A COMPETITIVE ELECTRICITY MARKET MODEL

WILLIAM W. HOGAN

Center for Business and Government
John F. Kennedy School of Government
Harvard University
Cambridge, Massachusetts 02138

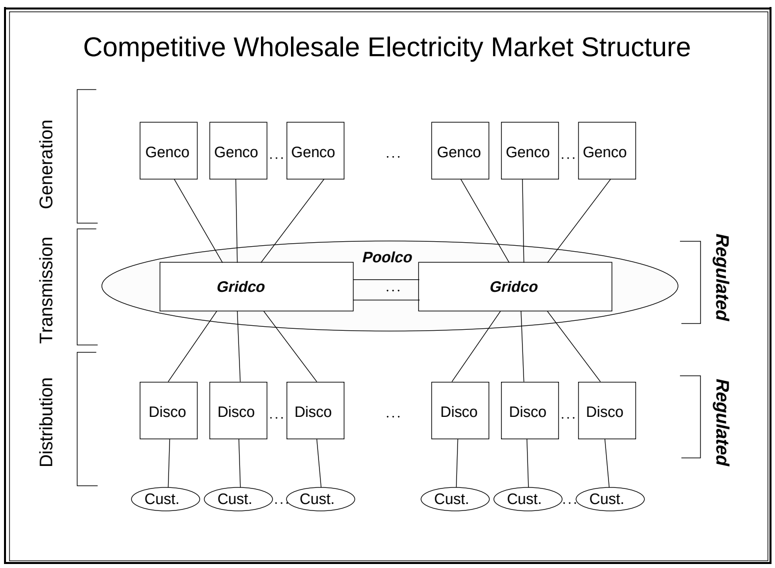
Draft
October 9, 1993

(Prepared for the Harvard Electricity Policy Group)

A COMPETITIVE ELECTRICITY MARKET MODEL

Summary

The electricity market is moving towards greater reliance on competition. Changing technology, new entrants in the generation market, and a legislative mandate to provide access to the essential transmission facility have accelerated a process of competition that will require major changes in the institutions and operations of the electricity market. Because of the special features of electricity supply, however, there are natural monopoly elements in electricity markets. Complete laissez faire competition is not desirable, and in the strictest sense would not be technically feasible. Even though increased competition appears inevitable, therefore, the specifics of how to implement an efficient competitive market are neither inevitable nor obvious. This paper outlines such a competitive market model.



Competitive Wholesale Electricity Market Structure

<u>Industry Organization.</u> The usual separation of the industry distinguishes among generation, transmission and distribution. The Figure illustrates the outlines of a competitive wholesale market structure that follows this traditional three-part segmentation and emphasizes competition in the generation market. The assumption here is that generation is a market with enough real or potential participants to enforce workable competition. Either there are enough separate generating companies to dilute any market power, or the individual generating units are under long-term contracts with many customers such that the economic interest in the plants is dispersed and there is competition in the short run. The need is to specify the minimum elements in the remainder of the system necessary to support competition in generation.

<u>Essential Facilities.</u> The competing generators must have access to the essential facilities that stand between them and their potential customers in the wholesale market. The thrust of the Energy Policy Act of 1992 is that transmission is this essential facility and open access to the transmission system is a necessary requirement for development and operation of a competitive market.

While there is no doubt that such access is required, there is more to transmission than simple connection to the wires. Here the analogy to the case of natural gas pipelines is instructive. There is far more to effective gas transmission access than simple connection to the interstate natural gas pipeline. Movement of gas from producer to consumer includes an array of services such as backup, storage, load balancing, scheduling, and so on. Access to these essential services is as important as physical connection to the pipeline in providing a competitive transmission opportunity for new entrants. If some participants in the market, such as the pipeline companies, have access to these services, while other participants, such as brokers, do not, then it is impossible to maintain the desired "level playing field" of competition. This reality led eventually to the natural gas open access rules in FERC Order 636 which required complete separation of the pipeline merchant function so that all participants in the market would have the same access to the many services that make up the essential monopoly facility.

Just as with natural gas, the essential facility of electricity transmission involves much more than simple connection to the wires. The least obvious feature is the separation of the transmission segment into Poolco and Gridco. The principal reason for the distinction is to accommodate the most important implication of the technical characteristics of electricity supply. In particular, the free-flowing grid requires coordination of short-term operations to maintain system stability and achieve least-cost dispatch. This coordination function operates most efficiently through a power pool which provides many services implicit in the economic dispatch. The dispatch provides an automatic source of backup supplies, short term excess sales, reactive power support, spinning reserve, and the many other services that are bundled in transmission. Without equal access to these many functions, new participants in the market will discover that they are at a competitive disadvantage relative to those who have access to the full array of benefits of a power pool.

Extent of Competition. Limitation of competition to include only the wholesale market is possible but is not necessary. Implicit in the Figure is a separation of the wholesale and retail markets, with regulated distribution companies purchasing in a competitive wholesale market but selling a bundled product to franchise customers. However, once the essential facilities of the pool and the grid are opened to all participants, the market could be extended to retail customers. In this case, the wires component of the distribution business would be recognized as a natural monopoly and an essential facility just as for the high voltage transmission grid. It would be less likely to find a similar dispatch function at the distribution level, and the other elements of the distribution function might be amenable to competitive market provision rather than bundled with wholesale electricity. This choice between wholesale and retail competition is a policy matter. There is no inherent feature of electricity supply that dictates the choice, and any of an array of combinations of open and restricted access would be possible.

Efficient Pricing. With all competing generators enjoying access to the pool dispatch, the natural requirement is for equal pricing for equal services. In the traditional pool with only the vertically integrated companies, average-cost pricing was consistent with the regulatory standard for cost recovery. In the competitive market with third-party access, opportunity-cost pricing provides the standard for consistent incentives and cost recovery in the short-run. Opportunity-cost pricing in the short run is a natural byproduct of an economic dispatch and guarantees the most efficient use of the electrical system.

Long-Term Contracts. The availability of a transparent short-run market with opportunity-cost pricing furnishes the ingredients for bilateral contracts that allocate risk and provide price certainty. In the presence of the short-run market, many variations on the theme of contracts for price differences will arise naturally. Suppliers with generation can sign contracts with customers and provide any desired mix of fixed and variable prices over some extended period. In the day-to-day operation of the market, customers pay and generators receive the short-run opportunity cost. The obligations under the long-term contract are met through strictly financial exchanges that provide the economic equivalent of a specific supply from a specific source. But the contracts need not and do not constrain the operation of the efficient dispatch. And the generation price difference contracts do not require regulatory oversight.

The incentives of short-run opportunity costs and the protection of long-run contracts provide the ingredients for investment in new generation capacity. The price certainty under the contracts makes it possible to craft the protection necessary for long-term financing. In place of the central planners determining reliability standards and reserve margins, the market converts quantity reliability issues into questions of price volatility. The market allows customers to express and pay for their individual preferences for reliability without requiring regulatory supervision. New capacity comes on line as the opportunity cost and price volatility in the short-run market creates the needed economic incentives for building new plants and creating new price hedges.

Transmission Economics. In the face of transmission congestion, economic dispatch produces different short-run prices at different locations. These locational prices capture in a simple form all the complexity of transmission interactions. And just as for generation, these opportunity-cost prices provide the incentives for investment in transmission to relieve congestion and gain access to lower cost supplies. Unlike the case for generation, however, these locational prices do not create a plausible bilateral mechanism for long-term contracts that allocate price risk. Only the pool operator has the information necessary to create the corresponding long-term contracts for price differences across locations. The contracts are multilateral in the sense that all users of the pool affect transmission constraints, and everyone must make a partial payment to compensate those who have paid for the economic benefit of the grid. Fortunately, the payments and the information needed to provide such protection are available automatically as part of the efficient pool dispatch with short-run opportunity-cost pricing. Hence a necessary support for the full competitive market is to provide access to the pool and the grid, and for the pool to operate a system of long-term transmission contracts that protect customer investment in the grid. These transmission contracts are only for payment of price differences across locations, and carry no requirements for determining the actual use of the grid, which is left to the efficient dispatch.

Market Impacts. This outline of the competitive electricity market addresses most of the major economic concerns raised in the discussion of regulatory reform. By definition the system maintains economic dispatch and system stability. Access to the pool and short-run opportunity-cost pricing guarantee comparable service. The obligation to serve changes from one of assuring average-cost supplies to assuring access to the grid and the pool dispatch. Reliability issues change from matters of central planning focussed on quantities to market choices based on prices. The services are not firm or interruptible, but vary depending on the choices worked out under market contracts. Opportunity costs are well defined and investment takes place when the opportunity costs provide sufficient incentives for the participants in the market. Investment decisions are made only in response to customer demands, and there is no regulatory requirement to make generation or network investments without contracts from customers. Transmission complexities are handled without appeal to contract paths, megawatt miles, or wheeling in or out. Secondary markets can operate to assure flexibility and efficiency. Long-term rights to the existing grid can be defined to assure protection of native load. And the system can operate with advances in real-time pricing or minor modifications of existing dispatch settlements systems.

In theory the competitive market outlined would continue to have problems with free-riders exploiting the economies of scale in the transmission grid. Furthermore, the design of the competitive market would not overcome any violations of the assumptions such as market power in generation. If there are social or environmental objectives that have been dealt with through public utility regulation, rather than through taxes or broader based environmental standards, then the competitive market model requires alternatives that focus on the reduced remaining monopolies of the wires and dispatch services.

Transition. The transition problems of moving from the traditional regulatory model to a more competitive market in electricity are a subject for further study. Although the competitive model does not require a "big bang," certain features are essential. Access to the pool and opportunity-cost pricing seem to be minimal requirements. Price difference contracts for generation would arise naturally. With any significant congestion, the pressure would rise for accompanying price difference contracts for transmission, which would have to be created and administered by the pool under regulatory oversight. Allocation of costs and rights for the existing system will require creativity for equity, but with a secondary market there would be no significant impacts on efficiency. Jurisdictional oversight need not require any legislative change, although various existing authorities would have to agree de facto on the accepted framework.

Once the competitive market model is defined, it sets a target and a standard for designing changes from the traditional status quo. The transition problems include many elements that must be considered simultaneously as part of an implementation strategy. Consistency of the broad model and strategy will provide a framework that separates the few necessarily revolutionary elements from the many more evolutionary changes. Analysis of the transition alternatives should focus on open access and pricing rules that are essential for the competitive goal, are feasible as a technical matter in electric networks, and are practical in dealing with stranded assets and other changing economic impacts.

A COMPETITIVE ELECTRICITY MARKET MODEL

William W. Hogan[1]

" 'Would you tell me, please, which way I ought to go from here?'
'That depends a good deal on where you want to get to,' said the Cat.
'I don't much care where--' said Alice.
'Then it doesn't much matter which way you go,' said the Cat." [2]

INTRODUCTION

This paper outlines a model of a reorganized electricity market that makes maximal use of competition. The electric industry and its regulation are in a period of restructuring and reform. A competitive market model reflecting and respecting the distinctive features of electricity supply can serve at least three functions as part of this process of reform. First, a well defined competitive model could guide interim steps during the transition by providing a goal or a destination for the end of the path of policy reform. Second, a consistent, efficient competitive model provides a standard of comparison for evaluating alternative market structure compromises that may be proposed as workable approximations. Third, a competitive market model could serve as a background for testing corporate strategies.

---

[1] Thornton Bradshaw Professor of Public Policy and Management, John F. Kennedy School of Government, Harvard University, and Director, Putnam, Hayes & Bartlett, Inc., Cambridge MA. This paper was prepared as contribution for discussion by the Harvard Electricity Policy Group with support from the Harvard-Japan Project on Energy and the Environment. Helpful comments have been provided by Steve Anderson, Robert Arnold, Ashley Brown, Ralph Cavanagh, Paul Joskow, Henry Lee, Thomas Parkinson, Richard Pierce, Howard Pifer, Larry Ruff, Michael Schnitzer, Jeffrey Tranen, and John White. The author is or has been a consultant on electric market reform and transmission issues for British National Grid Company, General Public Utilities, Duquesne Light Company, Electricity Corporation of New Zealand, National Independent Energy Producers, and Trans Power of New Zealand. The views presented in this paper are not necessarily attributable to any of those mentioned, and the remaining errors are solely the responsibility of the author.

[2] Lewis Carroll, Alice's Adventures in Wonderland, Macmillan, London, 1865.

Although competition has been introduced in parts of the industry, and alternative reforms have been implemented in other countries, there is not complete agreement on the outline of a competitive market model as applied to the case of electricity. Given current technology, complete reliance on a competitive market is not indicated. Only certain segments of the electric industry are now competitive or could be made competitive; other segments are best organized as monopolies and regulated accordingly. In addition, special conditions of the interconnected transmission grid affect the feasibility and efficiency of alternative forms of organization of a "competitive" market for electricity.

The model described here builds on judgments about what is possible under current technology, largely uninhibited by the limitations of current regulatory and market institutions. Most importantly, the assumption is that electric power will continue to be produced most efficiently and chiefly in generating facilities removed from the final customer, connected by a free-flowing transmission grid. A competitive market model for this world must recognize and accommodate the special requirements of the transmission system.

A summary of a few key features of the traditional U.S. market provides a starting point that helps as background and suggests critical that issues must be addressed as part of the process of market reform.

BUNDLED SERVICE AT AVERAGE COST

The traditional view of electricity as a natural monopoly conditioned the development of an industry with several features that are important by way of contrast with the normal operation of a competitive market. Most importantly, the many products and services that constitute production and delivery of electricity to final customers were bundled together in a vertically integrated system of supply. Although there are important differences for certain customers or certain times of use, the steps of generation, transmission and distribution are often integrated in a single firm, and the final customer does not have access to the separate services. Geographic franchise protection accompanies regulation of the vertically integrated firm.

This vertically integrated monopoly has been subject to cost-of-service regulation. From the present perspective, the most important feature of this regulation is reliance on average-cost pricing principles rather than the marginal-cost prices of competitive market theory. At the same time, the electric industry has achieved great efficiency through the use of cooperative efforts such as joint development of new investment, in both generation and transmission, and through operation of coordinated dispatch in formal or informal power pools.

Achieving these efficiency gains has required operating criteria that respect marginal cost, but these internal operating decisions are coupled with prices for the final customer that reflect average cost. This hybrid system has produced complicated practices that seem natural only to those familiar with them through long use. For example, the "split savings" mechanism frequently applied in allocating energy costs in power pools often requires estimation of the results of many hypothetical generation dispatches, to determine the total and individual savings obtained from the pool operation, savings that will be split among the participants according to a sharing formula.[3] On average, customers pay the average cost, but the prices differ for the same product depending on ownership of generating assets. This is only an example, but the complexity of this split-savings mechanism defines a standard which should encourage those proposing alternative market systems that look complex at first glance, but are merely different.

Reliance on average-cost pricing, with the averaging occurring over periods that are long by electrical standards, obscures short-run signals about opportunity costs. Hence administrative decisions are required for many of the most important choices in the electric system, such as setting capacity margins or determining reliability standards that have major influences on overall system cost. In addition, the complications of network interactions have been submerged in the internal operations of a few players in the vertically integrated market, where the acknowledged fiction of the transmission contract path combines with pricing through postage-stamp rates that have had little or no relation to the underlying economics of this essential facility.[4]

Traditionally, transmission capacity in the network is defined in terms of the limits on aggregate interfaces, but the resulting "transfer" limits are known to be load dependent, which requires periodic reallocation of capacity rights to reflect new load conditions.[5] The system worked well in the traditional market, where the few participants could agree on reasonable

---

[3] For instance, the New England Power Pool (NEPOOL) calculates the savings of the actual hourly dispatch compared to each company's hypothetical individual "own load" dispatch and then distributes the savings according to a formula that NEPOOL chairman Robert Bigelow has described as "complicated enough that very few people understand it," (personal communication).

[4] Federal Energy Regulatory Commission, "Transmission Pricing Issues," Staff Discussion Paper, Inquiry Concerning the Commission's Pricing Policy for Transmission Services Provided by Public Utilities Under the Federal Power Act, Washington, DC, June 1993, p. 2.

[5] W. Hogan, "Electric Transmission: A New Model for Old Principles," The Electricity Journal, March 1993, pp. 18-29.

sharing of rights based on fairness and engineering common sense, not on precise payments under contract. But transfer capacity rights for the many new entrants are difficult to define and impossible to guarantee in the aggregate, and the new participants may be unable or unwilling to rely on the good judgement of a future committee that will periodically re-divide the transmission pie.

Alternatives to these and other traditional practices will be required as the industry opens access to the essential facilities. The pressure of competition will force change, and properly designed new institutions could facilitate a transition to and operation of a competitive market.

COMPETITIVE MARKET

By definition, optimal central planning and good regulation can produce an economical, bundled electricity service. Likewise, a competitive market could achieve similar results. The focus here is not to make an argument in favor of either approach. Rather, the assumption is that the explicit policy choice has already been made and, along with the force of technological change, is driving the electricity market towards greater reliance on competition.[6] Because of the special features of electricity supply, however, there are natural monopoly functions and complete laissez faire competition is not desirable. Even though increased competition is inevitable, therefore, the specifics of how to implement an efficient competitive market are neither inevitable nor obvious. The purpose here is to outline the elements that will be critical in developing a more competitive market.

A competitive market includes open access with unrestricted entry by new participants willing to absorb ordinary business risks. For the wholesale electricity market, the Energy Policy Act of 1992 (EPAct) gives the FERC the authority and responsibility to require open access to the electricity transmission system,[7] and open access with comparable service is a standard widely accepted as a key hallmark of the prospective electric market.

The ideal case of the competitive market presumes a large number of competitors with no barriers to entry or exit. Few markets meet all the conditions of the ideal case, with various

---

[6] P.L. Joskow, "Regulatory Failure, Regulatory Reform, and Structural Change in the Electric Power Industry", Brookings Papers on Economic Activity: Microeconomics, Brookings Institution, Washington, DC, 1989, pp. 125-208, provides an overview of the history leading to greater use of competitive forces in the United States. For an update on the challenge ahead, see P. Joskow, "Electricity Agenda Items for the New FERC," The Electricity Journal, June 1993.

[7] Energy Policy Act of 1992, Pub. L. No. 102-486, 106 Stat. 2776, 1992. Section 721 amends section 211 of the Federal Power Act.

deviations in the nature of costs, number of real or potential participants, bundled products, or possibilities for collusive behavior. In practice, the interest is in workable competition, not the perfect case. And workable competition, compared to other realistic and less competitive alternatives, may exist even in the absence of an ideal market. However, the ideal case provides the simple benchmark where participants do not have market power in the sense of being able to maintain sustained and substantial profits that would disappear with significant new entry. Ultimately the competitive market model must be examined as to the degree of workable competition that is feasible in the electricity market.

In the absence of transaction costs or economies of scale and scope, open access and competition are inconsistent with both bundled products and average-cost pricing. In a competitive market with bundled products, new entrants have an incentive to provide the unbundled components of the service whenever the components can be offered to customers at a competitive price. Starting with bundled products and average-cost prices, the natural dynamics of competition lead new entrants to seek out and serve those customers and provide those products which can be supplied at the margin for less than the average cost implicit in the bundled price. This is cream skimming. At the same time, for a profitable firm, price cannot be less than its marginal cost and the process of new entry prevents the price from straying far above aggregate marginal cost. In the limit, therefore, this cream-skimming dynamic tends toward unbundling until all products are provided and customers served at marginal cost.[8] By a related argument, we can arrive at the familiar conclusion that in equilibrium the competitive market must provide the least cost solution for meeting demand. Hence the efficiency of the

[8] Rodney Frame, "Characteristics of a 'Good" Retail Wheeling System," NERA, Electric Utility Business Environment Conference, Denver, Colorado, March, 1993.

competitive market operates inherently with unbundled products and marginal-cost pricing. To the extent that current operations of the industry meet the efficiency test, the individual products and services must be unbundled for critical purposes such as power dispatch.

The products and services bundled in provision of electricity to the final customer are too many and complicated to identify with confidence the exact separation that will evolve in a competitive market. To be sure, transaction costs are not absent, and economies of scale and scope are important in some segments of the electric industry. But substantial unbundling of products and services is to be expected. Surely electricity provided at different times will be treated differently, just as it is today for many customers. Surge protection, backup supply, reactive power, spinning reserve, price hedging, hookups, maintenance, and so on are potential final or intermediate products that could be separated or bundled depending on the economics of supply and the nature of customer demand. Without defining all the details, it is possible to outline segments of the industry that might evolve or should be created as part of a competitive electricity market.

Electricity Market Structure

One possible segmentation of the industry distinguishes those elements that could conform to the competitive model and operate in a private market from those elements that are likely to require some form of price and service regulation.[9] The segments requiring continued regulation isolate the special operating or investment characteristics of electricity and provide the opportunity to examine the essential requirements in the essential facilities.

