viability/incentives for investment. On both efficiency and viability dimensions, cost-based pricing provides valuable indicators of alignment. The viability implications are clear -- failure to recover costs is not sustainable. On efficiency, while cost is not the only determinant, transmission pricing to recover the long-run incremental cost of prudent investments provides important signals to the market on efficient entry. Providing the correct economic signals to

consumers and to generators about the short run operating conditions of the grid and providing the owners/operators of the grid with the correct long run economic signals for investment in new capital stock are critical elements in both the operation and the future development of the electricity supply system.

The marginal cost pricing principles that we have discussed above in the context of efficient industry organization also provide the basis for pricing transmission services. The cost of a unit of electrical service to a customer at any point in time and at any location in the system is comprised of:

1. The marginal cost of providing the last unit of energy (system lambda);
plus
2. the cost of losses (and other variable costs) associated with delivering energy to any point in the system;
plus
3. the cost of system reliability -- i.e. the cost incurred when strict economic system operation can not meet all of the load. This includes emergency purchases, loading of generators out of the economic merit order to overcome regional generation/transmission capacity shortages, activation of interruptible load contracts, and load shedding.

This defines what is commonly known as the short run marginal cost (SRMC) of electricity service. Strictly speaking, this is the cost per unit supplied at any instant in time, even though in practice, SRMC's are measured over half-hourly or hourly time intervals. Furthermore, SRMC's vary quite significantly over the day, as each of the above components change over time. If a system is correctly designed, SRMC's averaged over the realized states of the world would equal the long run marginal cost (LRMC) of meeting an increment of demand at a particular point in the system. This follows from the basic investment criterion of Net Present Value (NPV) ≥ 0 . In this particular case, what this implies is that investment should be undertaken up to the point when the cost of a new unit of investment and its use should equal the expected cost (over the lifetime of the investment) of the SRMC's. These simple principles of marginal cost give a well founded and defensible basis for a transmission pricing scheme. Figure 9 below illustrates the different elements of SRMC for any given bus in the system (assuming all else unchanged).

It is important to make the distinction here between setting prices equal to short run marginal costs (which would imply pricing in close to real time) and short run marginal cost based pricing. The latter does not require that all participants in the market see and respond to half-hourly or even daily prices, but that the prices charged for the service be aggregated over longer time periods based on what SRMC's are expected to be during that period. In this way SRMC can be, and often is, used as the basis for setting tariffs that hold for time periods ranging from seasons to years. Simple Time Of Use (TOU) or peak / off peak rates are, for instance, aggregations of expected SRMC's.

Figure 9: Elements of SRMC at a Bus in the Transmission Network

Based on the above principles, a methodology for transmission pricing consists of the following steps (see the Technical Appendix for details):

1. Determine allowed revenue level for the system as a whole or a sub-system

The allowed revenue would be determined for a specific period, e.g. one year, or a three-month season. This could be on the basis of embedded costs as at present, or on the basis of incremental costs or opportunity costs if this is permitted.

2. Determine transmission zones

Based on marginal cost maps developed through various system models (such as "MAPPS") other means, divide the region into zones based on marginal cost. Each zone would cluster contiguous load and generation buses with fairly similar marginal costs. The actual number of zones would depend on the level of aggregation/disaggregation that is required. Even though there are considerable seasonal differences in the system which reverse power flows, the zonal configuration itself will be quite stable, since this is determined by the major system constraints.

3. Develop differentiated transmission charges, using marginal costs as an allocative basis

Having set the total revenue level, what remains is to develop a set of transmission charges that would recover this revenue level in a way that meets the desired criteria for transmission tariffs. In allocating revenue requirements to tariffs, there are basically four dimensions along which differentiation is possible across customers:

- a. by location (i.e., transmission zone)
- b. by time (e.g., by season or time-of-day)
- c. by system usage (e.g., by load factor or peak-coincidence)
- d. by reliability (e.g., firm and non-firm).

There are several schemes which may be adopted to allocate charges to different users according to their zonal SRMC's. These should be based on an efficiency rationale such as the following. The basic problem is to recover what, in the short run, are essentially the fixed cost of transmission assets. There is a well-developed literature on alternative methods of allocating such fixed costs which recognizes that various allocation rules have efficiency and sustainability consequences (e.g. [Braeutigam, 1980]). In the present context, these rules would have to be further extended to account for the specific features of transmission assets and usage, including firm and non-firm usage, time-of-use differences in valuation, and assets devoted to system coordination and reliability. We develop in the Technical Appendix one zonal approach in detail, which is quite similar to that currently employed in Norway. The essence of this approach is that it allows ex ante pricing of transmission capacity to all comers on a non-discriminatory basis. It does not reflect the nodal SRMCs since this would overly complicate transmission pricing with very few gains for even abstract efficiency and with considerable loss in transparency in the market. As we argue in the Appendix, the relative magnitude of transmission costs in the electric power value chain are in the range of 10% of total cost. Moreover, only a fraction of this 10% is variable in the short run and therefore includable in the logic of SRMC. Thus, an approximation to SRMC through ex ante zonal pricing with major portions of transmission fixed costs collected through subscription fees sacrifices little if anything on efficiency grounds and gains considerably on market transparency compared to complex approaches to transmission pricing, such as nodal prices with ex post reconciliation procedures.

The approach we recommend in the Appendix is to allocate transmission charges at each bus as a fraction of total revenue requirement (which itself could be zone-specific), with the fraction being determined by an-SRMC based weight at each individual zone. Thus, loads at high SRMC buses would contribute relatively more towards revenue requirements than loads at low SRMC buses. In contrast, generators at high SRMC buses would contribute relatively less towards revenue requirements than generators at low SRMC buses.

Nodal versus Zonal Pricing of Transmission Services

There is a continuing debate about how precise price signals for transmission services must be in time and space in order to reasonably reflect marginal costs and provide accurate market signals. Perhaps the most pointed form this debate has taken¹⁵ is in the discussion of whether full-scale nodal transmission pricing is desirable or whether zonal pricing is on balance, a better candidate for transmission pricing. As we have discussed above, it should be emphasized that the starting point for **development** of zonal transmission tariffs is, indeed, a one-time computation of node-specific marginal costs under various scenarios. Advocates of zonal pricing, such as the authors of this paper, suggest averaging such node-specific marginal costs across relatively homogeneous transmission zones to obtain an average zonal marginal cost to be used as the basis of transmission rates, which are fixed for a reasonable period of time (e.g. one year). Advocates of nodal transmission pricing (e.g., [Hogan, 1992]), on the other hand, prefer to have prices remain at the level of detail of these node-specific marginal costs, usually in real time with prices adjusting (for example) 48 times in a day. Notwithstanding the apparent efficiency benefits of nodal pricing, There are several reasons why zonal pricing is to be preferred. These include, foremost, the following:

(i) Elasticity Shrinkage: The cost of transmission is a relatively small (in the order of 10%) component of the total electricity price. Noting that total price elasticity of electricity demand is low to begin with, the transmission price elasticity of (total) demand is only 10% of (total) price elasticity of demand. In other words, the economic efficiency benefits of fine tuning of transmission pricing signals diminish very rapidly.

(ii) Transaction Costs: Transmission service providers will remain regulated entities for the foreseeable future. The complexities of revenue reconciliation, revenue requirements, comparability reviews and capability assessments are going to be difficult enough in zonal pricing, reset annually or semi-annually. They would appear to be almost impossible under the added complexity of nodal resets. But if transmission prices are to be fixed for a reasonable

¹⁵ For example, in the restructured U.K. electricity supply industry.

length of time, it should be clear that the required scenario averaging across time will not benefit much from the added complexity of having to do this averaging at each node. As a further problem in regulatory complexity, if issues of shareholder and customer cross-flows are raised, these will be more difficult to sort out in a nodal pricing environment than under zonal pricing, where zones and recoverable embedded costs can be clearly identified with respect to native customers and ownership boundaries.