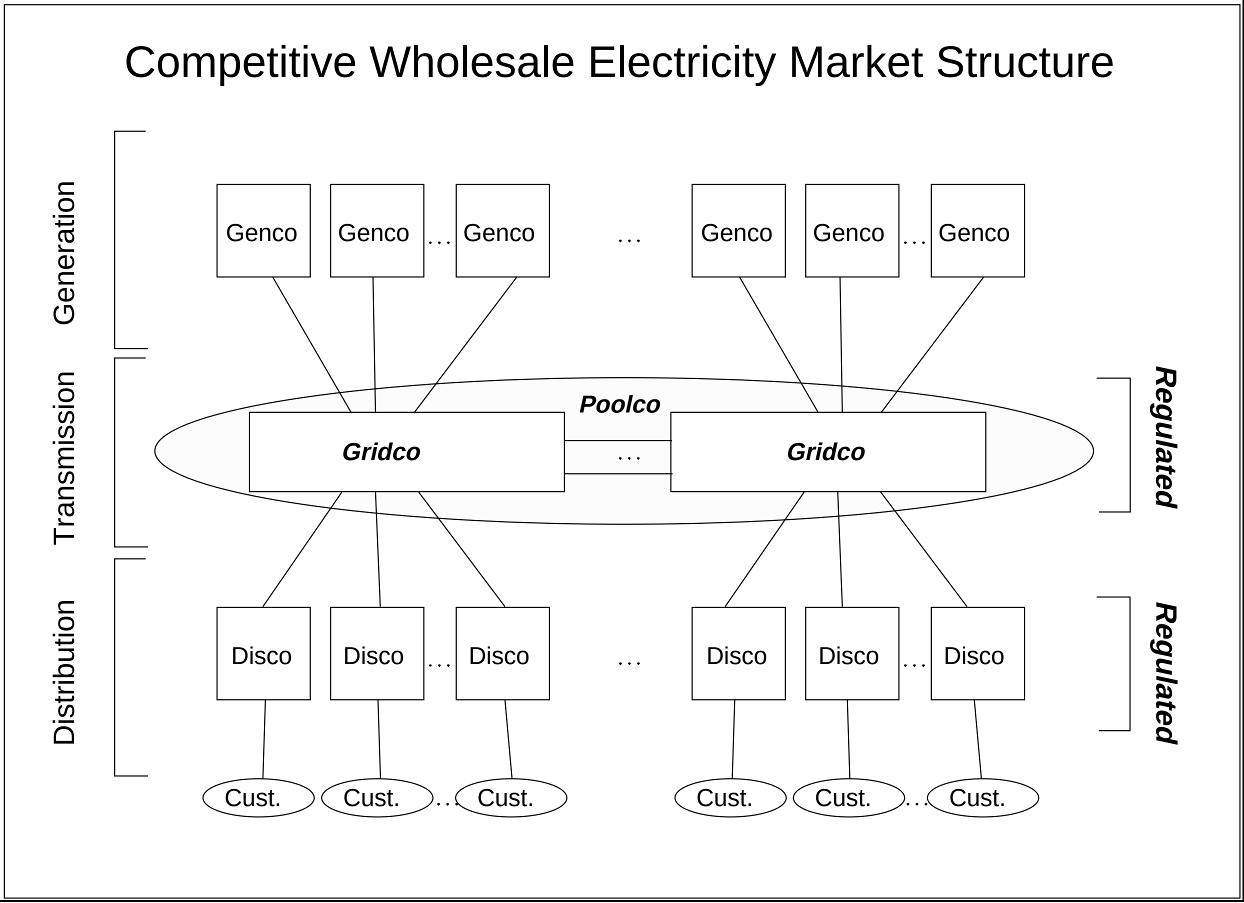
The usual separation of the industry distinguishes among generation, transmission and distribution. Figure 1 illustrates the outlines of a competitive wholesale market structure that follows this traditional three-part segmentation and emphasizes competition in the generation market. The assumption here is that generation is a market with enough real or potential participants to enforce workable competition. Either there are enough separate generating companies to dilute any market power, or the individual generating units operate under long-term contracts with many customers such that the economic interest in the plants is dispersed and there is competition in the short run. The need is to specify the minimum elements in the remainder of the system necessary to support competition in generation.

The competing generators must have access to the essential facilities that stand between them and their potential customers in the wholesale market. The thrust of the EPAct of 1992 is that transmission is this essential facility and open access to the transmission system is a

---

[9] This characterization of one organization of a competitive market relies heavily on the ideas of many others, especially my colleagues at Putnam, Hayes & Bartlett. See, for instance, the report "Establishing Competitive Access to the National Grid -- An Interstate Transmission Access and Pricing Proposal," Putnam, Hayes & Bartlett, March 1992, prepared by principally by Thomas Parkinson and submitted by New South Wales as part of the review of reform options for Australia; L. E. Ruff, "Competitive Electricity Markets: Economic Logic and Practical Implementation," International Association for Energy Economics, 15th Annual International Conference, Tours, France, May 1992 (Revised June 1992).

Figure 1



necessary requirement for development and operation of a competitive market.

While there is no doubt that such access is required, there is much more to transmission than simple connection to the wires. Here the analogy to the case of natural gas pipelines is instructive. There is far more to effective gas transmission access than simple connection to the interstate natural gas pipeline. Movement of gas from producer to consumer includes an array of services such as backup, storage, load balancing, scheduling, and so on. Access to these essential services is as important as physical connection to the pipeline in providing a competitive transmission opportunity for new entrants. If some participants in the market, such as the pipeline companies, have access to these services and other participants, such as brokers, do not, then it is impossible to maintain the desired "level playing field" of competition. This reality led eventually to the natural gas open access rules in FERC Order 636 which required complete separation of the pipeline merchant function so that all participants in the market would have the same access to the many services that make up the essential monopoly facility.[10]

Just as with natural gas, the essential facility of electricity transmission involves much more than simple connection to the wires. With an eye towards identifying these analogous components in the case of electricity, Figure 1 goes beyond the usual Genco, Transco, and Disco trilogy. The least obvious feature is the separation of "Transco" into Poolco and Gridco. This choice is made in part with an eye towards the institutional structure of regulation in the United States, where the Gridco(s) construction and maintenance of transmission wires might be subject to state regulation and Poolco control of dispatch to regional or national supervision. However, the principal reason for the distinction is to accommodate the most important implication of the technical characteristics of electricity supply. In particular, the free-flowing grid requires coordination of short-term operations to maintain system stability and achieve least-cost dispatch. This coordination function operates most efficiently through a power pool which provides many services implicit in the economic dispatch. The dispatch provides an automatic source of backup supplies, short term excess sales, reactive power support, spinning reserve, and the many other services that are bundled in transmission. Without equal access to these many functions, new participants in the market will discover that they are at a competitive disadvantage relative to

[10] R. Pierce, "Reconstituting the Natural Gas Industry from Wellhead to Burnertip," Energy Law Journal, Vol. 9, No. 1, 1988, pp. 1-57. R. Pierce, "Update on Reconstituting the Natural Gas Industry, Draft, September 1993.

those who have access to the full array of benefits of a power pool.[11]

To be effective, therefore, open access to transmission requires open access to the essential facility of the pool dispatch. Although this dispatch coordination could be provided by loosely coupled control centers, it would not be possible to allow unilateral or bilateral operating decisions by all generators or customers connected to the system. Some form of multilateral coordination is essential. For sake of the present discussion, this central operating control is treated as the function of the dispatcher and it is easiest to describe Poolco as a single entity or as a close coordination among a small number of entities.

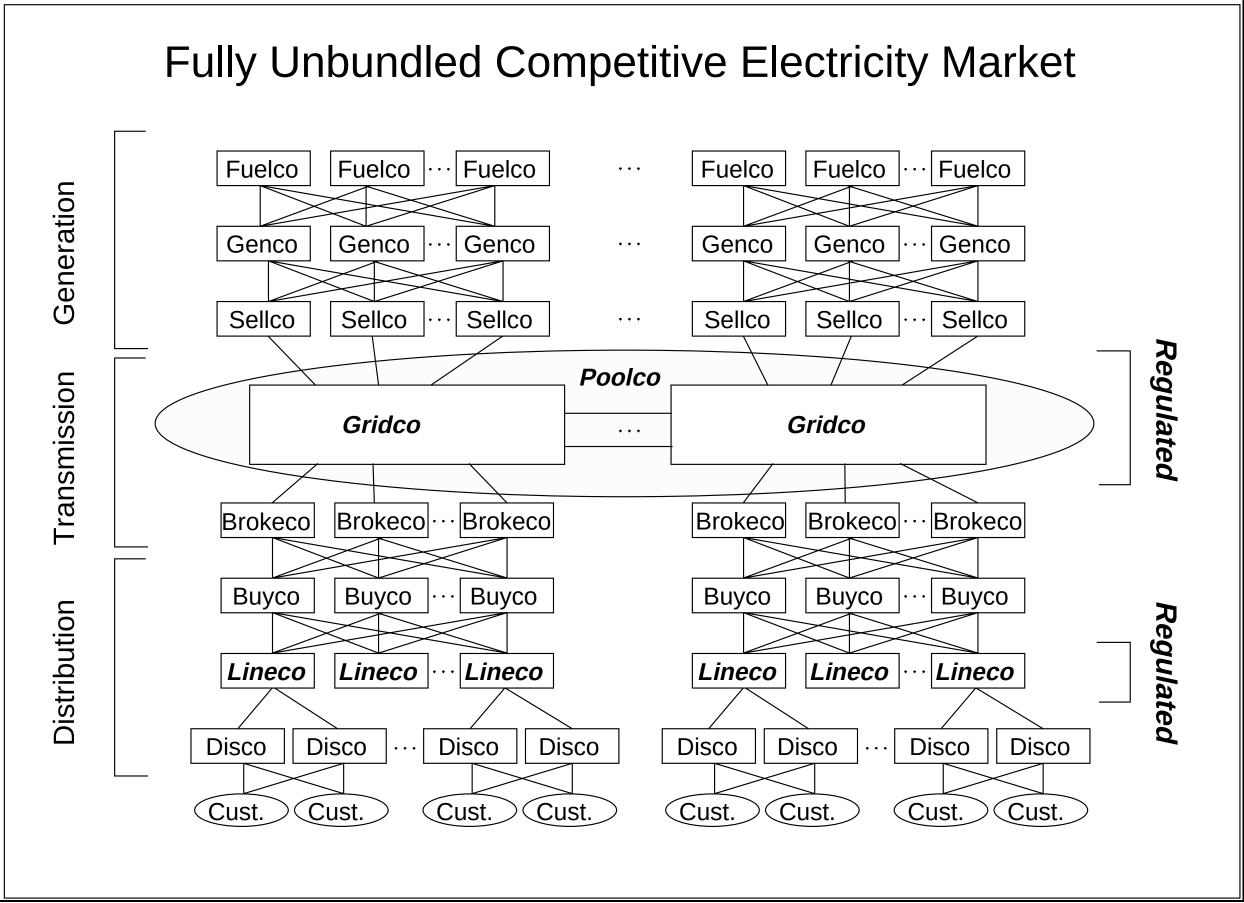
Limitation of competition to the wholesale market is possible but is not necessary. Implicit in Figure 1 is a separation of the wholesale and retail markets, with regulated distribution companies purchasing in a competitive wholesale market but selling a bundled product to franchise customers. However, once the essential facilities of the pool and the grid are opened to all participants, the market could be extended to retail customers. In this case, the wires component of the distribution business would be recognized as a natural monopoly and an essential facility just as for the high voltage transmission grid. It would be less likely to find a similar dispatch function at the distribution level. The other elements of the distribution function might be amenable to competitive market provision rather than bundled with wholesale electricity. This choice between wholesale and retail competition is a policy matter. There is no inherent feature of electricity supply that dictates the choice, and any of an array of combinations of open and restricted access are possible.

For ease in of the present discussion, it is convenient to outline the market structure

---

[11] L. Ruff, "Competitive Electricity Markets: The Theory and Its Application," forthcoming in a Kluwer published volume on electricity markets edited by Michael Einhorn.

Figure 2



that encompasses a more fully unbundled potential, recognizing that a more limited version of competition is possible in principle. As shown in Figure 2, this further segmentation separates possibly competitive functions within generation and distribution. Not all these segments need be distinct; competitive firms may operate in more than one segment as long as competition remains. However, the key segments where technology dictates monopoly or close coordination, separation of control or even ownership, and open access to essential facilities, are indicated as the regulated activities subject to some form of continuing public oversight.

The market segmentation relies on a number of critical assumptions about the nature of technology and the organization of ownership and control. Chief among these are:

- Economies of scale and scope in generation power plants are small relative to the size of the market, and can be achieved as needed through joint ventures that leave control dispersed. Ownership and control of generating assets is widely dispersed.

- Economies of scale and scope are very large for transmission and distribution wires, which are natural monopolies and essential facilities. Regulation of construction and operation of the wires is indicated.

- Central coordination of generation dispatch is essential to preserve the short-run stability and integrity of the system. Economic dispatch is a natural monopoly and regulation is indicated.

- All other services and products exhibit limited economies of scale or scope and could be unbundled and provided by many competing entities.

These assumptions suggest a segmentation of the market into principal functions, distinguished by the reliance on competition or the need for public regulation. The several segments and their main features include:

<u>Generation</u>

Fuelco Purchases fuels for electricity generating plants. There are many sellers and many buyers in regional and national markets.

Genco Operates and maintains existing generating plants. The Gencos interact with the short term market acting on behalf of the plant owners to bid into the short-term power pool for economic dispatch. There are many participants with existing plants and no barriers to entry for construction of new plants.

Sellco Markets long-term power supply compensation contracts to provide price hedges for customers and generators. May also participate in decision making for development of new Gencos. There are many participants and no barriers to entry.

Transmission

Poolco Dispatches existing generating plants and operates a short-term market. Operates a system providing long-term transmission compensation contracts. System control interactions require monopoly operation or close coordination. This segment is regulated to provide open access, comparable service and cost recovery.

Gridco Constructs and maintains the network of transmission wires. Network interaction and scale economies call for monopoly provision and entry barriers. This segment is regulated to provide non-discriminatory connections, comparable service and cost recovery.

Brokeco Matches buyers and sellers as brokers of long-term power supply and transmission compensation contracts. There are many potential participants and no barriers to entry.

Distribution

Buyco Purchasing long-term power supply and transmission compensation contracts for final customers. There are many potential participants and no barriers to entry.

Lineco Constructs and maintains distribution wires connecting transmission grid to final customers. Network interactions and scale economies call for monopoly provision and entry barriers. This segment is regulated to provide non-discriminatory connections, comparable service and cost recovery.

Disco Provides services to final customers including connection and billing. There are many potential entrants and no barriers to entry.

Within this framework, its possible in principle to organize all but three of the segments as unregulated competitive markets. Although some of these segments, such as generation, may have current ownership patterns which lead to concentration of control, this condition is not dictated by the technology. For the three remaining segments--Poolco, Gridco, and Lineco--the technology is such that for the foreseeable future there will be restricted entry and necessary regulation of some form.

To provide an overview of the operation of the market, it is natural to distinguish between the short-run operations managed by Poolco and long-run decisions that include investment and contracting. The system is much simpler in the very short run when it is possible to give meaningful definition to concepts such as opportunity cost. Furthermore, the long run is just a succession of short runs. In ideal competitive markets, without economies of scale and other complications, there is a natural connection between long run and short run that, for example, equates short- and long-run marginal costs in equilibrium. This handy simplification from competitive market theory is assumed, often implicitly, in proposals for the electricity market, as is implicit in proposals for incremental pricing of transmission. However, due to economies of scale in transmission, this handy condition of short- and long-run equilibrium at the margin is not valid in the case of electricity, and it may be poor even as an approximation. Close attention to the connection between short- and long-run decisions, therefore, isolates unique features of the electricity market.

Short-Run Market

The short run is a long time on the electrical scale, but short on human scale--say half an hour. The short-run market is relatively simple. In the short run, locational decisions have been made and power plants, the transmission grid, and distribution lines are all in place. Customers and generators are connected and the work of buyers, sellers, brokers and other service entities is complete. The only decisions that remain are for delivery of power, which in the short-run is truly a commodity product.

On the electrical scale, much can happen in half an hour and the services provided by the system include many details of dynamic frequency control and emergency response to contingencies. Due to transaction costs, if nothing else, it would be inefficient to unbundle all of these services, and many are covered as average costs in the overhead of the system. How far unbundling should go is an empirical question. For example, real power should be identified and its marginal cost recognized, but should this extend to reactive power and voltage control as well? Or to spinning reserve required for emergency supplies? For the sake of the present discussion, focus on real power and remember that reactive power may be unbundled as well, but assume that further unbundling would go beyond the point of diminishing returns in the short-run market.[12]

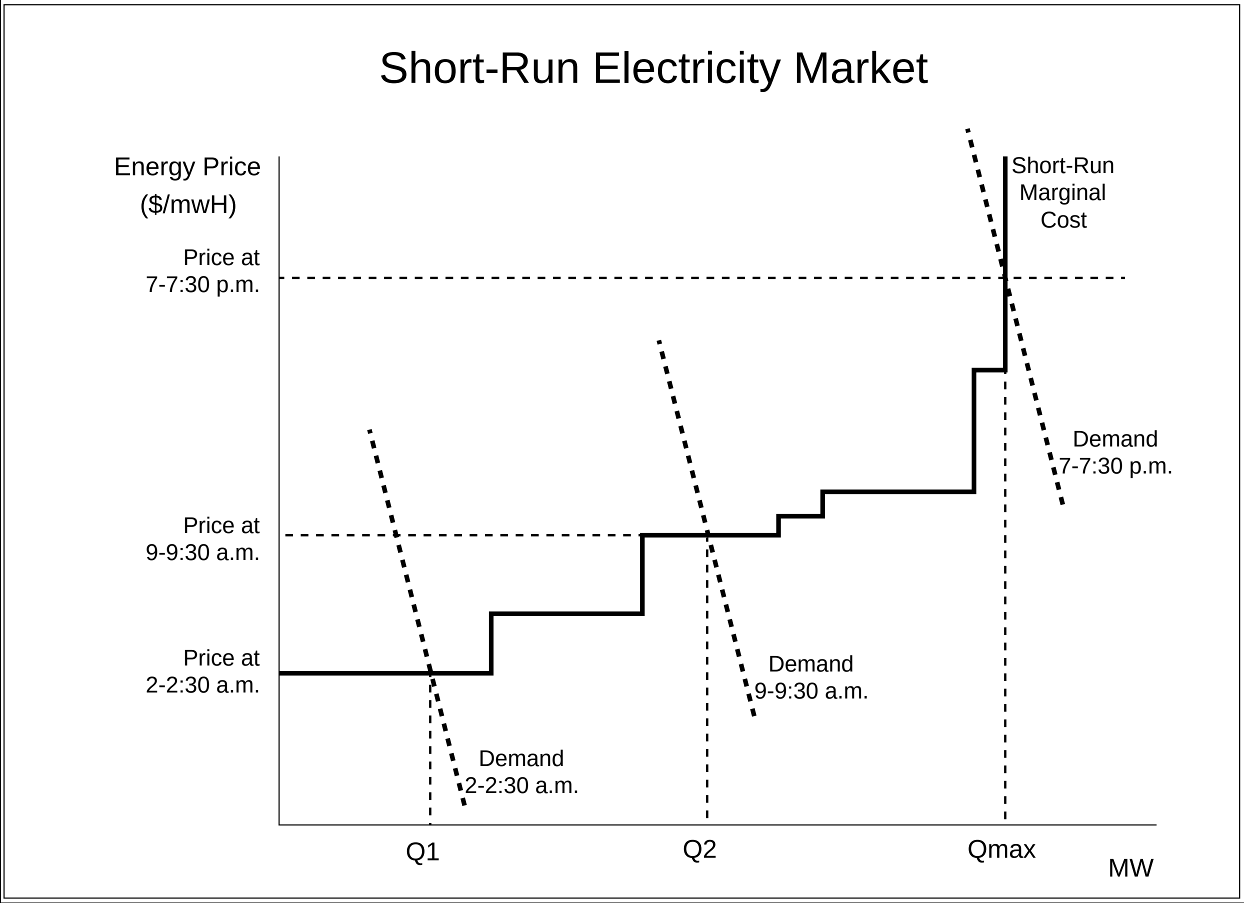
On human scale, over the half hour, the market operates competitively to move real power from generators to customers. Generators have a marginal cost of generating real power from each plant, and customers have different quantities of demand depending on the price at that half hour. With these conditions in a fully decentralized market, with no information costs, customers and generators would search and trade until an equilibrium developed at the market price where supply and demand balance.

This short term market result is illustrated in Figure 3. The collection of generator costs stack up to define the generation "merit order," from least to most expensive. This merit order defines the short-run marginal-cost curve which governs power supply. Similarly, customers have demands which are sensitive to price, and higher prices produce lower demands.

---

[12] W. Hogan, "Markets in Real Electric Networks Require Reactive Prices," Energy Journal, Vol. 14, No. 3, 1993, pp. 171-200. B. Ring, G. Read and G. Drayton, "Optimal Pricing for Reserve Electricity Generation Capacity," Proceedings of the 29th Annual Conference of the Operational Research Society of New Zealand, Auckland, New Zealand, August 1993.

Figure 3



As illustrated simply in Figure 3, the same supply curve is assumed to apply throughout the day, but three different periods have been selected to show different levels of customer demand. In the early morning, there is little demand, and the market equilibrium price settles at the marginal running cost of the cheapest generators. Later in the morning, demand increases and so does the equilibrium price. At this time, every customer that actually consumes power pays the market clearing price, and every generator running is paid this same price. For the generators, the differences between the market price and their individual marginal costs are the short-run profits that make a contribution to recovery of capital.

At the peak period in the evening, the equilibrium price is very high, with all capacity in use. Here the dispatchable demand is setting the market price above the marginal cost even of the last, most-expensive generator, and all generators earn a short-run profit. The equilibrium price still measures the opportunity cost, but at the peak period the marginal opportunity is not to generate more power but rather to forgo that last increment of demand.