(iii) Market Transparency: The most important role that transmission plays in the evolving electricity market is to facilitate an efficient and active energy market, since this is where most of the benefits of competition are going to come from. From the point of view of the energy market place, stability and transparency of transmission prices will be an important driver of efficiency. Thus, given all the other changes taking place in electric power, simple efficient and stable zonal prices can be an important ingredient for both transmission providers as well as GenCos and DisCos attempting to understand the evolving market place, and make appropriate long and short-term contracts for transmission service.

(iv) Transmission Cost Structure and Stability of Cashflows: Given the desirability of reasonable stability of cashflows from transmission services, zonal pricing provides significant advantages. First of all, it is worth noting that the current cost structure of transmission is largely fixed (although this may change as transmission providers substitute generation and load management contracts for new capacity investment to meet network constraints). Furthermore, if stranded investment recovery and asset revaluation are involved, understanding the interaction of market and regulatory constraints (and arguing credibly for acceptable regulatory relief) required to predict transmission cashflows will be quite difficult under a real-time nodal pricing regime. The existence of risk hedging forward arrangements such as contracts for differences will not obviate the need for settlements on a real-time basis, especially since such forward arrangements cannot be mandated on market participants.

In general, the experience to date with zonal pricing (e.g. in the U.K. which rejected real time nodal transmission pricing and opted for a marginal cost based zonal approach) has indicated that rather stable marginal cost patterns emerge at the zonal level. This would indicate that efficiency gains from nodal transmission pricing may be rather small, even if one neglects the very large and evident transactions costs of regulatory and competitive interactions. Thus, it would seem that a rather substantial burden of proof rests in this case with those who would advocate moving beyond zonal pricing to real time nodal transmission pricing.

7. The Role of Financial Instruments and Intermediation

The expansion of market-based activities in the electricity industry has been accompanied by the growth of financial contracting arrangements. These are for both transactional reasons (as the number of players increases rapidly with the vertical unbundling of the industry) and for the purpose of better allocation of risk across different segments of the industry. We will first discuss the emerging and potential role of financial instruments such as forwards, futures, swaps and options in the new industry structure, turning thereafter to a discussion of the critical role of intermediation in promoting the use of these instruments and enhancing overall efficiency and competition.

Whereas the old vertically integrated structure was dominated by long term (forward) contracts and prices fixed over long time intervals, the new industry structure will be characterized by a more even balance between spot markets and forward contracting arrangements. In this way, prices will be better reflective of the value of services provided and received, and risk can be borne by those who can do so at the cheapest cost.

The risk associated with electricity supply and consumption can be broadly divided into price (financial) and quantity (physical) risk. Price risk arises because the price of electricity fluctuates quite significantly on a temporal basis, much more than other energy commodities such as natural gas or oil. In the England and Wales spot market, the market "clears" each half-hour and prices can vary by an order of magnitude during the course of each 24 hours.¹⁶

¹⁶ For example, on Wednesday, Feb 14, 1996, the provisional England and Wales Pool selling price ranged from £8.94 per MWh at 0600 hours to £109.43 per MWh at 1730 hours (Financial Times, Feb 14, 1996).

Price risk also arises because of spatial price differences in electricity caused by congestion and losses in the transmission system. In the context of contracting and risk management, the lack of perfect correlation in contemporaneous prices at two different locations is termed basis risk.