This description of the equilibrium economics of a decentralized short-run market could apply to any product with many producers and many consumers. The special complication in the case of electricity arises because the technology does not permit the many leisurely offers, acceptances and trades that are implicit in the search for an equilibrium in the decentralized market model. On the electrical time scale, all those within half-hour dynamics preclude unilateral or bilateral decisions by the participants in the market. Preserving electrical stability and achieving efficiency within the half-hour requires some form of centralized, or at least closely coordinated, dispatch of supply and demand.

This complication presents no insurmountable difficulties, but it does make electricity different from other products. The problem has been there always, and the traditional solution has been to operate the system with a close approximation of centralized control and least-cost dispatch. Generators and customers do not act unilaterally; they provide information to the dispatcher(s) to be used in a decision process that will determine which plants will run at any given half hour. Power pools, especially tight power pools in the United States, provide the model for achieving the most efficient dispatch given the short-run marginal costs of power supply. Although dispatchable demand is not always included, there is nothing conceptually or technically difficult about this extension.[13] The centrally dispatcher controls operation of the system to achieve the efficient match of supply and demand.

In principle, and now in practice in the United Kingdom, this central dispatch can be made compatible with the market outcome. The fundamental principle to exploit here is that for the same load, the least-cost dispatch and the competitive-market dispatch are the same. The principal difference between the traditional power pool and the market solution is the price charged to the customer. In the traditional power pool model, customers pay and generators receive average cost, at least on average. Marginal cost implicitly determines the least-cost dispatch, but customers and generators do not face the marginal cost in the price. Of course, this reliance on average-cost pricing is not necessary. This is apparent from the current practice in the U. K. where the power market operates essentially consistent with Figure 3 but transactions take place at marginal rather than average cost.[14]

An important distinction between the traditional central dispatch and the decentralized market view is found in the source of the marginal-cost information for the generator supply curve. Traditionally the cost data come from engineering estimates of the energy cost of generating power from a given plant at a given time. However, relying on these engineering estimates is problematic in the market model since the true opportunity costs may include other features, such as the different levels of maintenance, that would not be captured in the fuel cost.

---

[13] Demand could be dispatched or with appropriate technology customers could respond to real-time pricing without explicit bids. Real-time pricing goes further than what is necessary for competition and efficient dispatch. For a discussion of real-time pricing, see F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R.E. Bohn, <u>Spot Pricing of Electricity</u>, Kluwer Academic Publishers, Norwell, MA, 1988.

[14] The U.K. market includes other features which cause the price to deviate from pure marginal cost, but the essential element is to determine the half-hourly equilibrium marginal cost which is known as the system marginal price (SMP).

Replacement of the generator’s engineering estimates, that report only incremental fuel cost, with the generator’s market bids is the natural alternative. Each bid defines the minimum acceptable price that the generator will accept to run the plant in the given half hour. And these bids serve as the substitute to guide the dispatch.

As long as the generator receives the market clearing price, and there are enough competitors so that each generator assumes that it will not be providing the marginal plant, then the optimal bid for each generator is the true marginal cost: To bid more would only lessen the chance of being dispatched, but not change the price received. To bid less would create the risk of running and being paid less than the cost of generation for that plant.[15] Hence, with enough competitors and no collusion, the short-run central dispatch market model can elicit bids from buyers and sellers. The dispatcher can treat these bids as the supply and demand curves of Figure 3, and determine the balance that maximizes benefits for producers and consumers at the market equilibrium price. Hence, in the short run electricity is a commodity, freely flowing into the transmission grid from selected generators and out of the grid to the willing customers. Every half hour, customers pay and generators receive the short-run marginal-cost (SRMC) price for the total quantity of energy supplied in that half hour. Everyone pays or receives the true opportunity cost in the short run. Payments follow in a settlements process with a single dispatch and single price that is simple by comparison with the settlements required under the multiple dispatches and multiple costs of traditional split-savings systems.

---

[15] This "incentive compatibility" property of the dispatch auction is not strictly true for those bidders who have a significant chance of defining the marginal price, but if the margin is small the distortion will also be small.

Transmission Congestion

This short-run market model is easy enough and workably approximated in the existing system in the United Kingdom. It could be readily adopted in tight power pools in the United States and elsewhere. However, this model implicitly relies on a critical assumption that all power is generated and consumed at the same location. In reality, generating plants and customers are connected through a free-flowing grid of transmission and distribution lines. The use of this transmission grid affects the short-run market model as summarized in Figure 3.

In the short-run, transmission too is relatively simple. The grid has been built and everyone is connected with no more than certain engineering requirements to meet minimum technical standards. In this short-run world, transmission reduces to nothing more than putting power into one part of the grid and taking it out at another. Power flow is determined by physical laws, but a focus on the flows, whether on a fictional contract path or on more elaborate allocation methods, is a distraction. The simpler model of input somewhere and output somewhere else captures the necessary reality. In this simple model, transmission complicates the short-run market through the introduction of losses and possible congestion costs.

Transmission of power over wires encounters resistance, and resistance creates losses. Hence the marginal cost of delivering power to different locations differs at least by the marginal effect on losses in the system. However, with a few exceptions, the marginal losses on high-voltage transmission grids are relatively small, amounting to only a few per cent of the cost of delivered power, and incorporating these losses does not require a major change in the theory or practice of competitive market implementation. Economic dispatch would take account of losses, and the market equilibrium price could be adjusted accordingly. Technically this would yield slightly different marginal costs and slightly different prices, depending on location, but the basic market model and its operation in the short-run would be preserved.[16]

Transmission congestion is another matter entirely. Limitations in the transmission grid in the short run may constrain long-distance movement of power and thereby impose a higher marginal cost in certain locations. In the simplest case, consider the generators in Figure 3 separated into low cost and high cost groups connected by a single transmission line. For sake of discussion, assume all the customers are located in the high cost region. Hence power will flow over the transmission line from the low cost to the high cost region. If this line has a limit, then in periods of high demand not all the power that could be generated in the low cost region can be used, and some of the cheap plants are "constrained off." In this case, the demand is met by higher cost plants that absent the constraint would not run, but due to transmission congestion are now "constrained on." The marginal cost in the two regions differs because of transmission congestion. The marginal cost of power on the low cost region is no greater than the cost of the cheapest constrained-off plant; otherwise the plant would run. Similarly, the marginal cost in the high cost region is no less than the cost of the most expensive constrained-on plant; otherwise the plant would not be in use. The difference between these two costs, net of marginal losses, is the congestion rental. An appendix illustrates the calculation and outlines the role of multiple contingencies in determining congestions costs in the network.

This congested-induced marginal-cost difference can be as large as the cost of the generation in the unconstrained case. If a cheap coal plant is constrained off and an oil plant which costs more than twice as much to run is constrained on, the difference in marginal costs

---

[16] In the United Kingdom, losses are ignored in setting prices and included in an average cost added to all consumption.

by region is greater than the cost of energy at the coal plant. This result does not depend in any way on the use of a simple case with a single line and two regions. In a real network the interactions are more complicated--with loop flow and multiple contingencies confronting thermal limits on lines or voltage limits on buses--but the result is the same. It is easy to construct examples where congestion in the transmission grid leads to marginal costs that differ by more than 100% across different locations.[17]

If there is transmission congestion, therefore, the short-run market model and determination of marginal costs must include the effects of the constraints. This extension presents no difficulty in principle. The only impact is that the market now consists of a set of prices, one for each location. Economic dispatch will still be the least-cost equilibrium. Generators will still bid as before, with the bid understood to be the minimum acceptable price at their location. Customers will bid also, with dispatchable demand and the bid setting the maximum price that will be paid at the customer's location. The economic dispatch process will produce the corresponding prices at each location, incorporating the combined effect of generation, losses and congestion.[18] In terms of their own supply and demand, everyone sees a single price which is the SRMC price of power at their location. If a transmission price is necessary, the natural definition of transmission is supplying power at one location and using it at another with the corresponding transmission price as the difference between the prices at the

---

[17] W. Hogan, "Markets in Real Electric Networks Require Reactive Prices," Energy Journal, Vol. 14, No. 3, 1993, pp. 171-200.

[18] F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R.E. Bohn, Spot Pricing of Electricity, Kluwer Academic Publishers, Norwell, MA, 1988. W. W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, No. 3, September 1992, pp. 211-242.

two locations.[19]

This short-run competitive market with bidding and centralized dispatch is consistent with least-cost dispatch. The locational prices define the true and full opportunity cost in the short run. Each generator and each customer sees a single price for the half hour, and the prices vary over half hours to reflect changing supply and demand conditions. All the complexities of the power supply grid and network interactions are subsumed under the economic dispatch and calculation of the locational SRMC prices. These are the only prices needed, and payments for short-term energy are the only payments operating in the short run, with administrative overhead covered by rents on losses or, if necessary, a negligible markup applied to all power. The dispatch and settlements process are handled by Poolco, with regulatory oversight to guarantee comparable service through open access to the pool. Something like this system is necessary, and a pool operation is the natural mechanism:

> "The process of defining the comparability standard will dominate the electric transmission service debate just as it did in the case of natural gas. And for many reasons having to do with backup, balancing and so on, a natural resolution of this debate will be to give all eligible producers and customers equal access to a tight power pool, with the pool operating to provide economic dispatch."[20]

---

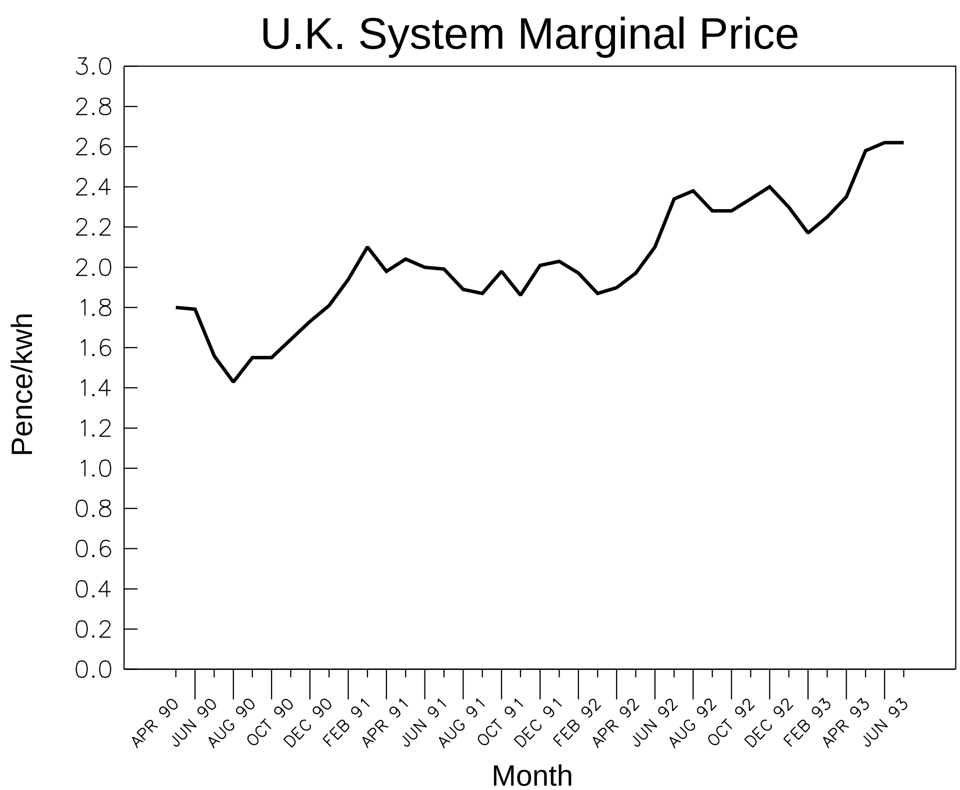
[19] Locational prices reflecting short-run generation, losses and congestion are part of the proposed transmission pricing regime for New Zealand as described in Trans Power New Zealand "Transmission Pricing 1993," , Wellington, New Zealand, February 1993.

[20] L. E. Ruff, "Competitive Electricity Markets: Economic Logic and Practical Implementation," International Association for Energy Economics, 15th Annual International Conference, Tours, France, May 1992 (Revised June 1992).

Long-Run Market Contracts

With changing supply and demand conditions, generators and customers will see fluctuations in short-run prices. When demand is high, more expensive generation will be employed, raising the equilibrium market prices. When transmission constraints bind, congestion costs will change prices at different locations.

Figure 4



Source: U.K. Office of Electricity Regulation, Pool Price Statement, July 1993

Even without transmission congestion constraints, the spot market price can be volatile. In the U. K., the "system marginal price" for generators, as shown in Figure 4, is calculated only in terms of the unconstrained dispatch. The figure shows the volatility in average monthly prices. The changes within a day or over the month have been greater, sometimes an order of magnitude greater. This volatility in prices presents its own risks for both generators and customers, and there will be a natural interest in long-term mechanisms to mitigate or share this risk. The choice in the market is for long-term contracts.

Traditionally, and in many other markets, the notion of a long-term contract carries with it the assumption that customers and generators can make an agreement to trade a certain amount of power at a certain price. The implicit assumption is that a specific generator will run to satisfy the demand of a specific customer. To the extent that the customer’s needs change, the customer might sell the contract in a secondary market, and so too for the generator. Efficient operation of the secondary market would guarantee equilibrium and everyone would face the true opportunity cost at the margin.

However, this notion of specific performance stands at odds with the operation of the short-run market for electricity. To achieve an efficient economic dispatch in the short-run, the dispatcher must have freedom in responding to the bids to decide which plants run and which are idle, independent of the provisions of long-term contracts. And with the complex network interactions, it is impossible to identify which generator is serving which customer. All generation is providing power into the grid, and all customers are taking power out of the grid. There are no bilateral transactions and there is no way to operate a secondary market in the actual deliveries of power. It is not even in the interest of the generators or the customers to restrict the dispatch and forego the benefits of the most economic use of the available generation. The short-term dispatch decisions are made independent of and without any recognition of any long-term contracts. In this way, electricity is not like other commodities.

This dictate of the physical laws governing power flow on the transmission grid does not preclude long-term contracts, but it does change the essential character of the contracts. Rather than controlling the dispatch and the short-run market, long-term contracts focus on the problem of price volatility and provide a price hedge not by managing the flow of power but by managing the flow of money. The short-run prices provide the right incentives for generation and consumption, but create a need to hedge the price changes. Recognizing the operation of the short-run market, there is an economic equivalent of the long-run contract for power that does not require any specific plant to run for any specific customer.

Consider the case first of no transmission congestion. In this circumstance, except for the small effect of losses, it is possible to treat all production and consumption as at the same location. Here the natural arrangement is to contract for differences against the equilibrium price in the market. A customer and a generator agree on an average price for a fixed quantity, say 100 MW at five cents. On the half hour, if the pool price is six cents, the customer buys power from the pool at six cents and the generators sells power for six cents. Under the contract, the generator owes the customer one cent for each of the 100 MW over the half hour. In the reverse case, with the pool price at three cents, the customer pays three cents to Poolco, which in turn pays three cents to the generator, but now the customer owes the generator two cents for each of the 100 MW over the half hour.

In effect, the generator and the customer have a long-term contract for 100 MW at five cents. The contract requires no direct interaction with Poolco other than for the continuing short-run market transactions. But through the interaction with Poolco, the situation is even better than with a long-run contract between a specific generator and a specific customer. For now if the customer demand is above or below 100 MW, there is a ready and an automatic secondary market, namely the pool, where extra power is purchased or sold at the pool price. Similarly for the generator, there is an automatic market for surplus power or backup supplies without the cost and problems of a large number of repeated short-run bilateral negotiations with other generators. And if the customer really consumes 100 MW, and the generator really produces the 100 MW, the economics guarantee that the average price is still five cents. Furthermore, with the contract fixed at 100 MW, rather than the amount actually produced or consumed, the long-run average price is guaranteed without disturbing any of the short-run incentives at the margin. Hence the long-run contract is compatible with the short-run market.

The price of the generation contract would depend on the agreed reference price and other terms and conditions. Generators and customers might agree on dead zones, different up-side and down-side price commitments, or anything else that could be negotiated in a free market to reflect the circumstances and risk preferences of the parties. Whether generators pay customers, or the reverse, depends on the terms. However, Poolco need take no notice of the contracts, and have no knowledge of the terms. Just such contracts have emerged in the U. K. market to provide price hedges against fluctuations in the pool price.

In the presence of transmission congestion, the generation contract is necessary but not sufficient to provide the necessary long-term price hedge. A bilateral arrangement between a customer and a generator can capture the effect of aggregate movements in the market, when the single market price is up or the single market price is down. However, transmission congestion can produce significant movements in price that are different depending on location. If the customer is located far from the generator, transmission congestion might confront the customer with a high locational price and leave the generator with a low locational price. Now the generator alone cannot provide the natural back-to-back hedge on fluctuations of the short-run market price. Something more is needed.

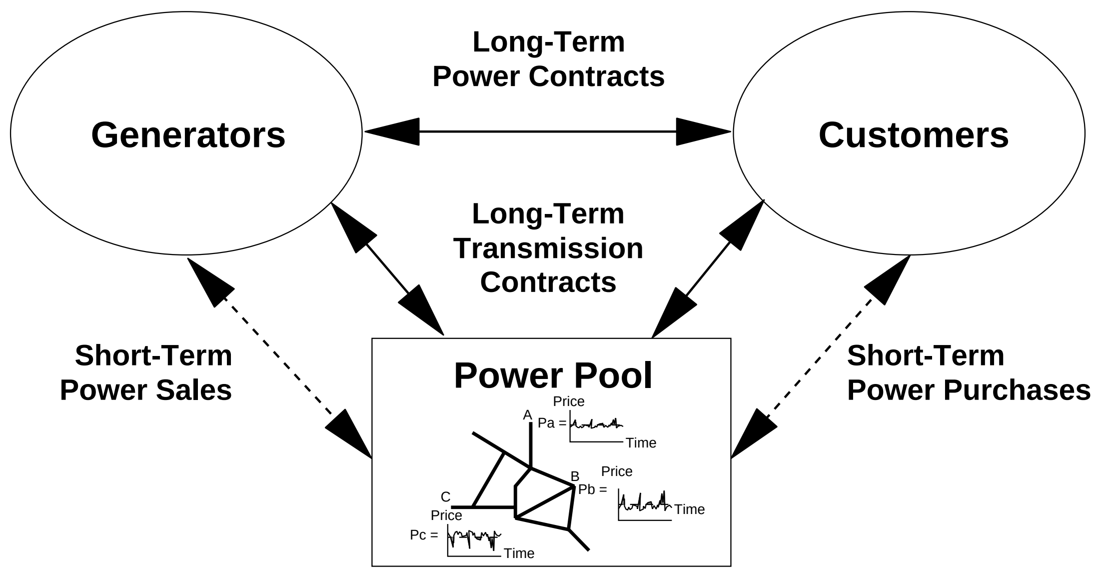
Transmission congestion in the short-run market creates another related and significant problem for Poolco. In the presence of congestion, revenues collected from customers will substantially exceed the payments to generators. The difference is the congestion rent that accrues because of constraints in the transmission grid. At a minimum, this congestion rent revenue itself will be a highly volatile source of payment to Poolco. At worse, if Poolco keeps the congestion revenue, incentives arise to manipulate dispatch and prevent grid expansion in order to generate even greater congestion rentals. Poolco is a natural monopoly and could distort both dispatch and expansion. If Poolco receives the benefits from congestion rentals, this incentive would work contrary to the goal of an efficient, competitive electricity market.

The convenient solution to both problems--providing a price hedge against locational congestion differentials and removing the adverse incentive for Poolco--is to re-distribute the congestion revenue through a system of long-run transmission contracts operating in parallel with the long-run generation contracts. Just as with generation, it is not possible to operate an efficient short-run market that includes transmission of specific power to specific customers. However, just as with generation, it is possible to arrange a transmission contract that provides compensation for differences in prices, in this case for differences in the congestion costs between different locations across the network.

The transmission capacity contract for compensation would exist for a particular quantity between two locations. The generator in the example above might obtain a transmission contract for 100 MW between the generator's location to the customer's location. The right provide by the contract would not be for specific movement of power but rather for payment of the congestion rental. Hence, if a transmission constraint caused prices to rise to six cents at the customer's location, but remain at five cents at the generator's location, the one cent difference would be the congestion rental. The customer would pay Poolco six cents for the power. Poolco would in turn pay the generator five cents for the power supplied in the short-run market. As the holder of the transmission contract, the generator would receive one cent for each of the 100 MW covered under the transmission contract. This revenue would allow the generator to pay the difference under the generation contract so that the net cost to the customer is five cents as agreed in the bilateral power contract. Without the transmission contract, the generator would have no revenue to compensate the customer for the difference in the prices at their two locations. The transmission contract completes the package.

When only the single generator and customer are involved, this sequence of exchanges under the two types of contracts may seem unnecessary. However, in a real network with many participants, the process is far less obvious but the net result is the same. Short-run incentives at the margin follow the incentives of short-run opportunity costs, and long-run contracts operate to provide price hedges against specific quantities. The structure of this market with contracts is illustrated in Figure 5. Poolco operates in the short-run market to provide economic dispatch. Poolco collects and pays according to the short-run marginal price at each location, and Poolco distributes the congestion rentals to the holders of transmission capacity rights. Generators and customers make separate bilateral arrangements for generation contracts. Unlike with the generation contracts, Poolco's participation in the transmission contracts is necessary because of the network interactions which make it impossible to link specific customers paying congestion

Figure 5



costs with specific customer receiving congestion compensation. In the aggregate the total congestion payments received by Poolco will fund the congestion payment obligations under the transmission contracts, but the congestion prices paid and received will be highly variable and load dependent. Only Poolco will have the necessary information, but the information will be readily available embedded in all the pool's locational prices.[21]

If the pool transmission capacity rights have been fully allocated, then Poolco will be

---

[21] W. W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, No. 3, September 1992, pp. 211-242. These transmission contract rights define directional rights that guarantee protection from changes in congestion rentals. Additional congestion payments from the grid to rights holders may be necessary to pass through all the congestion rents inherent in short-run, locational, marginal-cost prices.

simply a conduit for the distribution of the congestion rentals. Poolco will no longer have an incentive to increase congestion rentals, because any increase would flow only to the holders of the transmission rights, not to Poolco. The problem of supervising the Poolco and Gridco monopolies would be greatly reduced. And through a combination of generation and transmission contracts, participants in the electricity market can arrange price hedges that could provide the economic equivalent of a long-term contract for specific power delivered to a specific customer.

Much of the work of the Sellco, Brokeco and Buyco intermediaries would consist of making arrangements for these price hedging contracts. There is no added challenge to selling, delivering or buying the actual energy in the half hour; all this happens automatically in the operation of the short-run power pool. The more complicated long-run products are the portfolios of compensation contracts to hedge movements in generation costs and congestion costs. These contracting businesses might handle a large amount of money, but they would be easy to enter and the margins would be thin, especially for the brokers in the middle with no connection to buyers or sellers.

Long-term Market Investment

Within the contract environment of the competitive electricity market, new investment occurs principally in generating plants, customer facilities and transmission expansions. In each case, corresponding contract-right opportunities appear that can be used to hedge the price uncertainty inherent in the operation of the competitive short-run market.

In the case of investment in new generating plants or consuming facilities, the process is straightforward. Under the competitive assumption, no single generator or customer is a large part of the market, there are no significant economies of scale and no barriers to entry. Generators or customers can connect to the transmission grid at any point subject only to technical requirements defining the physical standards for hookup. If they choose, new customers or new Gencos have the option of relying solely on the short-run market, buying and selling power at the locational price determined as part of the half-hourly dispatch. Poolco itself makes no guarantees as to the price at the location. Poolco only guarantees open access to the pool at a price consistent with the equilibrium market. The investor takes all the business risk of generating or consuming power at an acceptable price.

If the generator or customer wants price certainty, then new generation contracts can be struck between a willing buyer and a willing seller, possibly through the intermediation of Sellco, Brokeco or Buyco. The complexity and reach of these contracts is limited only by the needs of the market. Typically we expect a new generator to look for a customer who wants a price hedge, and the generator defers investing in new plant until sufficient long-term contracts with customers can be arranged. The generation contracts can be with one or more customers and may involve a mix of fixed charges coupled with the obligations to compensate for price differences relative to the pool price. But the customer and generator will ultimately buy and sell power at their location at the half-hourly price.

If either party expects significant transmission congestion, then a transmission capacity contract would be indicated. If transmission capacity rights are for sale between the two points, then a contract can be obtained from the holder(s) of existing rights. Or new investment by Gridco can create new transmission capacity rights. In the case of transmission investment, economies of scale and network interactions loom large, unlike the case assumed for generation. Hence, because of economies of scale it is expected that for any given transmission investment there will be a material change in the pool prices through reduced congestion rentals. In addition, the network interactions will create many potential beneficiaries.

These facts typically will require that any transmission expansion be organized by a consortium of transmission investors who negotiate a long-term contract that allocates the fixed cost of the investment and the corresponding allocation of new transmission capacity rights. The Gridco, as a regulated monopoly, builds the lines in exchange for a payment that covers the capital cost and a regulated return. Gridco does not make transmission investments without long-run contracts signed by willing customers who will pay the fixed costs and recover any future congestion revenues. The Poolco participates in the process only to verify that the newly created transmission capacity rights are feasible and consistent with the obligation to preserve the existing set of rights on the existing grid. Unlike the ambiguity in the traditional definition of transmission transfer capacity, there is a direct test to determine the feasibility of any new transmission capacity rights for compensation, while protecting the existing rights, and the test is independent of the actual loads.[22] Hence, incremental investments in the grid are possible anywhere without requiring that everyone connected to the grid participate in the negotiations or agree to the allocation of the new transmission capacity rights.

Investments by Linecos to expand the local distribution system follow the same principles as for Gridcos, and the distinction between the transmission grid and distribution lines

---

[22] The test is that the net loads implicit in the capacity rights are feasible in the absence of any other loads on the system. See W. W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, No. 3, September 1992, pp. 211-242.

is not essential; the two could operate as the same regulated entity. However, the complexities of network interactions are related to loop flows. To the extent that there are local distribution lines that isolate the power flow from the grid, with no loops back to the grid, it may be efficient to organize and operate these businesses as separate local monopolies. The structure of short- and long-run pricing is the same as for the transmission grid. Short-run usage is charged at short-run marginal cost, including losses and the effects of any local congestion. Long-run investments are made under contracts with customers who agree to pay the fixed charges for capital return and recovery. These investments create local Lineco-managed capacity rights to collect local Lineco congestion rentals, should they arise. The local functions analogous to the pool dispatch, perhaps in controlling local demand, are solely limited to the local lines, having by definition have no network effects, and can remain as part of the functions of the regulated Lineco.

Other Services

The many other services that make up the traditional bundled delivery of electricity can be provided by the various competitive companies. For example, the Disco can connect customers, provide billing services, provide customer surge protection, and so on. In practice, Buycos and Discos may be the same company. As long as there are no barriers to entry and such companies operate in a competitive environment, prices and service quality are regulated only by the market. However, the Discos must be operated independently of the Linecos, which do have a franchise and a local monopoly, and are regulated accordingly. The separation should be at least in terms of control, with the Linecos providing "comparable service" to all the Discos.

In practice this separation may require separation of ownership.

Table I summarizes these components of this competitive electricity market. This structure and its key elements--access to essential facilities including the wires and pool dispatch, use of short-run marginal-cost pricing, and reliance on long-term contracts to provide economic hedges rather than specific performance--are not far from actual operations or proposed reforms in other systems. This competitive market is the essence of the design of the current U. K. system, with the notable difference of the lack of locational short-run prices.[23] The essential elements of this market model are close to the proposed reforms in New Zealand.[24]

---

[23] S. Littlechild, "Competition, Monopoly and Regulation in the Electricity Industry," U. K. Office of Electricity Regulation, June 1993. U. K. Office of Electricity Regulation, "Pool Price Statement," Birmingham, England, July 1993. The U. K. design is plagued by incentive problems created in unusual and unnecessary features of the calculation of market prices coupled with too few generators to ensure competition.

[24] Wholesale Electric Market Study, "A Managed Transition Toward a Facilitated Market: Rationale," Wellington, New Zealand, October 1992. Trans Power New Zealand "Transmission Pricing 1993," , Wellington, New Zealand, February 1993.

Table I: Competitive Electricity Market Model				
Segment	**Function**	**Structure**	**Regulation**	**Jurisdiction**
Fuelco	Purchases fuels for generating plants	Many sellers and many buyers in regional and national markets	Competitive market	Unregulated
Genco	Operates and maintains generating plants	Many participants with existing plants and no barriers to entry	Competitive market	Unregulated
Sellco	Markets long-term power supply compensation contracts	Many participants and no barriers to entry	Competitive market	Unregulated
Poolco	Dispatches generating plants and operates a short-term market with long-term transmission compensation contracts	System control interactions require monopoly operation or close coordination	Regulated to provide open access, comparable service and cost recovery	***Regional or national***
Gridco	Constructs and maintains transmission wires	Network interaction and scale economies call for monopoly provision and entry barriers	Regulated to provide non-discriminatory connections, comparable service and cost recovery	***Regional or state***
Brokeco	Matches buyers and sellers as brokers of long-term power supply and transmission compensation contracts	Many potential participants and no barriers to entry	Competitive market	Unregulated
Buyco	Purchases long-term power supply and transmission compensation contracts for final customers	Many potential participants and no barriers to entry	Competitive market	Unregulated
Lineco	Constructs and maintains of distribution wires	Network interactions and scale economies call for monopoly provision and entry barriers	Regulated to provide non-discriminatory connections, comparable service and cost recovery	***State***
Disco	Provides services to final customers including connection and billing	Many potential entrants and no barriers to entry	Competitive market	Unregulated

COMPARING MARKETS

Moving to the competitive market model outlined here would result in a variety of changes relative to the traditional regulatory market. Recognizing these differences, or in some cases the similarities, assists in evaluating transition strategies and designing approximations to the competitive ideal.

Economic Dispatch

Economic dispatch of the traditional central planning model is preserved in the competitive market based on Poolco. The centerpiece of the competitive market is a centrally coordinated dispatch built on customer bids. Poolco manages this dispatch to protect against multiple contingencies and achieve the least-cost dispatch and determine the market-clearing locational prices based on short-run marginal costs. With many competitors on both the supply and demand side, there is no chance for any single player to affect the market price, and the resulting bidding incentives align the bids with opportunity costs. Hence the Poolco dispatch is the efficient economic dispatch in the short-run.

Comparable Service

In the unbundled competitive market, the regulatory obligation to define and require "comparable service" is limited to those services provided by the residual monopolies. With free entry in buying, brokering, billing, and so on, a competitive market can discipline and guide the selection and pricing of products and services in the competitive segments. The residual monopolies are for Poolco, Gridco and Lineco. Here the principal services are in providing full participation in the dispatch and connections to the wires. Once this access is available, the customers and suppliers can compete in the competitive segments of the market. Since there is a separation of the regulated monopoly function from all other services, there is little difficulty in defining or enforcing comparability.

Of course, comparability does not mean that all services or all prices are the same. For example, customers in the pool are at different locations and purchase or supply power at different times. Each of these--different locations and different times--leads to different marginal costs that should be reflected in competitive prices. Price differences do not imply undue discrimination; price differences imply only the discrimination inherent in marginal costs.

Obligation to Serve

Under a competitive market organization, the obligation to serve differs dramatically from the traditional regulatory regime. In the case of regulated, average-cost pricing and service, central decision makers must make the choices that provide the appropriate balance of supply and demand. Most notably, the electric utility is obliged to serve the customer at the average cost of supply, even though the marginal cost of meeting that demand may be different and perhaps much higher than the average cost. With the marginal cost greater than the average price, it is necessary to impose a legal obligation to make sure that the demand will be met.

In the competitive market, however, with customers paying the marginal cost of electricity, there is no imbalance at the margin. With the proper definition of marginal cost, any customer can have as much or as little electricity as demanded at that price. Hence there is no necessity to impose an obligation on an electric utility to provide electricity or any other service available in the competitive spheres of the market. In these competitive domains, where customers have choices and pay marginal costs, the market balances supply and demand.

The only requirement for an obligation to serve, therefore, arise in the non-competitive, regulated portion of the market. This is limited to the Poolco, Gridco, and Lineco segments of the industry. Here the obligation to serve takes on a special and restricted meaning. In the case of the Poolco, the obligation is to accept and include as part of the dispatch pool all parties on a comparable basis, and to allow these parties to buy and sell power at locational marginal-cost prices in the pool. For the Gridco and the Lineco, the obligation is to connect customers and negotiate in good faith any requested enhancements of "wires" capacity with appropriate creation of new capacity rights collecting congestion rentals. The Gridco and Lineco can and must impose technical standards on new construction, and make the judgement with Poolco about the feasibility of new capacity rights. But the Gridco and Lineco are "obligated to serve" by meeting any feasible customer requests where customers agree to pay the regulated fixed charges to cover return of and on investments in the wires. The regulated segments have no obligation to review or be responsible for resulting investments or their subsequent economic consequences.

Security Concerns and Capacity Reserves

Given the history of vertical integration and centrally managed power dispatch, early proposals for market operations in electricity raised criticisms that unbundling threatened the security of the grid. If nothing else, the experience with the U. K. operation of a short-run pool with bidding and spot pricing has demonstrated convincingly that there is nothing inherent in the market structure which affects reliability. The essential requirement for a centrally coordinated dispatch remains, which protects the grid and ensures that the lights stay on, but this pool can operate to support a market with unbundled supply and demand without affecting reliability.

Secure economic dispatch coordinated through Poolco respects the current contingency limits on power flows and generation limits. In the short-run, the effect of security constraints in the system is to limit transmission and force out-of-merit generation which gives rise to short-run congestion costs. These short-run marginal costs are the true opportunity costs of "security," and in the competitive market customers face these opportunity costs. In the long-run, investment in the grid is undertaken when customers find it economic to reduce these congestion costs. In this sense, the evolution of the grid and its "reliability" is determined by the market. Since all demand is always met, at the short-run locational marginal cost, there is no non-price curtailment of demand and no need to build excess capacity to provide a reserve margin.[25] There may be high prices at times, and price will rise until demand can be met at marginal cost. But there is no separate or additional security or reliability problem. Hence, security in the short-run is maintained through the security constrained dispatch, and security in the long-run is priced and provided through the market for long-run investments to increase generation and transmission capacity.

With the elimination of a need for central planning to maintain capacity reserves, competitive market pricing and the economics of investments differ from traditional practice. Notably, there is no longer any regulated "capacity requirement" or the associated short-run

[25] This discussion simplifies from special cases such as loss of a direct connection line where no service can be provided and the price is, in effect, infinite. Obviously there is "curtailment" in this circumstance. However, in the competitive market there is never curtailment because of congestion where others are higher in the priority queue and the customer is "curtailed" even though the customer would prefer to receive service and pay the current market price. All customers who can be served and are willing to pay the market price will be served.

capacity payments. Transmission of capacity--a strange but common construct in the traditional world--disappears along with the capacity deficit and surplus payments. Traditionally, plants which would be too expensive to operate could continue to receive short-run capacity payments required under regulatory rules intended to maintain an adequate reserve margin. But with reserve margins no longer the responsibility of the central planner, electricity pricing in the short-run is limited to payments for energy only. Long-run investments in generation and transmission continue to be made, but solely in anticipation of profiting in the short-run market relative to the condition in the absence of the investment. Through the use of contracts between generators and customers, the risks can be shared. But in the end, the short-run opportunity costs are the only costs that can be avoided by investment, and these costs define the fundamental economics of investment decisions. Generation plants which are too expensive to run, even at true marginal cost, have no value.

Firm versus Interruptible Transmission and Generation

Firm versus interruptible service is another traditional institution that evolved of necessity in the world of average-cost pricing. When not all demand could be met, there was a value in distinguishing between "interruptible" uses that were of lower value and "firm" uses that were of higher value.   In the competitive market, this distinction vanishes. At the margin, all uses are equally firm or interruptible. If the customer or supplier is just indifferent at the margin, then a change in price leads naturally to some change in supply or demand. There is no obligation to serve, no capacity margin, and no distinction between different types of demand or supply.

The protection that came traditionally from "firm" rights reappears in the competitive market in long-run contracts for compensation for price differences. The compensation arises either for marginal generation costs or for transmission congestion. In the case of generation costs, the "firm" compensation contracts can be arranged independently through negotiations between buyers and sellers. In the case of transmission congestion, long-run contracts can be arranged through investment in the grid with congestion compensation administered by Poolco.

Opportunity Cost

Under traditional regulatory rules, it is difficult to define and utilize concepts such as opportunity cost and incremental cost. Especially in the case of transmission investments, where economies of scale and network interactions loom large, there may be no satisfactory definition of long-run variants of these concepts. At a minimum, application of these ideas, which is prominent in the EPAct and early FERC transmission pricing proposals,[26] appears to require analyses of supply and demand conditions over a horizon of years or decades, to evaluate alternative possible uses of the transmission grid and long-term power exchanges.[27] Given the complexities of network interactions, such analyses imply requirements for long-run forecasts of

---

[26] The Northeast Utilities (NU) principles allowing embedded or incremental cost pricing are set forth in FERC decisions: Northeast Utilities Service Company, Opinion No. 364-A, 58 FERC ¶ 61,070 (1992), *reh'g denied*, Opinion No. 364-B, 59 FERC ¶ 61,042 (1992), *order granting motion to vacate and dismissing request for rehearing*, 59 FERC ¶ 61,089 (1992), *affirmed in part and remanded in part sub nom*. Northeast Utilities Service Company v. FERC, Nos. 92-1165, *et al*. (1st Cir. May 19, 1993). The Penelec decision on the allowed "or" rather than "and" treatment of embedded cost and opportunity cost is in Pennsylvania Electric Company, 58 FERC ¶ 61,278 (1992), *reh'g denied and pricing policy clarified*, 60 FERC ¶ 61,034 (1992), *reh'g rejected*, 60 FERC ¶ 61,244 (1992), *appeal pending,* No. 92-1408 (D.C. Cir. filed Sept. 11, 1992).

[27] Section 723 of the EPAct authorizes FERC to collect and publish detailed information on transmission facilities. The FERC decision of September 29, 1993, approved a final rule on the information utilities must provide on their transmission systems. The new FERC Form 715 includes requirements for power flow studies and information on the incremental cost of adding units under system dispatch.

both the timing and the location of supply and demand for power.

Such forecasts may be prepared by someone, but the competitive market model moves this responsibility from the central planner and the regulator to the customer preparing to make an investment. The only "opportunity" cost arising with any form of regulatory supervision is in the Poolco calculation of the short-run locational marginal cost of power. And the regulator needs only supervise the process to guarantee the integrity of pool operations as they unfold and the availability of information for customers to analyze the market. There is no need for the regulator to prepare or endorse forecasts of future opportunity costs. All those risks are left to the customers to absorb in the market place. If the customer anticipates a market with high short-run opportunity costs, the customer can purchase existing capacity rights or invest to create new capacity rights that provide compensation for and protection from changing and volatile short-run opportunity costs. But the benefits and the risks remain with the customers; the regulator's only guarantee is that the customer will have access to the competitive market at the short-run market price.

In the case of transmission, the opportunity costs in many cases may be nothing more than the (small) cost of marginal losses. If there is excess capacity in the grid, either because of economies of scale or because of ex post realizations of load differing from ex ante forecasts, the opportunity cost of using the grid can be very low. Only in the case of congested systems, with the associated out-of-merit generation, will there be a significant opportunity cost for use of the transmission grid.

Deep versus Shallow Investments

Under the traditional regulatory schemes, transmission rights have been difficult to define, and the benefits are seen as widely dispersed. The central planner deciding on transmission investments has been unable to attribute the benefits of investments to particular customers or match the benefits against the costs. For those investments that are clearly attributable--the so-called "shallow" investments such as the final connection wire running from the grid to the customer--cost responsibility is obvious. But for "deep investments" embedded in the network the argument often has been that cost attribution is difficult or impossible, and everyone should share the costs as part of a network charge. Inevitably this practice confronts customers with obligations for investments they do not want for benefits they do not receive.

In part this difficulty with transmission investments arises from the inherent nature of scale economies and network interactions. But a large part of the difficulty arises from the inability traditionally to identify and assign the benefits of transmission investments. By contrast, in the competitive market model it is possible to identify the benefits of investment and to assign some or all of those benefits to the customers who agree to incur the costs. These transmission benefits appear in the form of lowered congestion rentals and the capacity-right contracts which protect these investments.

The problems of scale economies and network interactions remain, but they appear now largely as free ridership, with some customers receiving benefits that cannot be denied them once the investment is made, but for which they refuse to pay. The challenge is to find a consortium of potential beneficiaries of transmission investments who can be assigned enough of the benefits to agree to long-term contracts to cover the fixed costs.[28] With economies of scale, typically there will be total benefits strictly greater than total costs, so in principle it is possible to assemble a consortium to cover the investment costs without full participation by all the beneficiaries.

With the consortium agreeing to pay the investment costs, in exchange for certain long-run protection against short-run changes in congestion costs, the remaining customers are largely unaffected, at least in terms of their own long-run rights. The investment process does not impose new, unwanted fixed charges to pay for enhancements that serve others, and existing transmission capacity rights for compensation are unaffected by the new investment.

Of course, customers without long-term rights--those who are relying solely on the short-run market--may well be affected by the new investment. Typically the transmission investment will raise short-run prices near generation and lower prices far from the same generation. Generators and customers at these locations may win or lose. Exposure to the risks comes with the decision not to purchase long-run protection. Later as demand increases, congestion costs will rise, raising the short-run market price. Those who make the choice to rely on the short-run market have no recourse to regulators to protect them from the business risk, or to stop investment in the generation or transmission that may affect market prices.

---

[28] The economic theory of club decision making addresses the process of joint decision and allocation of benefits within a group. Buchanan, J. M., "An Economic Theory of Clubs," Economica Vol. 32, February 1965, pp. 1-14. S. C. Littlechild, "Common Costs, Fixed Charges, Clubs, and Games," Review of Economic Studies, Vol. 42, No. 1, January 1975, pp. 117-124. For applications in the context of telephone networks, see G. R. Faulhaber and S. B. Levinson, "Subsidy-Free Prices and Anonymous Equity," The American Economic Review, Vol. 71, No. 5, December 1981, pp. 1083-1091. For a discussion in the context of power pools but without explicit consideration of networks, see S. R. Herriott, "A Long-Run Cost Allocation Problem in the Political Economy of Electric Utility Power Pools," Journal of Regulatory Economics, Vol. 1, No. 1, March 1989, pp. 69-86. Since transmission is not as important as generation, generation economics may swamp transmission costs and make agreement easier within the transmission investment club.