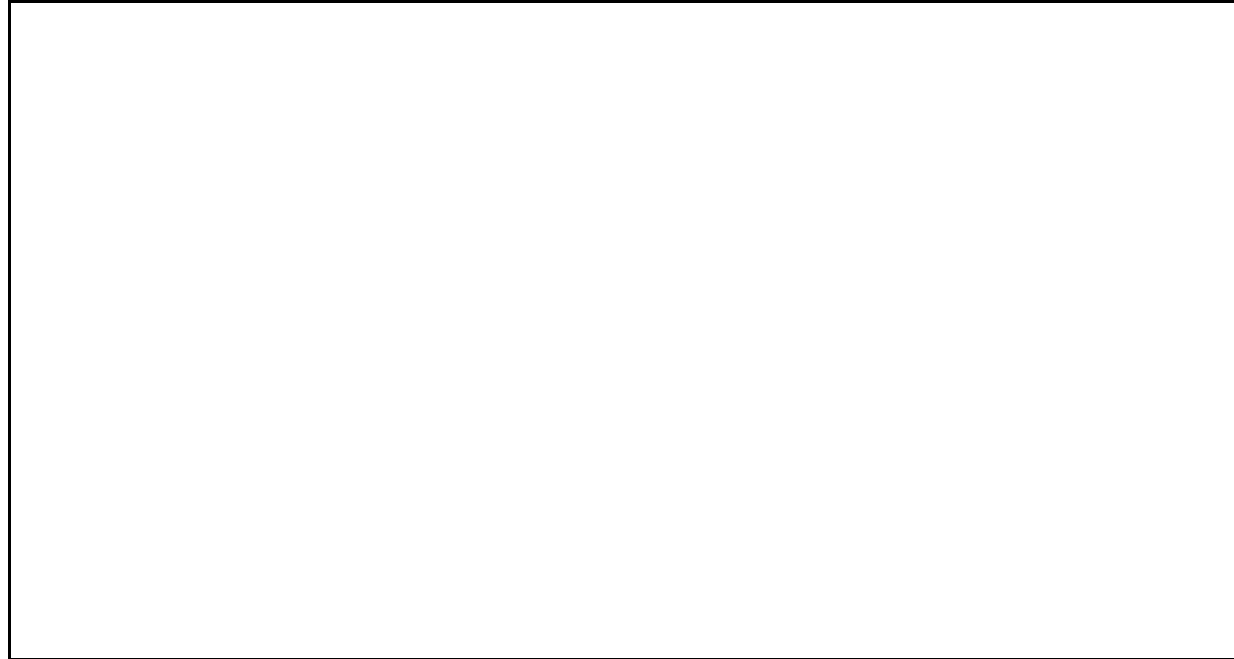
Quantity risk depends on the reliability of supply and demand. While the supply of electricity has traditionally been highly reliable in the U.S. and other industrialized countries, in the new market environment consumers will only pay for the reliability they need and suppliers cannot always be assured of take-or-pay contractual safeguards when demand fluctuates. Whereas in the old industry structure, interruption options (discounts) may have been used as a cover to subsidize supply to certain demand segments (see [Crew and Fernando, 1994]), they will have a very important role in the new market regime in efficiently allocating resources. This is evident from the rapid growth of load management in the England and Wales electricity system after it was unbundled and privatized.

The need to manage these risks will increase the demand for and hence supply of risk management instruments, closely paralleling the process that evolved following the deregulation of the natural gas industry in the 1980's. The first natural gas futures contract was launched by the New York Mercantile Exchange (Nymex) in April 1990 based on the spot natural gas price at the Henry Hub gas pipeline intersection in Louisiana and rapidly emerged as one of the most successful products launched by Nymex.¹⁷ The launch of this futures contract was followed shortly thereafter by various innovative contracting arrangements introduced by Enron and other market intermediaries, including fixed-for-floating swaps, basis swaps and a variety of financial options based on the Nymex futures contract and Henry Hub spot market.

This process has already begun in electricity. An early form of risk management tool introduced in England and Wales with the privatization of the industry was the "Contract for Difference (CFD)". In its most basic form, a CFD is a swap contract between an electricity generator (producer) and supplier (consumer) in which the price in the electricity pool is swapped for a fixed contract price. Under such a CFD, the supplier would pay the fixed contract price to the generator and the generator would pay the half-hourly pool price to the supplier. Since the generator received this price from the pool for its generation and the supplier paid it to the pool, the net effect was to guarantee a fixed (contract) price to both parties. Figure 10 below illustrates the workings of such a basic fixed-for-floating swap in electricity.

Several variations of the CFD are currently in use in England and Wales (see [Hoare, 1995] for a discussion). These CFDs are typically negotiated bilaterally between large generators and suppliers in the UK. A shorter-term contract known as the Electricity Forward Agreement (EFA) has also been introduced in the UK and is conceptually very similar to the fixed-for-spot swap agreement described above except that contract periods are much shorter (e.g. the same four-hour period for one week). Unlike futures contracts, these EFAs are not exchange traded. Unlike forward contracts, EFAs are brokered transactions, with trades being facilitated by electronic screens. Perhaps due to the domination of the bilaterally agreed CFDs, the EFAs have not become widely used in the UK. This highlights one of the limitations associated with the rigid England and Wales industry structure and monopsonistic power pool -- the lack of opportunities for intermediation, traditionally the source of most financial innovations.

¹⁷ For a discussion, see [Fitzgerald and Pokalsky, 1995].



In the United States, the introduction of the first Nymex electricity futures contract was approved a few weeks ago by the Commodity Futures Trading Commission.¹⁸ This futures contract is based on the spot price at the Palo Verde switchyard in Arizona. A second futures contract has been subsequently approved for trading at the California-Oregon border. Thus, a buyer of such a futures contract would (say three months ahead) effectively fix the price which he pays for electricity at the time the contract comes due. At the time of maturity, the futures price converges to the spot price. Thus, if the spot price is higher than the three-month ahead futures price, he would gain the difference between the two, which would exactly offset the increased price that he would have to pay in the spot market. On the other hand, if the spot price is lower, he would lose the difference. In both cases, his effective price is the futures price transacted three months ahead. A consumer wishing to hedge price risk would buy these futures whereas a generator wishing to hedge price risk would sell futures. Like in other futures markets, this market will also be open to speculators and arbitrageurs who will serve to enhance its liquidity and eliminate market inefficiencies.

Once these futures markets begin to operate actively, other financial innovations are likely to follow shortly thereafter. In addition to the physical options discussed earlier associated with various interruption features, call and put financial options written on electricity futures will take positions on the futures price relative to a specific strike price. Market participants can use these options to place either a cap or floor on the electricity price. Thus, typically a buyer of electricity would purchase a call option to ensure that his electricity price does not exceed a specific level (the strike price of the option) whereas a seller of electricity could ensure a floor on the electricity price by buying a put option.

Apart from fixed-for-spot swaps as in the UK or in the US natural gas industry, we would also envision the development of basis swaps which would operate in a manner similar to those in the gas industry. A consumer in Houston will not be able to buy Palo Verde futures contracts to perfectly hedge his price risk since the price of electricity in Houston will not be perfectly correlated with the Palo Verde spot price. By undertaking a basis swap with an intermediary such as Enron, this consumer would pay Palo Verde spot and receive Houston spot, thereby completely eliminating his basis risk associated with the spatial price differences between Houston and Palo Verde. Figure 11 illustrates.

¹⁸ "CFTC OKs First NYMEX Electric Futures Contract", the *Electricity Daily*, Jan. 29, 1996.

As seen from the above example, intermediaries are crucial for the structuring and liquid operation of several of the potential markets in electricity. When the scope for intermediation is limited, it is our view that the scope for market competition is also limited. In concluding this section, we briefly review the role of intermediaries in the new unbundled electricity industry.

In the old vertically integrated structure of the electric utility industry, there was little scope for intermediation, since all transactions along the value chain were internalized within a single company. However, the trends toward emergence of full-fledged intermediation have been evident for some time, paralleling the trends toward greater competition. Power pooling and exchange arrangements across groups of vertically integrated utilities have been a first step in this direction. Whereas these arrangements were originally conceived for reliability reasons, to spread the physical risk of supply shortfalls or demand spikes across a wider base, they have more recently become a means of economizing electricity supply sources in a given region. Furthermore, facilitated by these power pools and wholesale access, transactions across utility boundaries have expanded rapidly, accompanied by the emergence of NUG's and IPP's as significant sources of generation. Some of these transactions have been intermediated by power brokers.

In the new industry structure which is envisioned in this paper, the role of intermediation is expected to expand quite rapidly. This is consistent with the view that intermediation is the "lubricant" of competitive markets, of which we would expect to see a proliferation in the new industry structure. Intermediaries will perform the following key roles in this structure:

a) Intermediate physical transactions

Physical transactions could be intermediated either by bringing buyers and sellers together in brokerage-type transactions, or by acting as dealers for the unbundled services which will be provided in the industry. For example, in the latter case intermediaries could deal in location and time-specific generation capacity or energy. Alternatively, intermediaries could rebundle unbundled services into specific forms as demanded by the marketplace.



b) Intermediate financial transactions

- Intermediate fixed-for-spot and basis swap contracts as illustrated above, to facilitate the management of financial risk associated with spot markets in electricity.