Protection of Native Load

Without a clear definition of the benefits in a transmission network, it is not clear how to defend previously implicit rights. Without a clear definition of opportunity cost, it is not clear how to compensate the grid owners for their opportunity costs. When the grid is bundled with generation, and third-parties cannot impose costs on the owners, the problems are internalized in the traditional central planning decision. In the open access market, however, with the introduction of third-parties who will utilize some of the benefits of the transmission grid, defining and measuring the appropriate transmission compensation is essential.

This is the essence of the response to the charge to protect native load as defined under the EPAct.[29] In the competitive market model, protection of native load reduces to defining the long-run capacity rights and then assigning those rights to the native load to match the assignment of responsibility for the fixed costs of transmission. In the short-run, all users of the transmission grid will pay the short-run opportunity cost, and holders of long-run capacity rights will receive the benefit of congestion rentals. If all the fixed charges for transmission are in the rate base of native load customers, then all the long-run capacity rights should be held for the benefit of the same native load customers.

Transmission capacity rights for compensation provide the protection for native load customers. However, everyone still has the right to use the transmission grid in the short-run and pay short-run marginal costs. Protection of native load does not prevent access to and use of the grid. Assignment of capacity rights to native load customers only preserves the economic

---

[29] Section 722 of the EPAct defines transmission services and pricing principles including the provision that "costs incurred in providing the wholesale transmission services, and properly allocable to the transmission services, are recovered from the applicant for such order and not from a transmitting utility's existing wholesale, retail, and transmission customers."

interests of the customers who are paying the fixed charges of the transmission grid.

Contract Path and Wheeling

The contract path fiction is not required by the traditional regulatory regime, but it is relatively simple and easy to implement. However, the contract path model is at odds with the competitive market model under open access and marginal-cost pricing. Fortunately, the contract path notion is not required in order to define transmission rights over the long run. Transmission capacity contracts for compensation do not require a contract path, and are defined only in terms of the differences in congestion rentals between two locations. There is no necessity to follow the path of power flow. All the many short-run network interactions are contained in the calculation of locational marginal costs.

Similarly, compensation for wheeling occurs automatically in the transactions at locational short-run marginal-cost prices. There is no need to identify the intervening utilities or the wheeling entities. Everyone connected to the grid is affected by every transmission action, and everyone pays the full short-run cost through the differences in locational prices. Those using the grid pay only the short-run costs, with embedded investment costs paid by those with the rights to collect the congestion rentals. The short-run marginal cost incorporates all the opportunity cost to be recouped in the competitive market. Pricing for flows on parallel paths is thus automatic, with no necessity to assign specific flows as passing through specific intervening utilities. The contract path is replaced by the contract network.[30]

[30] With charges for use of the network are based solely on short-run marginal costs, there is no need for ad hoc adaptations of the contract path such as the "megawatt-mile" proposals as in A. F. Mistr and E, Munsey, "It's Time for Fundamental Reform of Transmission Pricing," Public Utilities Fortnightly, July 1, 199, pp. 13-16.

Just as the contract path is only an approximation of the flow in a network, transfer capacities measured at "interfaces" are only an approximation of the constraints in a network. In reality, multiple contingencies, voltage limits, thermal limits and dynamic stability conditions interact to constrain the feasible transactions. Many combinations of load, and therefore many capacity right configurations, are possible. But once established as feasible, these capacity rights can be described in terms of the right to collect the difference in congestion costs at two different locations. These rights are compatible with operation of the market and pricing based on short-run locational marginal costs.

Secondary Markets

Availability of a secondary market is essential in maintaining price efficiency for long-lived assets under changing conditions. The existence of such a secondary market, with the freedom to trade, is usually sufficient to guarantee efficient prices. The initial allocation of asset rights is important as a matter of the distribution of wealth, but once allocated the market will operate to determine the opportunity value of assets.

In the traditional regulatory market, secondary trading is typically cumbersome or in many cases prohibited. Under average-cost-based regulation, the regulators have often felt obligated to prevent windfalls in a secondary market, with the expected result of an inefficient allocation of assets, construction of new facilities when barred access to excess old facilities, and so on. Introduction of property rights and the ability to trade in the market is a principal mechanism for movement away from a regulated market into an efficient competitive environment.

In the competitive market model, there is an active secondary market. The principal long-run assets that can be traded are the unregulated facilities and contracts for compensation. Facilities such as generating plants can be bought and sold in the market for whatever price can be negotiated between willing buyer and seller. The same applies to trade in any long-run contracts for price differences that have developed between generators and customers. These trades take place without regulatory oversight, or even knowledge, as the market responds to changing conditions and customer preferences.

The only secondary market that requires central participation, if not regulatory oversight, is the market for trades in transmission capacity rights. Although in principle this too could evolve as an arms-length market without participation by Gridco or Poolco, other information economies may require participation by the transmission entities. In particular, the physical "terms-of-trade" between two transmission capacity rights are far from obvious. It may well be that 100 MW between Pittsburgh and Baltimore can be exchanged for 200 MW between Pittsburgh and New York, or 150 MW between Baltimore and New York. The network interactions are nonlinear, and only the central dispatch operator has the information required to know which exchanges are feasible.

The value of the rights would differ even more depending on the forecasts of short-run prices and congestion rentals, but this is a complication that is left to the market participants. However, even with this risk imposed on the right owners, the Poolco and Gridco may have to assist in operating a secondary market for transmission rights just to identify and verify which trades are possible. The right-holders can then negotiate to identify which trades are economic. Poolco and Gridco would allow any trades that were feasible, as long as the parties could agree.

Two-Part Tariffs

The competitive market model outline here relies on pricing according to two-part tariffs including a fixed and variable charge. This approach contrasts with the alternative of a one-part tariff with a single price based on usage. The importance of this distinction arises because of the large capital costs throughout the industry coupled with uncertainty or the assumed economies of scale in transmission.

In the absence of capital costs, a single short-run price for usage would suffice to maintain the efficiency of the competitive market solution. Obviously, if there are no capital costs, then there are only short-run costs and a short-run product, and the only price would be the short-run price which would cover all the costs. Even with capital costs and no uncertainty or no economies of scale, a single short-run price would provide and efficient solution in equilibrium. Capital investments would be made just to the point that balanced short-run and long-run marginal costs and recouped the marginal capital investment. In this world, a single price derived in a short-run market is optimal in the long-run, guaranteeing both efficiency and revenue recovery.[31]

With uncertainty and capital costs, it may be that actual demand, as realized after making a capital investment, might be higher or lower than anticipated. Hence the short-run efficient price would be too high or too low to cover the capital costs of the investment. On average, the expected short-run price is adequate, but it would be only luck if the actual price

---

[31] For example, as Grant Read has noted, constant-returns to scale in transmission is an important and implicit assumption in the real-time pricing model outlined in F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R.E. Bohn, Spot Pricing of Electricity, Kluwer Academic Publishers, Norwell, MA, 1988.

turned out to be the expected price. Or, if the short-run market did not operate, and there was only a single long-run price, then there would be either too little use of the capital facility, when demand was low, or there would be some administrative curtailment, when demand was high. Such traditional regulation is inefficient, and investors in large capital facilities like electricity generating plants are often unwilling to take the risk associated with full reliance on the short-run market. Without regulation to enforce a uniform long-run price, the natural alternative is for long-term contracts that provide a assurance of recovering fixed costs. Although these contracts can have many forms, when viewed in conjunction with a secondary market for short-run power such contracts can be converted into an equivalent fixed and variable cost format. The contract allows efficient long-run investments and efficient use of the facilities in the short-run. This form of protection against uncertainty in the short-run is a standard rationale for the use of long-run contracts and effective two-part tariffs.

With economies of scale, which are assumed to apply to the wires business only, reliance on the short-run market alone becomes even more problematic. In the presence of economies of scale, by definition the ex post short-run marginal costs for an optimal investment can be too low to recover the costs of the investment. With economies of scale, it is not possible to guarantee that the last small increment of investment just balances the last small change in short-run marginal costs for the simple reason that the investment does not come in small increments. With optimal investment made in large chunks, to exploit the economies of scale, the proper criterion is to compare the short-run costs with and without the investment. If the difference is great enough, the investment should take place, but then the ex post short-run marginal cost will not recover the cost of the investment.

An extreme example--an investment with infinite capacity that yields zero short-run opportunity cost, perhaps a bridge--illustrates the point. If the initial short-run costs are large, then the investment may be desirable. It is better to build the bridge than to stay on one side of the river. But after the investment is made there is no way to pay for the capital costs by collecting the short-run efficient price of zero! If the bridge has excess capacity, it is best not to charge a toll for its use. This extreme case also suggests the obvious answer. If the investment is worthwhile, those who benefit from the investment should agree in advance to pay the fixed charge of the investment and then use the facility without restraint. In the case of public facilities like with the bridge problem, we often use taxes as the source of the fixed charge payment. In a market, this would be a two-part tariff with a fixed charge under a long-run contract and a variable charge for use that happens to be set at zero.

This application of two-part tariffs is especially apt in the case of transmission. For example, in its open access regulations for natural gas pipelines under Order 636, FERC has employed a fully fixed charge for capacity rights and a variable cost for actual use that includes no contribution to capital.[32] The case is even more compelling for electric power transmission where it is more difficult to produce an acceptable definition of use that is tied to any subset of the network. Any attempt to recover all or some of the fixed costs of the network based on variable charges confronts the problem of defining who is using which facility, and why. In a free-flowing electrical transmission grid, power may flow in ways that are beyond the control or intent of the user. And if as can be the case, the short-run opportunity costs are significantly different than some pro-rata allocation of the long-run capital costs, a one-part tariff that averages

[32] FERC, [18 CFR Part 284] Order No. 636, Washington, DC, April 8, 1992.

fixed costs along with variable costs can produce inefficient incentives and contentious debate over complex and changing allocations. However, in the competitive market model outlined here these problems for transmission are handled through the effective application of a two-part tariff. Short-run variable costs of transmission use are paid through the automatic pricing of the pool operations. Long-run fixed charges are agreed to under contract as part of the process of deciding on new investments in the grid.

Network Transmission Service

New unbundled users of the transmission grid have sought a flexible form of access that is often referred to as network service. "While there is no universally accepted definition of network service,"[33] the basic idea is to allow use of the grid without restriction to particular point-to-point transactions or without paying additional charges for each change in point of receipt or point of delivery. In the case of no transmission congestion, this form of network service is provided automatically by the competitive market supported by the Poolco dispatch. All users are allowed to change their sales and purchases of power, and hence their implicit transmission. Prices for power at different locations reflect only the relatively modest impact of differential losses and--with no congestion--the market price for long-run capacity-rights-to-collect-congestion-rentals should be at or close to zero. Hence, without transmission congestion everyone has a well-defined network service with no materially different charges depending on location.

---

[33] Federal Energy Regulatory Commission, "Transmission Pricing Issues," Staff Discussion Paper, Inquiry Concerning the Commission's Pricing Policy for Transmission Services Provided by Public Utilities Under the Federal Power Act, Washington, DC, June 1993, p. 13.

In the face of transmission congestion, however, the situation could be different. If "network service" means that flexibility of points of delivery and receipt comes without constraint or cost, then it is clear that not everyone can have network service. With transmission congestion, by definition, some users must curtail their demand or pay more to respect and reflect the transmission constraints. The pool dispatchers would not be able to guarantee economic dispatch and meet all "network service" requests without violating the short-run constraints that limit the total system. Assuming that network service should not be imposed at the expense of giving up the benefits of economic dispatch, the need is for a definition that is economically sensible even in the face of transmission congestion. From this perspective, the natural specification of network service is to allow complete freedom to use the grid as long as the user pays the short-run opportunity cost of both losses and congestion. However, there would be no additional charges or limitations on use of the grid.

This form of network service is provided to everyone without discrimination under the operation of the competitive market model. All users have the opportunity to bid at different locations, and the implicit short-run price of transmission reflects only the short-run opportunity costs. There is no other limitation and no additional charge to any user, and all users are treated the same. In the short-run with constraints, transmission pricing will produce congestion rentals that will be paid in turn to the holders of long-run capacity rights. These transmission capacity rights will be available in the market, and all users can purchase different combinations of rights or invest in transmission enhancements in order to increase capacity and the available rights. The price hedge defined in the rights is primarily for point-to-point transactions. With the grid providing these perfect hedges from point-to-point, a market can develop for different portfolios of contracts which should allow customers to construct alternative hedges at a competitive market price. In the aggregate, however, not everyone can have a perfect hedge against all price changes, and someone must bear the risk of short-run changes in congestion rentals. If network service incorporates these short-run risks, it is embodied in the competitive market model. If not, then someone else must bear the cost of congestion, and not everyone can have network service.

Ex Ante versus Ex Post Pricing

In a fully real-time market sellers and buyers would interact through multilateral negotiations with the pool dispatcher, changing offers for supply and demand in response to changing prices. This continuous interaction is necessary if trades take place at the bid price. Such real-time pricing might be difficult to implement in an electricity market, although it remains a subject of active research and experimentation.[34] The alternative approach is to interact through the pool with bids and pricing based on the equilibrium price of the efficient dispatch. The price paid would not be the bid price; rather all transactions would take place based on the actual dispatch at the equilibrium price. In principle, the bids could be current up to the moment of dispatch, and the locational prices determined would be the same as real-time prices. Since the bid would not affect the price paid or received for most participants, there would be no necessity to change the bid continuously in response to changing market conditions. All that would be necessary would be to change bids when circumstances changed the minimum or

[34] Bahman Daryanian, Roger E. Bohn, Richard D. Tabors, "An Experiment in Real Time Pricing for Control of Electric Thermal Storage Systems," IEEE Transactions on Power Systems, Vol.6, No. 4, November 1991, pp. 1356-1365.

maximum price that would be accepted.

In practice, for administrative convenience, it is necessary to fix the bids at some point, say a day ahead as is done in the U. K., even though the actual conditions of availability of plants and other factors may change over the ensuing day. Hence the debate arises whether to base the transaction prices on the ex post conditions of the actual equilibrium dispatch or on a hypothetical dispatch based on the ex ante conditions of some version of the day ahead bids. For example, the U.K. system determines the market price using the ex ante unconstrained equilibrium, but then of necessity dispatches the system to reflect the constraints. The differences in costs are averaged across all users in the "uplift" which is added to the market price.[35] A common argument is that the ex ante prices should apply so that customers know how to respond to market conditions. However, this not necessary for dispatchable generators or load, which by definition can be controlled in real time by the pool dispatcher. And for the load which has a longer lead time, the decisions a day ahead can be made on a forecast of the prices.

The principal problem with making payments based on actual ex post load and hypothetical ex ante prices is the creation of incentives for gaming the pricing system. For instance, a generator may declare a plant unavailable, to increase the estimated day ahead price at a location, and then make the plant available after the prices are set, while demanding the ex ante price for the actual volume generated.[36] An ex post pricing system is not subject to these

---

[35] The U. K. "uplift" consists principally of payments to generators to account for "operational outturn" arising because of constraints in the transmission system. This average cost added to the short-run marginal price has been increasing at a trend of 2.2% per month since the launch of operation of the short-run market. U. K. Office of Electricity Regulation (OFFER), "Pool Price Statement," July 1993, p. 37.

[36] This was the experience under the particular rules of the U.K. system, which uses ex ante prices but then pays for ex post quantities. U. K. Office of Electricity Regulation (OFFER), "Pool Price Statement," July 1993.

inefficient incentives. To the extent that an ex ante price is deemed necessary, the preferred approach would be to set day ahead prices and quantities as short term contracts, with a settlement for differences based on the actual dispatch quantities and prices.[37]

For both reasons of simplicity and to avoid certain counterproductive incentives, therefore, the pricing system settlements process should be based finally on the ex post prices and ex post quantities of the actual dispatch. With a bidding system run through Poolco, this settlements process would be easy to implement. Such charging mechanisms are already in place with the split-savings systems of tight power pools which are based on actual costs of generation calculated ex post using actual quantities. The only difference would be conversion from the split savings implicit prices, based on average costs and hypothetical re-dispatches of the system, to market based marginal costs at each location including the effect of congestion and losses. Once the conceptual hurdle is leaped, the competitive market settlements system would require a minimum of changes from traditional operations.

Information Requirements

Both the traditional and competitive markets require vast amounts of information to operate successfully. Control of dispatch, for example, requires real time monitoring and quick responses. The difference between the highly regulated and the more competitive market is in how much of the information necessarily passes through central hands and is monitored by regulatory bodies. For example, if long-term contracts between customers and generators are

[37] L. E. Ruff, "Competitive Electricity Markets: Economic Logic and Practical Implementation," International Association for Energy Economics, 15th Annual International Conference, Tours, France, May 1992 (Revised June 1992).

deemed to require regulatory oversight and specific performance, then it will be necessary for dispatchers to be informed of these contracts and to modify the generation dispatch accordingly. However, if the competitive market is operated in the short-run based only on short-run marginal costs, the dispatcher need know nothing of the contracts. And if the generation market is competitive with open access to the essential pool and grid facilities, as in the competitive market model, the regulators need not monitor or approve the long-run generation contracts.

In the short-run for the competitive market model, the only contracts that require central information are the transmission capacity rights to collect congestion rentals. Although Poolco must be aware of and handle the distributions under these capacity rights, the contracts have no effect on the dispatch itself and can be handled entirely through an essentially transparent accounting process as part of the ex post settlements system.

For long-run investments in transmission, there is again a difference in what will require central control and participation. However, the requirements for the central operators and the regulators are substantially less in the competitive market model. In the case of traditional regulation, it is necessary for a utility to decide on expansion, and for the regulator to certify the need for the investment, that will later be recovered from the ultimate customer. Typically there is little participation by the customer in this process other than as an intervenor in regulatory proceedings. Hence the regulator and the utility bear the responsibility for analyzing the highly uncertain prospects of the prospective transmission investment for which the customer will later be obliged to pay. And with no well-defined assignment of the potential transmission benefits, it is difficult to match the costs with the benefits. This system necessitates very detailed studies across many uncertain conditions, with huge associated information requirements, and even then little can be guaranteed.[38]

By contrast, these information requirements are substantially reduced in the competitive market model. Here the question that must be addressed is not the full range of capabilities of the transmission system under a wide variety of generation and load conditions, which is a complicated challenge. Rather the issue reduces to an evaluation of the feasibility of the particular allocation of transmission capacity rights under the loads induced by the actual capacity rights, and no others. In the subsequent operation of the short-run market, changes in load and generation will affect the congestions rentals, but not the capacity rights. This is an important simplification which should greatly reduce the regulatory burden. Customers decide on the capacity rights that they want to obtain, and negotiate with Gridco to provide the least-cost expansion of the grid needed to create those rights. The regulators need supervise the process only to avoid the problem of gold plating the transmission grid. Otherwise Gridco has no incentives for or against any particular investment. Regulators need not certify need. And Poolco need only certify that the resulting expansion of capacity rights is feasible and compatible with the existing rights, which is a relatively straightforward process that involves testing the load flow under a single set of conditions.

Of course, all the uncertainties about actual transmission flows and capacities dependent on loads will remain with the competitive market model, but this uncertainty is translated from uncertainty about the magnitude of the transmission right to uncertainty about the magnitude of the congestion rentals. Furthermore, this short-run price uncertainty moves onto the shoulders of the transmission customers--including generators, brokers, consumers, and all

[38] Apparently such a system operating with regulatory oversight is anticipated by FERC as summarized in the September 29, 1993, decision on the required transmission information filing.

other possible participants in the market--who either accept the risk or agree in advance to pay for transmission capacity rights that protect them from the risk. The costs are now well defined and the transmission rights are clearly defined and clearly assigned. The uncertainty is in the valuation of the transmission rights, for which the customer may want to do many of its own analyses to estimate the value of expanded capacity. But the customer is required to pay for investment only in the case where the customer sees the value in expanded capacity, not because someone else has decided act as the customer's unwanted agent to expand the transmission system. As in any other market, the participants make their own decisions to pay for certainty or take the risks.

ISSUES UNRESOLVED

The competitive market model accommodates many of the features of electricity production and use within a framework that replaces central planning with market choices. However, in principle, there remain areas where the suggested market structure may not support the most efficient outcome. Alternative institutions or public regulation may be indicated.

Free Riders

The assumed economies of scale and network interactions in the wires businesses can create circumstances where transmission investment is not economically optimal. Inevitably, incremental investments in transmission can change marginal costs, and these changes affect all users of the system. The total net benefits could be positive, and greater than the cost of investment. However, if the result is a permanent lowering of marginal costs for many users, there is an incentive for some to "free-ride" on the investments of others. Once the investment is made, others capture some of the benefits of lower short-run prices whether or not they have contributed to the investment.

Besides the equity problem and the perception of fairness, the potential for free-ridership creates potential efficiency problems similar to the prisoner's dilemma. If the investment is efficient, and everyone could agree to share the costs, the investment would be made. But because of the incentive to free-ride, it may be difficult to forge and enforce agreement among a large enough coalition to permit the investment to go forward, and an efficient investment could be foregone.

The efficiency problem is not only one of under-investment in transmission. With the change in short-run prices, the transfer of congestion rents induced by an incremental investment might be enough to motivate the winners to expand transmission, even though the economic value net of the transfers is less than the investment cost.[39]

The importance of these problems is an empirical matter that deserves investigation if it is viewed as a serious complication. By assumption, the absence of economies of scale and the relatively small size of incremental investment in other sectors of the market, such as generation, limit the problems of free-ridership and inframarginal change. These are concerns only in the transmission sector. However, most investment in the transmission system would arise as a result of new investment in generation and the accompanying need to obtain transmission capacity rights that would complement generation contracts. Since the cost of the

---

[39] N. G. Mankiw and M. Whinston, "Free Entry and Social Inefficiency," Rand Journal of Economics, Vol. 17, Spring, 1986.

generation for each such deal is likely to be much larger than the cost of transmission, there is a natural pressure to force agreement on a reasonable allocation of the transmission cost in order to complete the generation deals. If these conditions prevail, the transmission tail is unlikely to wag the generation dog.