- Facilitate the development of more standardized futures-type financial instruments in electricity, based at high volume "hubs" such as Palo Verde and California-Oregon, across the country.
- Intermediate swap-type arrangements where the delivery of electricity at one point in the system or during a specific time interval will be swapped for delivery at another point or over another time interval.

In broad terms, by engaging in these types of transactions, intermediaries will facilitate the emergence of liquid markets and thereby the strengthening of market forces.

8. Looking to the Future

This paper has focused on the rebundling side of physical and financial transactions in electric power following the unbundling of the basic elements of the electric power value chain to achieve greater transparency and non-discrimination to enable competition. In the process, we have suggested some answers to the key questions of structuring the ISO and pricing transmission services. Our approach is grounded in the realization that transmission services are central to making the market, even though they are not a significant driver of the retail cost of electric power. Thus, the key to a successful transition to an unbundled power market will be to ensure that transmission service is priced in a sufficiently transparent and simple fashion that it facilitates competition in generation and that the ISO is in a position to coordinate bilateral and pool markets in a neutral fashion. In addition, we have discussed various approaches to providing incentives to TAPs and the ISO to invest and to maintain capital stock, to seek out least-cost alternatives for transmission support services and to provide open access to all comers.

A number of open research questions remain, however. These include: models for the efficient integration of long-term (e.g., bilateral energy) and short-term (e.g., spot energy) contracts; models of market intermediation including interface to intermediation for environmental "markets" ; and models for markets involving both firm and non-firm energy use. These models can build on the organizational and ownership principles articulated here.

Appendix on Transmission Pricing

We develop below a general formula for SRMC-based allocation and analyze the impacts of its use. In this scheme of allocation, transmission charges at each bus are allocated as a fraction of total revenue requirement, with the fraction being determined by the SRMC based weight at each individual zone. Thus, loads at high SRMC buses would contribute relatively more towards revenue requirements than loads at low SRMC buses. In contrast, generators at high SRMC buses would contribute relatively less towards revenue requirements than generators at low SRMC buses.

There are several analytical approaches to capturing the general flavor of this logic. We follow the standard Ramsey approach in constructing an analytical approach.¹⁹ We consider two cases based on the organization of the network system:

- ! a **multilateral or "network service"** arrangement in which power is sold to and bought from a common pool, and
- ! a **bilateral or "point-to-point"** arrangement in which a bilateral contract is used for electric power transactions between a generator and a customer.

In the former case, the SRMC at individual buses is used as described above to allocate revenue requirements and derive transmission charges. In the latter case, we derive efficient contracts based on the specific generation and load zones for the contract, and SRMC differentials between these zones.

Efficient Transmission Pricing for Network Service Contracts

We will first consider the case of transmission pricing where all power produced is sold to a common pool from which it is purchased by consumers. It is convenient though not necessary to assume that transmission services are being provided by a single transmission company (GridCo). Energy-related transmission charges (those not collected as connection fees from generators) are borne by consumers and each consumer's cost of transmission is a function of the SRMC at the bus where he is located. Let $P_T(j)$ be the transmission price during time-period T per unit of energy delivered at bus j under optimal system dispatch. According to the standard Ramsey formula, the price paid by customers of transmission service should vary inversely proportionally to the demand elasticities of these customers. Given the level of available information and the size of the typical customers, we will make a first-order approximation that these elasticities are equal across customers. In this case, the Ramsey formula reduces to:

¹⁹ For details on Ramsey pricing, see [Crew and Kleindorfer, 1986]. Note that we are considering multiple owners of transmission assets, and the Ramsey problem here would therefore have a number of breakeven constraints, one for each owner. Rather than pursue this in detail, we separate the problem here into two problems: the first considers efficient pricing to raise sufficient revenues to allow payments to transmission asset owners to allow all transmission losses and costs to be covered. In a second step, we then specify an allocation of these revenues to transmission asset owners which allows each of them to nearly break even, but which also provides some incentives for efficient maintenance and expansion of the transmission network itself.

where $k < 1$ is a positive constant and $C_T(j)$ is the marginal transmission cost (discussed further below) per unit of energy delivered at bus j . We can rewrite this expression as:

where $\mu = 1/(1-k) > 1$. Given that k is to be set so that transmission costs, including a reasonable return on capital, are exactly recovered, we see that $P_T(j)$ must satisfy

where RR^T is the total transmission revenue requirement (discussed below) to be recovered in time interval T . From this, we can solve for μ to obtain

so that, from above,

Let us now briefly discuss the definition of $C_T(j)$ in more detail. Recall that

$C_T(j)$ = Expected value of short-run marginal transmission costs for energy supplied at bus during time interval T . Note that this energy will be supplied from other buses in the network based on its optimal dispatch.

$C_T(j)$ would be estimated, using for example the MAPPS model, by taking expected values across reasonable scenarios which might obtain for the time interval in question. Given the complexity of transmission network costing, it is unlikely that an explicit analytical basis for $C_T(j)$ can be developed for general networks. Intuitively, however, the form of $C_T(j)$ can be written as follows:

where $L_T(j)$ is the expected transmission loss throughout the network per marginal unit of energy extracted at j , λ_T is the expected marginal cost of generation required to supply the transmission losses $L_T(j)$, and $C_{ET}(j)$ represent marginal externalities associated with the supply of a marginal unit at j . Such externalities could be positive or negative and result from such issues as congestion costs, out-of-merit-order operation of plants and other transmission externalities. This expression may be thought of as the expected value of such marginal externality costs plus unit transmission losses times unit generation costs when j increases load by one unit. Given this general structure for $C_T(j)$, the import of the above uniform-elasticity Ramsey structure is to determine unit energy charges for transmission based on a constant mark-up above short-run marginal transmission costs. The higher the losses in serving a given customer from a given supply bus, and the more inefficient the generation which is called into play to make up these losses, the larger will be the transmission price paid.

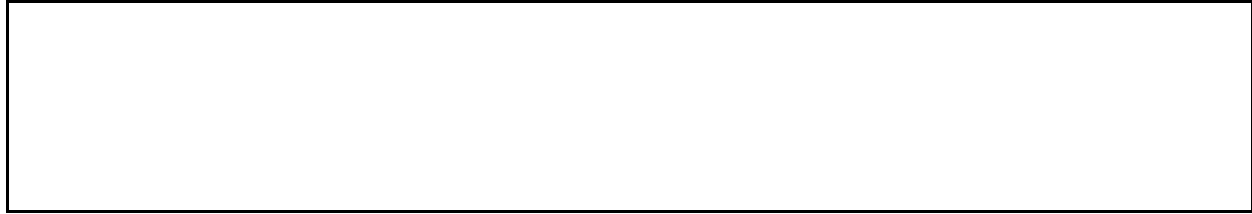
While the above approach provides an economically efficient basis for collection of transmission revenues from consumers on a multilateral basis, it does not directly lend itself to pricing transmission along bilateral contract paths. This is important since many wheeling contracts are negotiated on a bilateral basis between generators and consumers. We turn to this issue next.

Efficient Transmission Charges for Point-to-Point Service Contracts:

We will consider the case of a bilateral contract between a generator at bus i and a customer at bus j . Let $P_T(i,j)$ be the transmission price during time-period T per unit of energy injected at bus i and extracted at bus j under this contract.²⁰ As in the previous case, we can use the Ramsey formula to obtain a basis for pricing transmission service in this bilateral contract. At the outset, note that given our previous definition of SRMC at a bus j , $C_T(j)$, the marginal cost $C_T(i,j)$ of a unit of power injected at bus i and extracted at bus j is simply the difference in marginal costs at the two buses. Thus,

Indeed, moving power from low marginal cost to high marginal cost buses is the means by which transmission adds value to the network. Following arguments that parallel the previous multilateral case, we can obtain the following formula for the transmission price $P_T(i,j)$ between buses i to j :

²⁰ The somewhat awkward language on injection and extraction is required here since, as in the multilateral case considered above, the actual energy supplied to j may not be that injected at i , but will depend on the entire network geometry and power flows at the time of the transaction.