Furthermore, the simplistic application of the free-rider argument ignores the dynamics of club decision-making where repeated decisions are required. If the next transmission investment will be the last, then the incentives to game the system may be pronounced. But if enough of the participants see the process as a repeated interaction among the players, there may be a tendency to arrive at an equilibrium that would be impossible to enforce without the incentives of anticipated future deals.[40]

These appeals to reason among the current participants in the market would carry little force when applied to future users. For example, consider the case of a new line built to a new industrial park. The owners of the first factory may be willing to pay for the line, and it would have a great deal of excess capacity and generate no congestion rentals. However, when the second factory connects to the same line, and there are still no congestion rentals, only the economists would be content with the solution that allocates all of the cost to the first investor. It may not be economically inefficient, but it will be unpopular.

In these cases and others, there may be some necessity for regulatory oversight or other mechanisms to address the most egregious problems where economies of scale loom large and market choices deviate far from an efficient solution. The challenge will be to define these procedures without sacrificing too much of the principle that new investments should not be

---

[40] Drew Fudenberg and Jean Tirole, Game Theory, Cambridge, MIT Press, 1992, Chapter 5, "Repeated Games," pp. 145-203.

made unless the customer who benefits is willing to pay.[41] This fundamental market discipline of consumer sovereignty is essential to preserve efficient investment and to replace reliance on ex post prudence procedures with ex ante long-term contracts.

Market Power

Given the traditional view of the electricity market as a natural monopoly with vertical integration, any reforms of the electric industry must address the issue of market power. The FERC in its policies has been concerned with mitigation of market power as an essential prerequisite for any participation in an unregulated market.[42] Without the conditions for workable competition, and in the absence of regulation, the electricity market would be subject to manipulation of prices and quantities by participants who possessed market power.

By assumption, the Poolco, Gridco, and Lineco segments have market power, and there is no obvious way to mitigate this power or operate in a decentralized, competitive environment. Hence the electricity market model assumes the need for separation, open access rules, and regulatory oversight, and the regulation is designed to promote the operation of competition in the other segments. These other segments are by assumption workably competitive, through a combination of enough participants, contracts with customers, ease of entry, and contestability. The issue, therefore, is the degree to which the assumption is wrong and market power might

---

[41] L. Ruff has suggested the phrase "extra-competitive" costs to refer to costs of efficient investments that, due to economies of scale, exceed the ex post market value of the investment. These costs might be allocated to other participants deemed to receive benefits, with others participating through voluntary long-run contracts. The rule might be that the Gridco could impose a fraction of these fixed charges, along with a corresponding share of the new capacity rights, as long as a majority of the costs were covered under voluntary contracts.

[42] The NU transmission pricing principles are motivated explicitly by the FERC goal of mitigating market power.

exist in other than the wires and dispatch functions, and how the competitive market model would fare in the presence of market power.

Although a complete analysis of market power issues would require a factual assessment, the most interesting case is in the generation market, where the stakes are high and ownership patterns might lead to a concentration of control that could be used to affect the otherwise economic Poolco dispatch. For example, this has been a problem of concern with the generation duopoly created in the U. K. privatization.[43]

A key feature of the competitive market model is reliance on locational spot prices for all short-run purchases and sales of power. The prices differ by location, and in certain instances, transmission constraints may force certain expensive generators to operate, and thereby create large differences in prices. The existence of high prices, with large differences across regions, is not in itself evidence of market power. In the face of transmission congestion, there are true opportunity costs which lead to very high values for generation located in particular locations. The associated short-term profits or "scarcity rents" are a normal part of the operation of an efficient competitive market, and they play an important role in providing the proper incentives in the market. In some cases, the high prices reflect the value of generation and provide the signal for reduced demand. And over the longer-run these short-run profits provide the incentive for new entry that is needed to keep the market efficient.

Existence of high prices or profits becomes a matter of policy concern only when market power leads to short-run or long-run monopoly rents that arise with self-interested

---

[43] S. Littlechild, "Competition, Monopoly and Regulation in the Electricity Industry," U. K. Office of Electricity Regulation, June 1993. U. K. Office of Electricity Regulation, "Pool Price Statement," Birmingham, England, July 1993.

limitations on supply to maintain prices that diverge from marginal costs. For a generator that is in a region that must be operated to respect transmission constraints--the so called "constrained on" generators--the generator may know its plant is protected from competitive entry and its plant will be the marginal source of supply over a wide range of prices, in violation of the competitive market assumption. In this circumstance, the bidding model operated by Poolco produces the result that the constrained-on generator is paid whatever it bids, and there is an incentive to manipulate the bid price and quantity to extract monopoly rents.

In the extreme theoretical case, with no local demand response, the generator could extract any price. More realistically, with dispatchable demand, the generator is a local monopoly and can raise prices and lower supply according to the classic monopoly model, continuing to raise prices and reduce demand as long as it is profitable, and produce an inefficiency in the local level of price and production. If such circumstances arise, there is a policy problem, and the competitive market model with Poolco dispatch does not inherently eliminate the detrimental effects of market power. However, neither does the model create the market power, and these market power distortions would need to be addressed no matter what market approach is adopted. Hence, the relevant comparison is the competitive model versus other alternatives.

One approach, as taken in the U. K., is to ignore locational differences in prices and charge only a uniform national price. This system is clearly inferior to the competitive market model in constraining the effects of market power. In this circumstance of a single national price, it is by definition still necessary to dispatch the "constrained on" plant, but the increased cost is averaged across all customers. The effect of the national average price model is to reduce the demand elasticity compared to the competitive market model where all the local increase in price would be seen in the local incentive for reduced demand. With the national average price diluting this incentive in the U. K. case, the demand elasticity is reduced and there is more room for the generator to bid a higher price for the constrained-on plant.

A simple alternative might be to charge the "marginal" price nationally, rather than the average price. While this would eliminate the problem of the reduced demand response, it would force inefficient reductions in demand everywhere. Furthermore, it would create additional profits for the other generating units, and to the extent that these units were owned by the same entity as owned the "constrained on" plant, there would be an added incentive for distortions in bids and the resulting prices.

Compared to these other market models, therefore, the competitive market model with locational pricing at least isolates and limits the inefficient incentives for manipulation of bids. And to the extent that the constraints are uncertain or changing, the interaction through the grid could greatly reduce the generator’s ability to be sure that a higher bid would not cause Poolco to redistribute generation patterns to avoid the need for this expensive plant.

From this perspective, the residual problems of market power that remain seem best viewed as the same as in other markets. There is no advantage in foregoing the Poolco dispatch or ignoring the real constraints that may exist in the transmission grid. To the extent that market power abuses remain, there may be no alternative to traditional regulatory approaches, but regulatory approaches applied within the natural structure of the competitive market model.

The symptoms of manipulation by generators would appear in a comparison between dispatch and running costs of a plant. In many instances, at least for nuclear and fossil plants, the energy cost of a plant is a good proxy for its short-run marginal cost. In the competitive market model, whenever the local price equals or exceeds the marginal cost, it should be running, except when the plant is removed for maintenance or for some other identifiable reason. Hence, questions of market power and bid manipulation would be raised for any plant which exhibited a consistent pattern of not running or running well below capacity in hours when the local price exceeded a reasonable estimate of its short-run marginal cost. In these cases, the form of regulatory intervention might be to limit market power by imposing conditions on the bid prices for plants that enjoy market power, but leaving the rest of the market to operate as a competitive bidding process into the Poolco dispatch.

Social Ratemaking and Externalities

Average-cost pricing and traditional regulation have included many practices geared to pursuing social objectives that are not obtained in decentralized operation of the market. The use of average-cost prices in itself creates a social problem whenever there is a significant deviation from marginal cost and its associated economically efficient use of resources. In addition, by definition the decentralized market does not recognize the impacts of environmental and other externalities that can reduce overall social welfare. Traditional regulation of the vertically integrated monopoly provides the opportunity to support programs or change prices that redirect activity towards the social objectives as interpreted by the regulators.

The competitive market model naturally eliminates a few of these problems. Reliance on marginal-cost pricing restores the equation between the costs of production and the private benefits of the use of electricity. Customers have the opportunity to purchase either more electricity or to change their demand for electricity through conservation and load management. Hence, the competitive market model removes one of the causes of distortion that has led to wide-scale use of demand-side management programs to provide assistance or subsidies to customers to act in a way that otherwise would be consistent with their own interest given the true cost of producing versus saving electricity.

Use of average- rather than marginal-cost prices is only one form of market failure, meaning a deviation from the results of the pure theory of a competitive market. Movement to marginal-cost pricing helps eliminate one of the most obvious complaints about the traditional system, but others will remain as a matter of principle. For example, to the extent that capital markets are imperfect and consumers cannot obtain credit for conservation investments, there may be another form of incentives for private decisions to produce or use electricity that deviate from the ideal.[44] To the extent that broader environmental standards do not capture the damage caused by production and use of electricity, there will be "externalities" which by definition are not included in the price of electricity and therefore not accounted for in the decentralized decisions of the competitive market model. These externalities can be internalized through regulatory intervention in the choices of production and use. The competitive electricity market model itself does not internalize these environmental impacts, but it is compatible with the use of market-based methods for internalizing externalities such as environmental taxes and marketable

[44] Even the existence of these alternative forms of market failure is subject to debate. For example, apparent high discount rates, so often cited as an example of the failure of consumers to make efficient conservation investments, can be explained as an optimal response to uncertainty and technological change. See A. Jaffe and R. Stavins, "The Energy Paradox and the Diffusion of Conservation Technology," Kennedy School of Government, Harvard University, July, 1993.

emissions permits.[45]

In addition to this standard argument about externalities and market failures, it has always been recognized that the market outcome and its social value depends on the distribution of wealth. For instance, even without externalities it has always been true that society has been concerned and often unwilling to accept the implications of curtailing supplies to low income customers who have not paid their electricity bills. Or social goals of universal minimum service have depended on subsidies from one class of customers to another, say higher prices from industrial customers in order to provide lower prices for residential customers. Attention to these realities may be far more common than the relatively esoteric worries of economists focussed on efficiency.

The existence of traditional rate regulators and regulation provided the venue for addressing these "market failures." Not all the problems of deviation from the socially optimal result can be solved by the competitive market model. The competitive market is best at matching costs and benefits, stimulating innovation, allocating risks to those willing to bear them, and so on. But these benefits accrue only for those costs that are recognized in transactions in the market, and only for the existing allocation of wealth. Redistributing wealth and capturing the effects of externalities will continue to be treated through some form of public oversight and regulation.

To the extent that electricity regulators carry the burden of influencing the market outcome to address these remaining problems with the market, the target of the regulation must

---

[45] For a discussion of market-based methods of environmental control, see R. Stavins and T. Grumbly, "The Greening of the Market: Making the Polluter Pay,"chapter 9 in Mandate for Change, Washington, DC, 1992, pp. 197-215.

be different than in the traditional model. Just as the introduction of competition involves unbundling products and services, so too must rate regulation unbundle and redirect its focus. The simple principle is that costs which are not consistent with the transactions in a competitive market-- a market with free entry and customer choice--can only be collected in those segments of the market that operate as a franchise monopoly. If the customer can choose between one supplier who pays the regulatory "tax" and another who does not, the choice will be clear. If the tax is to be collected, the buyer must have either only one choice or face the tax on every alternative.

In the traditional market, the fully bundled electricity product was provided in a franchise monopoly, and added costs imposed by regulation were bundled as part of the overall cost of the product. In the competitive market model, the only franchise monopolies are in the wires and dispatch activities of the Gridcos, Linecos and Poolco. Hence these will be the only places where regulators operate to set prices or define access rules. Now the added costs imposed to achieve social objectives can be added only for services that amount to a fraction of the price of the overall service. Although costs vary from company to company, on average generation, transmission, and distribution account for 77%, 7%, and 16% of the cost of delivered electricity.[46]

If local regulators have control only over the Lineco franchise and are interested in correcting the failures of the market, only the Lineco charges offer the opportunity for effective intervention. The impact of any such internalization on Lineco costs could be pronounced.

---

[46] Based on percentage allocation of variable expenses according to FERC accounts, and share of electric utility plant for fixed costs. Other and general administrative expenses allocated pro rata to generation, transmission and distribution. Department of Energy, Financial Statistics of Selected Investor Owned Utilities 1989, Energy Information Administration, Washington DC, January 1991.

Presumably the Lineco charges will be only a portion of the traditional distribution charges. If the increased regulatory cost is expressed solely as an added charge on the franchised monopoly Lineco service, which is only 16% of the cost of the delivered electricity, the magnitude of the tax will be increased by more than a factor of six relative to the share in the bundled service, and thereby made far more visible. This will not simplify the life of a regulator.

One further implication is that in the transition moving from traditional regulation to the competitive market model, it will be important to avoid long-term commitments to charge lower embedded cost rates for Lineco type services, and then open access to all other segments of the market that give customers a choice of suppliers of energy and related services. If the Lineco service is the only mechanism for a local regulator to collect a tax to support an "above market" cost, then the commitment to low embedded cost rates eliminates the only mechanism for addressing market failures. This could be a critical issue in the transition to open access to the wires, where attention to this narrow part of the problem without a vision of the operation of the market as a whole could produce an early set of decisions that create unwanted constraints and incentives that preclude later necessary steps in a successful transition to the competitive market model. This failure of internal coherence of the reforms has been cited by leaders in the natural gas industry as the most significant problem in the expensive transition in the open access regime of the U. S. natural gas pipeline industry,[47] or in the problems that developed after the launch of the U. K. competitive electricity market model.[48]

---

[47] J. Skilling, Enron Corporation, Presentation to Edison Electric Institute, Finance Committee Annual Meeting, New York, May 1993. For a discussion of the transition costs, see R. Pierce "Gas Industry Transition Costs and Their Allocation," Draft, September 1993.

[48] S. Littlechild, "Competition, Monopoly and Regulation in the Electricity Industry," U. K. Office of Electricity Regulation, June 1993.

TRANSITIONS

Unlike Alice, those interested in pursuing the electricity market changes mandated by the EPAct should know where they "want to get to." Therefore, as the Cheshire Cat would advise, the path to follow in the transition is importantly influenced by the destination. If the competitive market model defines the long-run target for the organization of the electricity market, then this goal should guide many of the choices that must be made during the transition. Describing and understanding the transition problems is a major challenge. The most important immediate decisions include issues of speed, direction, and who pays for the trip.

Big Bang

Not every piece of the ultimate competitive market model must be adopted at the same time. For instance, in the services provided by Disco, it may be most efficient to follow something like the telephone industry transition model where the traditional utility provided the service through the new Lineco, subject to regulation. Prices of the combined company could be regulated, but with the Disco charges separated from the Lineco charges for purposes of transparency. Over time, new entrants could emerge to compete with Disco, and after enough entry had developed, and with the experience of the practicality of separating Lineco from Disco, the old Disco would be fully independent and deregulated. In effect, separation of charges provides the opportunity for entry of competitors without the disruptive step of separating Disco and Lineco on the first day. Regulation of Disco charges in a transition provides the protection needed until viable competition emerges in the supply of distribution services.

Other innovations may require a minimum set of changes that occur simultaneously.

For example, movement to locational prices is naturally implemented as a one-time reform. Prices either differ by location, or they don't. However, there is no logical necessity to move immediately to short-run marginal costs in calculating the locational prices. Through the use of transitions formulas and long-term contracts, it is entirely possible to map out a transition path where the structure of pricing and billing follows the Poolco model, but the net effect is to start with average-cost prices at each location and move over a period of time to fully marginal-cost based prices.[49]

Implementation of the Poolco bidding and dispatch system would be a one-time change, but the change would not be very large for existing power pools which already use economic dispatch. Only the settlements system would be new, moving from payments based on split-savings to payments based on marginal cost. And new participants to the pool can be added over time, or connected pools can merge, without requiring a complete reorganization of the entire U. S. interconnected grid before any changes can be made.

These and other timing issues are important topics for definition, research and debate. It is clear that some steps must be taken early, including the commitment to move to the competitive market model. But other actions can be delayed in a gradual adaptation that allows for incremental change and learning along the way.

[49] The Trans Power New Zealand proposal adopts this approach with initial locational prices identical to the current system and a five year transition to fully marginal-cost prices including losses and congestion costs. See Trans Power New Zealand "Transmission Pricing 1993," , Wellington, New Zealand, February 1993.

Central Pool

Operation of a spot market with marginal-cost prices is easiest to envision when there is literally a tight central pool with single dispatcher. This would not be much different from the current tight power pools in the U. S. or the spot market in the U. K., with only the modification of using locational marginal costs. However, with the many more loosely coordinated dispatch control centers in the U. S., it is sometimes suggested that the competitive market model with its associated pricing and rights could not exist without wholesale institutional change to implement a single national power pool with a central dispatcher. Hence an appeal to the assumed requirement for a central pool is used to deflect further consideration of the possibility of implementing a workably competitive electricity market model.

Both the validity and impact of this critique are empirical matters, but they have little effect on the decision to adopt the competitive market model described here. Either a single central dispatcher is not necessary and we can move easily to a competitive market model; or the institutional reform of a more coordinated dispatch would be highly desirable and would reinforce movement to a competitive market model..

The need is not for a single dispatcher but for an efficient dispatch. One way to achieve the efficient dispatch is through the use of the bidding system and central selection of the economic use of generation or demand management. However, it is possible as a matter of principle to have interacting control centers implementing the direct management of generation and load and monitoring opportunities for short-run energy purchases or sales. To the extent that the existing amalgam of tight and loose power pools, control centers and economy sales achieve the benefits of economic dispatch, there is no need to create a new central dispatcher. All that is needed is to change the pricing and settlements systems to reference marginal costs rather than average costs or split savings.

If the current system is working, therefore, in the sense that no significant opportunities for short-term trades go wanting because of failures to communicate, then the required institutional change is confined to the pricing and payments practices. However, if the current system is broken and not achieving something close to the efficient dispatch, then more would be required and more would be indicated. If the present pool and control center interactions preclude or ignore significant opportunities for economical short-run trades, then for this reason alone it would be desirable to pursue and implement a more coordinated system of dispatch decisionmaking. But the primary purpose of the institutional reform would be to achieve the short-run benefits of an improvement in a significantly distorted dispatch. The benefit of being able to implement marginal cost pricing and open access in a competitive market would be a valuable by-product, but not the sole motivation.

The relevant issue, therefore, is the degree to which current dispatch practices support an efficient short-run use of the system. To the extent that otherwise economical short-run trades are precluded by the current institutional arrangements, the task independent of implementation of the competitive market model is to design reforms that move closer to the efficient dispatch. It should not be necessary to move all the way to the national power pool before any gains can be made. And a workable version of the competitive market model should be able to operate in conjunction with any reasonable system of efficient dispatch.

Jurisdiction

Regulatory oversight of the remaining franchise monopolies could rest with local, regional or national entities in the case of the Lineco and Gridco wires businesses. In the case of the dispatch operations of Poolco, either regional or national oversight is required, with the natural application of coordinated dispatch crossing many state boundaries.

Transition to this new industry organization and regulatory jurisdiction would not require a sudden or wide-scale reallocation of regulatory responsibility and authority. The existing combined distribution companies, with the embedded Linecos, are subject typically to state regulation. Eventual separation of the Buyco and Disco functions, if necessary, would leave the Lineco standing as the remaining company. The same state regulation of this essentially local monopoly franchise could continue.

The Gridco, separate from dispatch, provides only the wires design, construction and maintenance functions. It works in collaboration with Poolco in defining the transmission capacity rights for compensation, but the regulated portion of the business, and the cash flow of fixed charges to pay for the lines, would be amenable to virtually any organization of regulatory oversight. Although this could be a national function, in principle at least there is no reason that states could not continue to regulate this activity in much the same ways as in the traditional model. Customers would have to commit to pay the fixed charges before construction. Regulation would be focussed on cost recovery and reasonable rates of return to the Gridco, but local regulation would not extend to the much more complicated decisions about the actual use of the transmission network.

The principal potential problem with state regulation of the Gridcos would be application of a "not-in-my-backyard" syndrome designed to prevent construction of new transmission lines that would yield benefits primarily for residents in other states. Other remedies might be available, such as compensation or appeals to protection under interstate commerce statutes. To the extent that these avenues are not universally accepted as adequate protection, the alternative would be to move control of the transmission line construction from state jurisdiction to federal jurisdiction. This political decision would be a major impediment to complete transition to the competitive market model. However, it is not an "all or nothing" obstacle, and problems could be addressed in turn during the transition.

This Poolco operation is the one area where some form of organization and oversight that transcends state boundaries would be essential. Fortunately, this institutional requirement need not be a major obstacle, at least in the United States. Existing pools, both formal and informal, already operate across state boundaries with little operational supervision by any regulatory authorities other than the FERC. Modifications in the pool practice of economic dispatch are not required, although the process will change to one involving both supply and demand bids, rather than dispatch based on engineering estimates of generation cost.