A further generalization, which could be explored if desired, is to set Revenue Requirements that are differentiated by zone, so that the above pricing formulae would be determined on the basis of the revenue recoverable within each zone from intra-zonal transmission.

Note in the above that we do not specify which partner (buyer or seller) to the bilateral contract would actually pay the transmission charges. The point here is that these transmission charges are unbundled charges to be paid by the partners to this contract. They represent the marginal costs imposed on the transmission system to serve the contract plus a markup to recover capital costs and possibly other fixed costs, e.g. some portion of stranded investment costs for one or other of the transmission asset owners.

The above discussion has been framed in terms of transmission charges based on per unit energy flows during T . Clearly, this is equivalent to capacity flows as long as customer load represents a 100% load factor. In practice, for both risk management reasons as well as revenue stability, the transmission component of a bilateral contract will take the form of a two-part tariff, covering energy losses as a percentage of total energy demanded plus a subscription fee per MW of required capacity. It is straightforward to translate the above Ramsey logic into this two-part tariff world. One simply deducts the marginal cost of energy losses from the price implied by the above formulae and the remaining per unit price is to be collected over the time interval T as a subscription fee.

References

- R. Braeutigam, "An Analysis of Fully Distributed Cost Pricing in Regulated Industries", *Bell Journal of Economics*, Vol. 11, Spring, 1980, pp. 182-96.
- Hung-Po Chao and Stephen Peck, "Market Mechanisms For Electric Power Transmission", *Journal of Regulatory Economics*, Vol. 10. July 1996, pp. 25-60
- M. A. Crew and P. R. Kleindorfer, The Economics of Public Utility Regulation, MIT Press, Cambridge, MA, 1986.
- M. A. Crew and P. R. Kleindorfer (eds), Special Issue on Incentive Regulation, *Journal of Regulatory Economics*, May, 1996
- M. J. Doane and D. F. Spulber, "Open Access and the Evolution of the U.S. Spot Market for Natural Gas", *J. of Law and Economics*, Vol. XXXVII (October 1994), pp. 477-517.
- M. Einhorn and R. Siddiqi (eds) *Electricity Transmission Pricing and Technology*, Kluwer Academic Publishers, Boston, 1996.
- C.S. Fernando, P. R. Kleindorfer, R. D. Tabors, F. Pickel and S. J. Robinson, "Unbundling the US Electric Power Industry: A Blueprint for Change", Working Paper, Tabors, Caramanis & Associates, Cambridge, MA, March, 1995.
- Jay Fitzgerald and Joseph Pokalsky, "The Natural Gas Industry" in *Managing Energy Price Risk*, Risk Publications, London, 1995.
- James Hoare, "The UK Electricity Market", in *Managing Energy Price Risk*, Risk Publications, 1995.
- William Hogan, "Contract Networks For Electric Power Transmission", *Journal of Regulatory Economics*, Vol. 4, pp. 211-242, September 1992
- Marija D. Ilic and Shell Liu, Hierarchical Power Systems Control: its Value in a Changing Industry, Springer-Verlag, London, 1996.
- L. Joskow and R. Schmalensee, Markets for Power, MIT Press, Cambridge, MA, 1984.
- David M. Newbery, "Power Markets and Market Power", *The Energy Journal*, Vol. 16, No. 3, 1995, pp. 39-68.
- B. J. Ring and E. G. Read, "A Dispatch Based Pricing Model for the New Zealand Electricity Market", in M. Einhorn and R. Siddiqi (eds), *Electricity Transmission Pricing and Technology*, Kluwer Academic Publishers, Boston, 1996, 183-206.
- N. W. Simmons, L.L. Garver and R. Mukerli, "Transmission Constrained Production Costing: A Key to Pricing Transmission Access", *Public Utilities Fortnightly*, February 1, 1993, pp. 52-55.
- Steve Walton, "Establishing Firm Transmission Rights Using a Rated System Path Model". *The Electricity Journal*, October 1993, pp.20-33