The principal innovation of Poolco in the competitive market model is in moving to a marginal- rather than an average-cost price settlements process for all energy dispatch. Some approximation of marginal-cost pricing is likely to be an inevitable change that develops with more and more competition in generation, and average cost differences from existing plants can be accommodated through the introduction of long-term contracts that give existing ratepayers the benefits or obligations associated with the existing plants.

Negotiation of the transition arrangements for generation contracts will not be easy, but it should not be complicated by the need to reorganize regulation of the power pools. The pool operations have not been subject to extensive oversight under the traditional model, and supervision of Poolco operations can be both relatively light and limited to a national authority focussed on enforcing open access to encourage competition. It would be possible and natural for this regulatory change to evolve under national regulation as part of the process of defining the rules for open access and pricing.

Allocation of Sunk Costs

Long-term generation contracts for existing plants focus on one of the major transition questions that fall under the heading of stranded assets. Movement from the traditional regulatory model to a competitive market model is unlikely to occur under conditions of perfect balance in the allocation of costs and value for existing assets. Facilities such as high-cost generation plants or obligations such as capitalized demand-side costs exist with only implicit arrangements for cost recovery and responsibility. Under the strict competitive market model, these above-market costs for stranded assets cannot be recovered. Customers with a choice will choose to avoid paying such costs. To be collected, therefore, these costs must be accepted under some form of contract or other obligation.

Although the stranded asset costs represent only a transfer of responsibility for sunk costs, and therefore have no direct economic efficiency implications, assignment of the obligation to pay these costs will be the major challenge of the transition. It is not unusual to have estimates of the potential exposure to such stranded assets as exceeding the equity of existing utilities, which would lead to technical or real bankruptcy.[50] For many, the large magnitude of these "above market" costs and the hope of avoiding them, reinforces or even provides the motivation for moving to the competitive market model. And for those who would be left to absorb the stranded costs, there is a strong motivation to prevent or delay the transition.

The only places to collect these stranded asset costs will be through the remaining franchise monopoly, direct government taxes, or transition contractual agreements where acceptance of the obligation is part of the social compact to move to the competitive market model. Defining and developing these alternatives remains as a major challenge.

Retail Wheeling

A key decision in moving towards a more competitive market model is the degree to which there will be de facto or de jure retail wheeling. The EPAct authorizes and requires wholesale wheeling, and FERC is charged with responsibility to develop transmission access and pricing provisions which implement these mandates.[51] However, retail wheeling to final customers is explicitly excluded from the new authorities under the EPAct, and this arena remains the responsibility of the states.

As important as it is, wholesale wheeling at its maximum constitutes only a relatively small portion of the total pie. Although all customers receive power from the wholesale market indirectly, only a small fraction of all consumers have direct access as wholesale customers. In 1991 the share of reported revenues from wholesale transactions ranged from a low of 6.2% in

---

[50] S. Anderson, "Electricity Transition Costs," draft, October 1993.

[51] Energy Policy Act of 1992, Pub. L. No. 102-486, 106 Stat. 2776 , 1992; sections 721-726.

the eastern MAAC region to a high of 13.7% in the central ECAR region.[52] If it is possible to limit open access and wheeling to the wholesale market, the competitive market transition problems will be greatly reduced. In principle, therefore, it is conceivable to imagine a world in which progress in improved efficiency of operation of the electricity market is pursued within the market including wholesale customers and utilities acting on behalf of the final customers.[53] In effect, this would leave the Buyco and Lineco functions united, although it would still be possible to separate the Disco services, eventually.

Even if state regulators choose not to allow blanket retail wheeling, however, there may be little comfort for the owners of stranded assets. The practical concern is not with a de jure restriction on retail wheeling but with the de facto impact of the pressure to avoid high costs and seek low costs. In a world with a little bit of competition, and a little bit of access--but a lot of stranded costs--there will be enormous pressure for more competition, more access, and more movement from high cost to low cost status. Customers will change from a retail to a wholesale classification. Local political leaders will seek special "economic development" rates for special customers. It is difficult in practice to moderate the competitive pressures. If retail wheeling is coming by design or despite the design of the regulators, it will have a major impact on the scope and urgency of the transition problems. The issues are many, but the overarching requirement is to have a plan for the transition before politically irreversible commitments are made. If customers get access to the market before an allocation of the cost of stranded assets,

---

[52] North American Electric Reliability Council (NERC) regions include Mid-Atlantic Area Council (MAAC) and East Central Area Reliability Coordination Agreement (ECAR). Data from Fitch Investors Service, <u>Electric Utilities' Competitive Risk</u>, July 12, 1993, pp. 26-30.

[53] R. Cavanagh, "The Great 'Retail Wheeling' Illusion--And More Productive Energy Futures," September 1993, draft.

the customers will be gone for good.

CONCLUSION

The competitive electricity market model defines one vision of a possible future for the industry. Although technology requires that a few critical segments of the industry remain as regulated monopolies, it is possible to design an internally coherent market that makes maximum use of competition, especially in the generation sector. The key requirement is to maintain efficient, open access to the essential facilities. The need for access to the wires through the Gridco and Lineco is obvious and has many analogies in other industries such as natural gas pipelines. The important special role of the dispatch process, through Poolco, is less familiar but is required by the distinctive nature of electricity provided through a free-flowing grid. With everyone treated equally in the short-run dispatch, pricing and utilization replicate the outcome of an efficient market without violating the operational requirements for system stability.

Once the competitive market model is defined, it sets a target and a standard for designing changes from the traditional status quo. The transition problems include many elements that must be considered simultaneously as part of an implementation strategy. Consistency of the broad model and strategy will provide a framework that separates the few necessarily revolutionary elements from the more evolutionary changes. Analysis of the transition alternatives should focus on open access and pricing rules that are essential for the competitive goal, are feasible as a technical matter in electric networks, and are practical in dealing with stranded assets and other changing economic impacts.

APPENDIX

Competitive Electricity Market Pricing

Introduction

Pricing in a competitive electricity market is at marginal cost. The many potential suppliers compete to meet demand, bidding energy supplies into the pool. The dispatchers at Poolco choose the least-cost combination of generation or demand reductions to balance the system. This optimal dispatch determines the market clearing prices. All consumers pay this price into the pool and all generators in turn are paid this price for the energy supplied.

Inherently, energy pricing and transmission pricing are intimately connected. The FERC has outlined objectives for transmission pricing that would be compatible with a competitive market. These objectives include:[54]

- Promote efficient use of and investment in the transmission grid and provide appropriate price signals to transmission customers. To the extent practicable, prices should accurately:

  - account for transmission constraints
  - reflect any prudent costs incurred as a result of transmission service
  - reflect the actual power flows of the transmission service
  - reflect the distance- and location-sensitive costs of the transmission service
  - reflect the prevailing direction of the flow, distinguishing between "with the flow" and "counter flow"

[54] Federal Energy Regulatory Commission, "Transmission Pricing Issues," Staff Discussion Paper, Inquiry Concerning the Commission's Pricing Policy for Transmission Services Provided by Public Utilities Under the Federal Power Act, Washington, DC, June 1993, pp. 7-8.

- Address any transition problems arising from the reform

    - Balance equity considerations associated with any reform with the potential efficiency improvements
    - Mitigate the hardships arising from any reform

- Allow customers an option to have stable prices over time

- Be simple to implement and administer

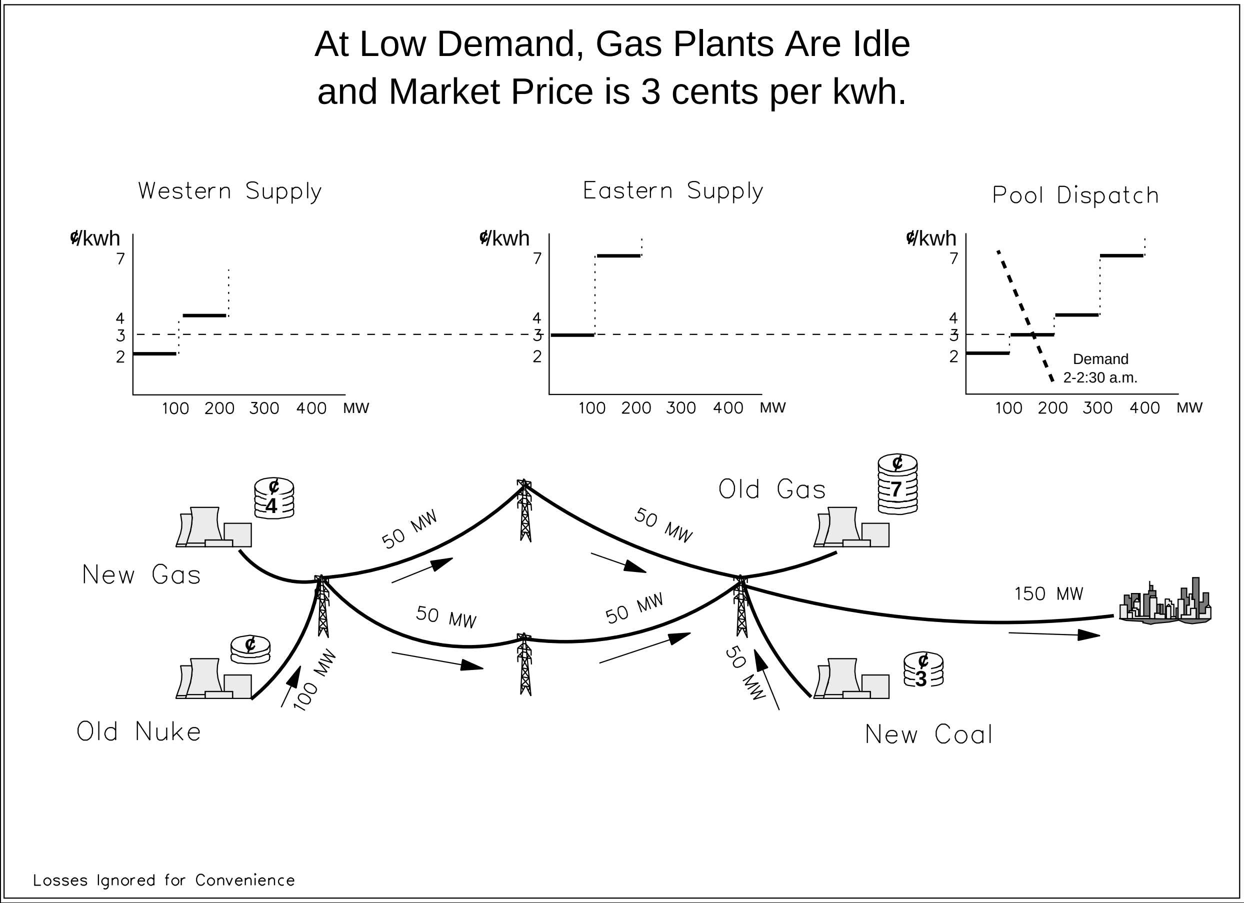
A series of examples of pricing in the competitive electricity market model illustrates the determination of prices and relates transmission constraints to congestion rentals that lead to different prices at different locations. These fundamentals provide the building blocks for an energy and transmission pricing system that addresses the several requirements of the FERC outline.

Economic Dispatch

Consider the simple market model in Figure 6, which will serve as the basis for a set of succeeding examples. In this market there is one load center, a city in the east, supplied by generators located far away in the west, connected by transmission lines, and by local generators who are in the same region as the city customers. The plants in the west consist of an "Old Nuke" which can produce energy for a marginal cost of 2 ¢/kwh and a "New Gas" plant that has an operating cost of 4 ¢/kwh. These two plants each have a capacity of 100 MW, and are connected to the transmission grid which can take their power to the market in the east.

The competing suppliers in the east are a "New Coal" plant with operating costs of 3 ¢/kwh and an "Old Gas" plant that is expensive to use with a marginal cost of 7 ¢/kwh. Again

Figure 6



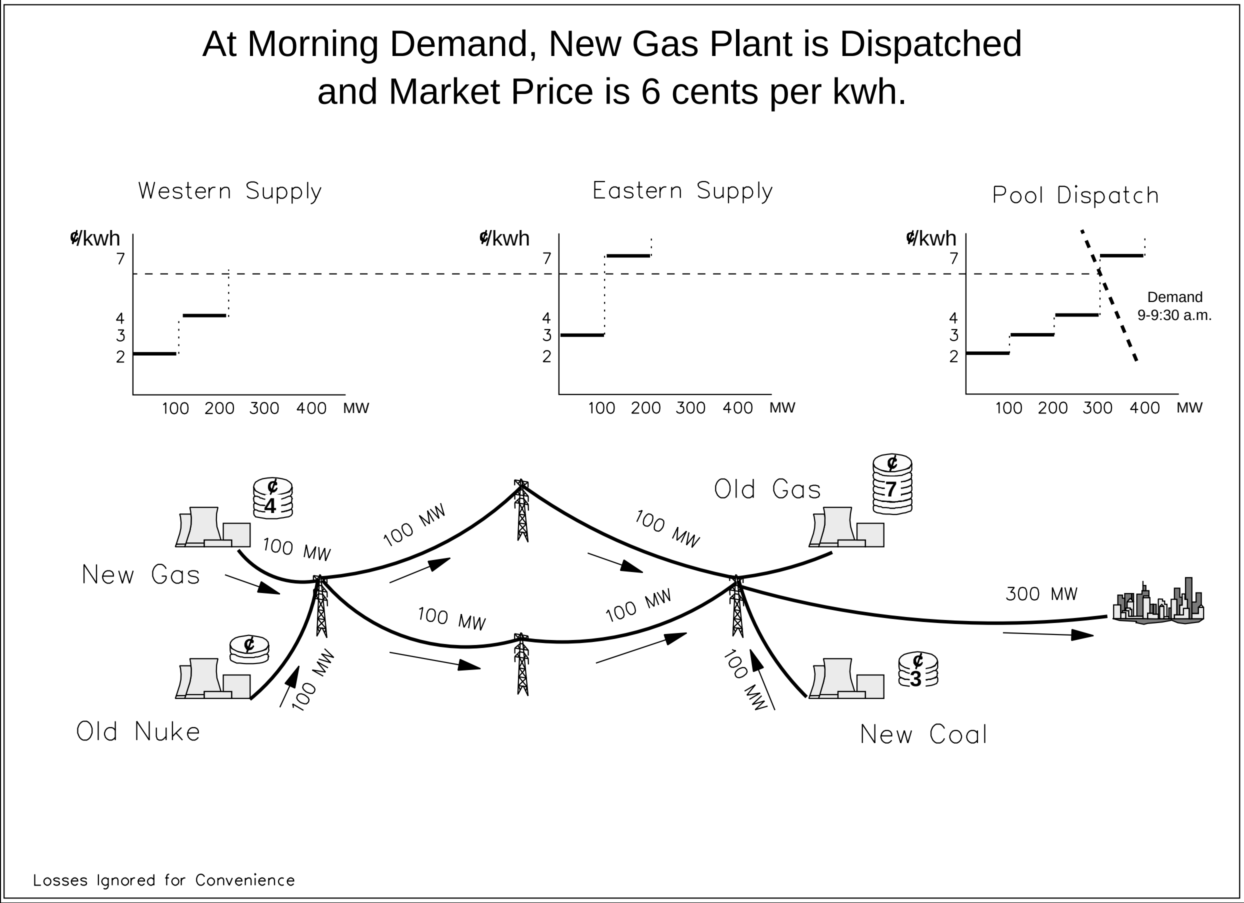
these eastern plants are assumed to have a capacity of 100 MW. The two plants in the west define the "Western Supply" curve, and the two plants in the east define the corresponding "Eastern Supply" curve. These supply curves could represent either engineering estimates of the operating costs or bids from the many owners of the plants who offer to generate power in the competitive market. For simplicity, we ignore transmission losses and assume that the same supply curves apply at all hours of the day.

Under low demand conditions, as shown in Figure 6 for the early hours of the morning, the supply curves from the two regions define an aggregate market supply curve that the Poolco dispatchers can balance with the customer demands. The aggregate market supply curve stacks up the various generating plants from cheapest to most expensive. The Poolco dispatchers choose the optimal combination of plants to run to meet the demand at this hour. In Figure 6, the result is to provide 150 MW. The inexpensive Old Nuke plant generates its full 100 MW of capacity, and the New Coal plant provides another 50 MW. The New Coal plant is the marginal plant in this case, and sets the market price at 3 ¢/kwh for this hour. Hence the customers in the city pay 3 ¢/kwh for all 150 MW. The New Coal plant receives 3 ¢/kwh for its output, and this price just covers its running cost. The Old Nuke also receives 3 ¢/kwh for all its 100 MW of output. After deducting the 2 ¢/kwh running cost, this leaves a 1 ¢/kwh contribution towards capital costs and profits for Old Nuke owners.

In this low demand case, and ignoring losses, there is no additional opportunity cost for transmission. The 100 MW flows over the parallel paths of the transmission grid. But there is no constraint on transmission and, therefore, no opportunity cost. Hence the price of power is the same in the east and in the west. In the short run, there is no charge for use of the transmission system.

If demand increases, say at the start of the business day, Poolco must move higher up on the dispatch curve. For example, consider the conditions defined in Figure 7. This hour presents the same supply conditions, but a higher demand. Now the Poolco dispatchers must look to more expensive generation to meet the load. The Old Nuke continues to run at capacity, the New Coal plant moves up to its full capacity, and the New Gas plant in the west also comes on at full capacity. The New Gas plant in the west is the most expensive plant running, with a marginal cost of 4 ¢/kwh. However, this operating cost cannot define the market price because at this price demand would exceed the available supply, and Poolco must protect the system by

Figure 7



maintaining a constant balance of supply and demand.

In this case, the result is to turn to those customers who have set a limit on how much they are willing to pay for electric energy at that hour. This short-run demand bidding defines the demand curve which allows Poolco to raise the price and reduce consumption until supply and demand are in balance. In Figure 7 this new balance occurs at the point where the market price of electricity is set at 6 ¢/kwh. Once again, the customers who actually use the electricity pay this 6 ¢/kwh for the full 300 MW of load at that hour. All the generators who sell power receive the same 6 ¢/kwh, which leads to operating margins of 2 ¢/kwh for New Gas, 3 ¢/kwh for New Coal, and 4 ¢/kwh for Old Nuke.

Once again, the Poolco dispatch in Figure 7 depends on excess capacity in the transmission system. The plants in the western region are running at full capacity, and the full 200 MW of power moves along the parallel paths over the grid to join with New Coal to meet the demand in the east. There is a single market price of 6 ¢/kwh, and there is no charge for transmission other than for losses, which are ignored here for convenience in the example.

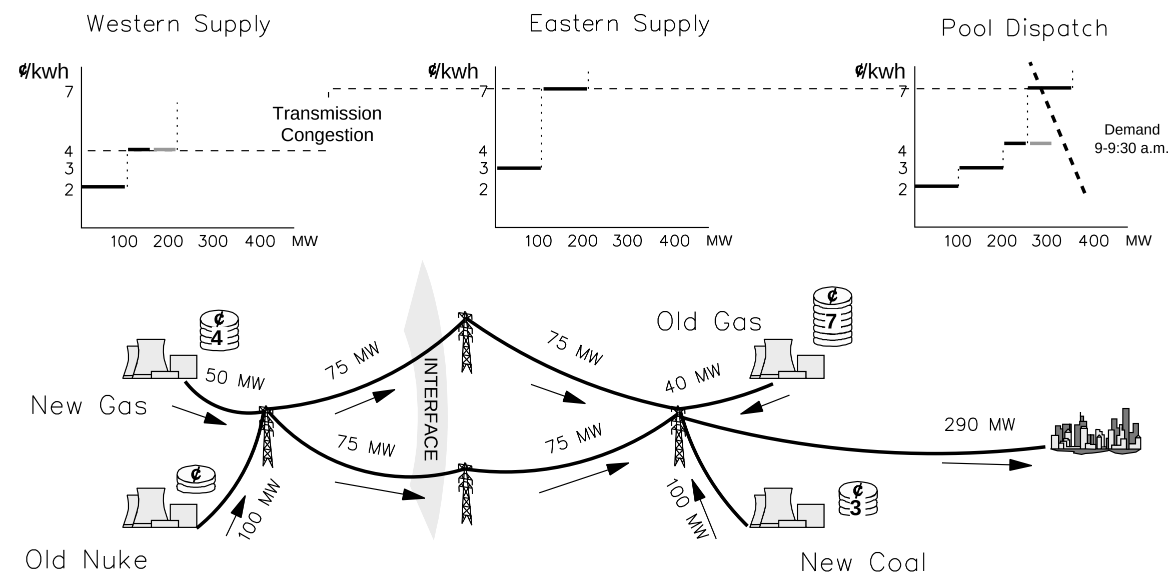
Transmission Constraints

With the plants running at full capacity, there might be a transmission constraint. To illustrate the impact of a possible transmission limit, suppose for sake of discussion that there is an "interface" constraint between west and east. According to this constraint, no more than 150 MW of power can flow over the interface.

As shown in Figure 8, this transmission constraint has a significant impact on both the dispatch and market prices based on short-run marginal costs. In Figure 8 the level of demand from the city in the east is assumed to be the same as in the case of Figure 7. However, now the Poolco dispatcher faces a different aggregate market supply curve. In effect, only half of the New Gas output can be moved to the east. To meet the demand, it will be necessary to simultaneously turn off part of the New Gas output and substitute the more expensive Old Gas generation which is available in the East. This new dispatch increases the market price in the east to 7 ¢/kwh and necessarily induces a further reduction in demand, say to a total of 290 MW. The New Coal and Old Gas plants receive this full price of 7 ¢/kwh for their 140 MW, which provides a 4 ¢/kwh operating margin or short-run profit for New Coal and allows Old Gas to cover its operating costs.

Figure 8

At Higher Demand, Transmission Interface Constraint Yields Congestion; Market Price is 7 cents in the East and 4 cents in the West.



Transmission Constraint Creates 3 cent Congestion Rental.

Losses Ignored for Convenience

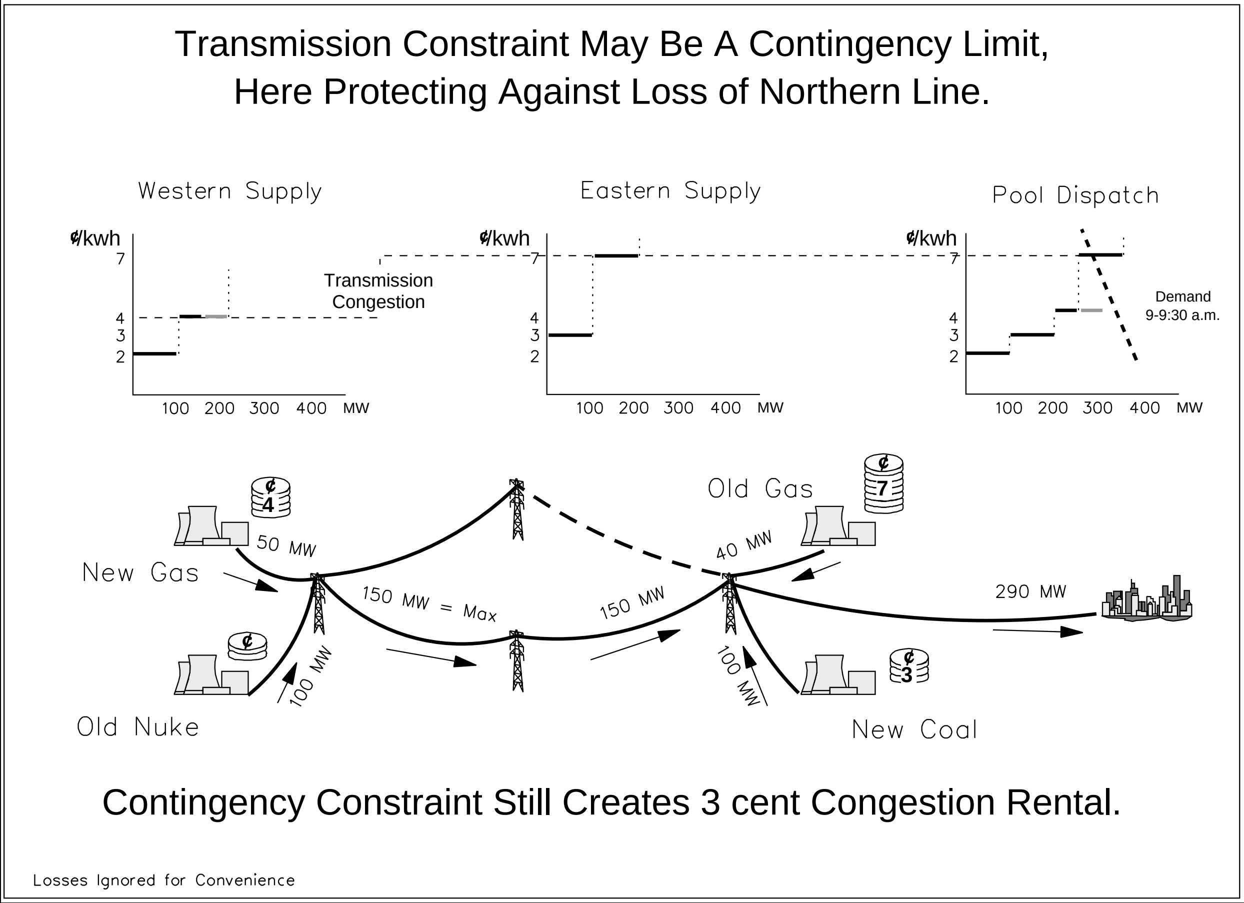
In the western region, however, a different situation prevails. The transmission interface constraint has idled part of the output of the New Gas plant. Clearly the market price in the west can be no more than the operating cost of the plant. Likewise, since the plant is running at partial output, the market price can be no less than the operating cost of 4 ¢/kwh. This is the price paid to New Gas and Old Nuke, which covers New Gas operating costs and provides Old Nuke an operating margin of 2 ¢/kwh.

The 3 ¢/kwh difference between the market price in the east and the market price in the west is the opportunity cost of the transmission congestion. In effect, ignoring losses, the marginal cost of transmission between west and east is 3 ¢/kwh, and this is the price paid implicitly through the transactions with Poolco. Electricity worth 4 ¢/kwh in the western region becomes worth 7 ¢/kwh when it reaches the eastern region.

The transmission "interface" constraint is a convenient shorthand for a more complicated situation handled by the Poolco dispatchers. The interface limit depends on a number of conditions, and can change with changing loads. Typically it is not the case that there is a 75 MW limit on one or both of the parallel lines through which power is flowing in the grid. In normal operation, it may well be that the transmission lines could individually handle much more flow, say 150 MW each or twice the actual use. At most normal times, the lines may be far from any physical limit. However, the Poolco dispatchers must protect against contingencies--rare events that may disrupt operation of the grid. In the event of these contingencies, there will not be time enough to start up new generators or completely reconfigure the dispatch of the system. The power flow through the grid will reconfigure immediately according to the underlying physical laws. Hence, generation and load in normal times must be configured, and priced, so that in the event of the contingency the system will remain secure.

For instance, suppose that the thermal capacity of the transmission lines is 150 MW, but the Poolco dispatchers must protect against the loss of a northern transmission line. In this circumstance, the actual power flows may follow Figure 8, with 75 MW on each line, but the Poolco dispatchers must dispatch in anticipation of the conditions in Figure 9. Here the northern line is out, and in this event the flow on the southern line would hit the assumed 150 MW thermal limit. This contingency event may never occur, but in anticipation of the event, and to protect the system, Poolco must dispatch according to Figure 9 even though the flows are as in Figure 8. In either case, the transmission constraint restricts the dispatch and changes the market

Figure 9



prices. The price is 4 ¢/kwh in the west and 7 ¢/kwh in the east, with the 3 ¢/kwh differential being the congestion-induced opportunity cost of transmission. This "congestion rental" defines the competitive market price of transmission.

Buying and selling power at the competitive market prices, or charging for transmission at the equivalent price differential provides incentives for using the grid efficiently. If some user wanted to move power from east to west, the transmission price would be negative, and such "transmission" would in effect relieve the constraint. The transmission price is "distance- and location-sensitive," with distance measured in electrical rather than geographical units. And the competitive market prices arise naturally as a by-product of the optimal dispatch managed by

Poolco.

Transmission Rights

The congestion rental received by Poolco provides the key to defining property rights in the transmission grid. In the face of transmission constraints, prices will be more volatile and it will not be possible for a generator to provide guaranteed price stability in the form of a long-term contract with a customer. Furthermore, customers and generators will have an incentive to expand the grid only if they recognize and believe that after making an investment in the grid there will be some protection against any future congestions costs. A simple way to define the property right and provide this guarantee is to assign the congestion rental not to Poolco but to the holder of the transmission right. In the transmission constrained cases of Figure 8 or Figure 9, the rights to transmit a total 150 MW of power might have been held by customers in the city, or by generators in the west. In either case, this right is defined only as the right to collect the congestion rental. The generators and customers would not control the use of the grid. Poolco would determine the efficient pattern of use through economic dispatch. Poolco would collect the congestion payments from the actual users of the grid and pay them in turn to the holders of the transmission rights.

With this definition of transmission rights, it is an easy matter for generators at Old Nuke and New Gas to arrange long-term contracts that provide price stability for customers in the city. For example, the owners of Old Nuke may have acquired rights to 100 MW, and signed long-term contracts that guaranteed to provide power delivered to the city at a price of 5 ¢/kwh. In the case of low demand as in Figure 6, the short-run price is only 3 ¢/kwh, which customers pay and generators receive through Poolco. Separate from Poolco, the customers pay Old Nuke the difference of 2 ¢/kwh owed under the contracts. If demand shifts to the higher case in Figure 7, the market price is 6 ¢/kwh, and again the customers pay and generators receive this short-run price through Poolco. In this event, the generators separately pay the customers the difference of 1 ¢/kwh required under the long-term contract.

When transmission constraints bind as in Figure 8 or Figure 9, the price paid by the customers to Poolco is 7 ¢/kwh, and the price received by the generators from Poolco is 4 ¢/kwh. If the generators own the transmission rights, then Poolco pays the generators an additional 3 ¢/kwh which allows the generators in turn to pay the customers the 2 ¢/kwh difference agreed to by contract. The owners of Old Nuke are always making an operating margin of 3 ¢/kwh, and the customers are always receiving the equivalent of 5 ¢/kwh electricity. Likewise, if the customers own the transmission rights, the customers receive the 3 ¢/kwh from Poolco and in turn pay 1 ¢/kwh to the generators. Again the owners of Old Nuke are always making an operating margin of 3 ¢/kwh, and the customers are always receiving the equivalent of 5 ¢/kwh electricity. Furthermore, in either case Poolco ends up with no transmission congestion rentals; Poolco serves only to pass through the congestion costs from the actual users of the grid to the holders of the property rights in the economic interest of the transmission grid.

This system of transmission rights and property payments back and forth may seem unnecessary and cumbersome in the case of the simple system of Figure 6 through Figure 9. After all, couldn't the Poolco dispatchers in effect assign the generation from Old Nuke to the long-term customers in the city? In principle, this specific performance model--assigning particular generation to particular users--is possible in this simple case, but it does not generalize into the more complicated reality of an interconnected grid with many different sites of load and generation, real and reactive power, thermal and voltage limits, and multiple contingencies. It is impossible in a real system to meaningfully assign any particular sources and destination of electricity, and attempts to do so can only serve to compromise the efficiency objective of maintaining an optimal dispatch which may require only partial use of plants in constrained regions, violating the assumptions of specific performance. However, the payments of congestion rentals from Poolco to the holders of point-to-point transmission rights do generalize to the more complicated case, and allow optimal dispatch for efficiency while accommodating long-run contracts for price differences and congestion rentals, contracts that provide both stability and the essential protection of investment in the network.

Transmission Rights in a Network

These examples of efficient pricing and transmission rights extend beyond the case of two locations to the general case of a network with the complications of electricity system operation and prices at many locations. Electric energy pricing in a pool with economic dispatch is closely connected with transmission pricing. With multiple participants in a pool the efficient choice is to price power at the short-run marginal cost at each location. With such efficient pricing, transmission use pricing appears automatically in pool pricing: the differences between locational prices are the opportunity cost prices for transmission. At a minimum, locational prices would reflect differences in marginal losses. To the extent that there is congestion in the transmission grid, locational prices would also differ by the cost of congestion induced by "out-of-merit" generation. All users pay or are paid by the pool at these short-run prices. Through these prices the pool would collect congestion rentals.

Changing prices at locations would create an interest in transmission rights that would protect those who invest in the grid. Because of the effects of network interactions, it is not possible to guarantee simultaneously both specific performance--particular plants operating for particular customers--and constrained economic dispatch. If specific plants must be run for specific customers, then the system operators do not have the freedom to provide the least-cost dispatch. If the operators provide the least-cost dispatch, then constraints may preclude using specific plants and require alternative plants to run.

However, it is possible to define point-to-point transmission rights that make payments to the right holders in the event of constrained transmission in the grid. These point-to-point price protection transmission rights could be defined in either of two equivalent ways:

> <u>Rental Right:</u> The right to collect the difference in congestion rentals for a specified quantity between two locations; or
>
> <u>Purchase Right</u>: The right of customer at one location to an economic transaction equivalent to purchasing a specified quantity at another location with transmission at the cost of marginal losses.

The total quantity of these rights can be defined for a given configuration of the network, and the rights guaranteed for any pattern of loads in the network. In a real system, the rights would respect all the constraints in the grid, including contingency constraints for thermal limits on lines and voltage limits at buses.

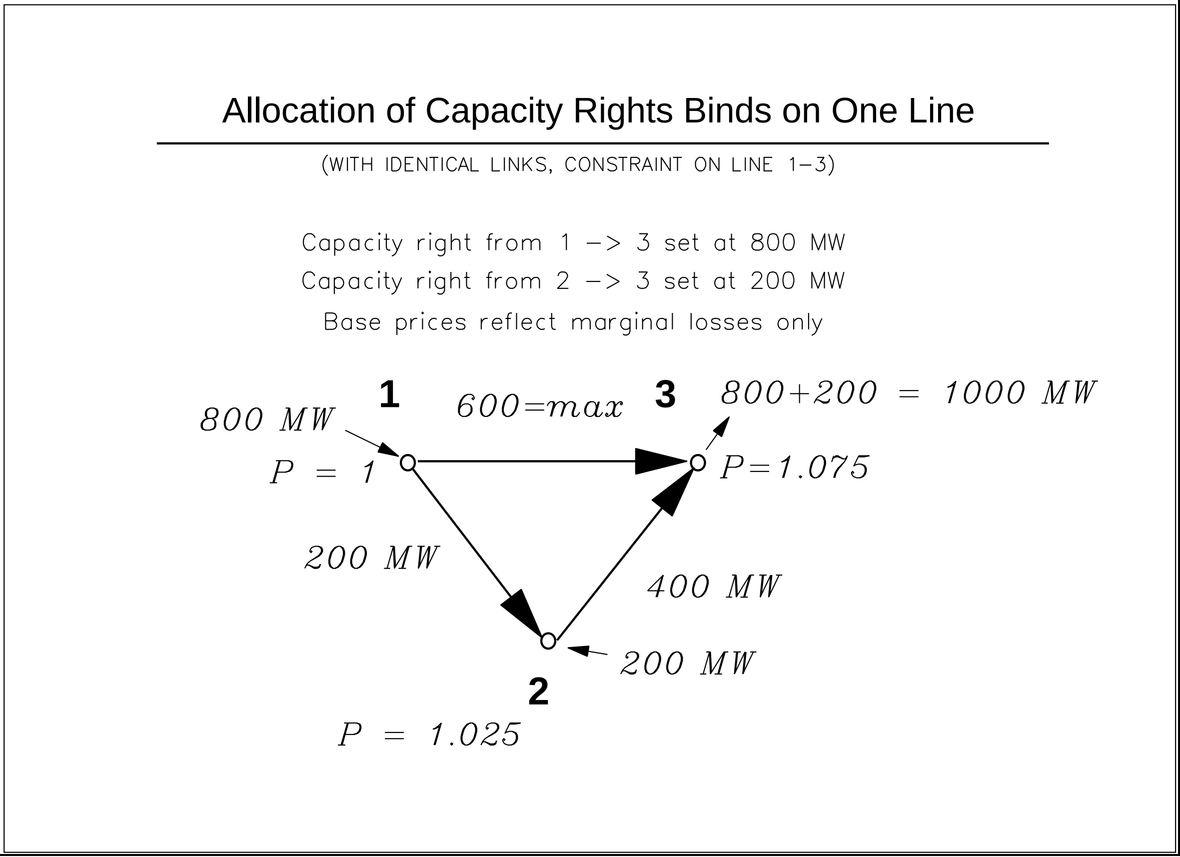
The point-to-point rights can be provided by the pool and allow protection of new investment in generation and transmission. The pool takes no risk in offering these rights because the revenue collected from users paying at short-run marginal cost is always great enough to honor the obligations under the rights. Under certain circumstances, the revenues collected by the grid will be greater than the obligations on the point-to-point rights, and there will be "excess congestion rentals." For example, if there are no point-to-point rights assigned, then none of the congestion rental would be paid out under such rights. In the more interesting cases, if point-to-point rights are assigned up to the limit of some transmission constraint, it is possible that some other network constraint will be binding in the actual dispatch. In these cases, there will be more than enough revenues to honor the rental or purchase rights, and the pool will face the added task of allocating the excess congestion rentals.

The natural assignment of any excess rentals is to those who are paying the fixed charges. There is no single rule for allocating the rights to the excess congestion rentals. The rule examined here is to share the revenues according to the share of the fixed charges. Hence, the owners of the grid who commit to pay the fixed charges have access to two types of well defined and tradeable rights: the point-to-point rental purchase right and a share in any excess congestion rentals.

Network Examples

For purposes of further illustration, consider the case of a three bus network with identical lines and identical thermal limits on each line. A three bus network is the minimum case needed to observe the network interaction effects of loop flow. Here we use the DC-Load approximation for real power only, and ignore contingency constraints. Reactive power and contingency constraints can be included without changing any of the fundamental points examined here.[55]

Figure 10



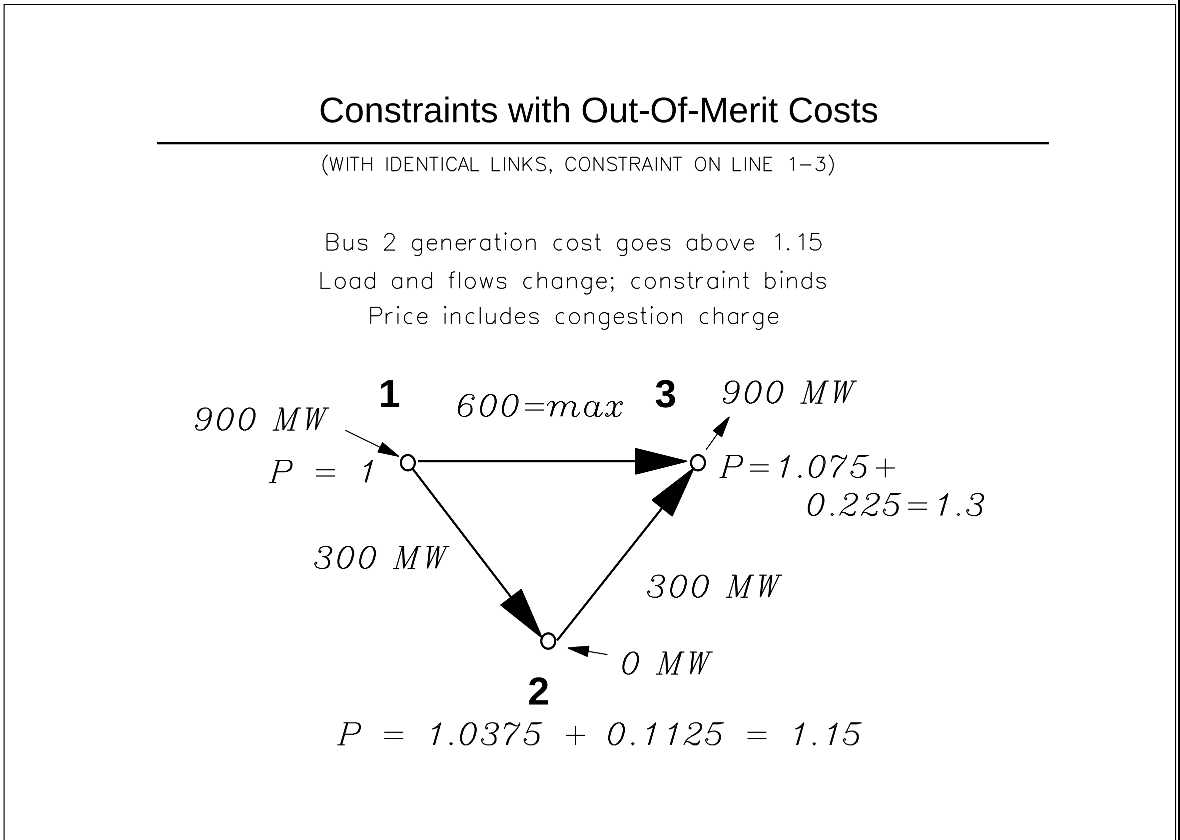
The base case model and an allocation of rights are shown in Figure 10. Here we assume that the desired transmission rights are for 800 MW from Bus 1 to Bus 3, and 200 MW from Bus 2 to Bus 3. Or, an equivalent definition is that the customer at Bus 3 has the right to purchase 800 MW at Bus 1 and 200 MW at Bus 2. The simultaneous allocation of these rights is feasible, but it does hit the thermal transmission constraint of 600 MW on the line between Bus 1 and Bus 3.

In Figure 10 the prices calculated for this dispatch are shown relative to the price at

55 W. W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, No. 3, September 1992, pp. 211-242.

Bus 1. In this instance there is no congestion and the prices cover only the cost of generation at Bus 1 and the marginal cost of losses. In this simplified case, the equilibrium required is that the marginal losses are linear in the flow on any link and are the same along any parallel path. Hence the marginal loss of one additional megawatt from Bus 1 to Bus 3 is 0.075, whether by the path 1→3 or via 1→2→3. There is no additional congestion cost, and hence there is no payment from the pool under the congestion rental contracts. Equivalently, the customers at Bus 3 buy their 800 MW from Bus 1 and 200 MW from Bus 2, just as specified in the transmission rights.

Figure 11



Of course, a change in the economics of generation could induce transmission congestion with the associated differences in prices across locations. In Figure 10, it was economic to generate power at Bus 2 and the actual economic dispatch is the same as the dispatch with simultaneous use of the allocated rights. In Figure 11, the assumed conditions change with an increase in the running cost of power at Bus 2 and the need to use expensive generation at Bus 3. If the gross load at Bus 3 is still 1000 MW, then part of the load must be met with local generation, which costs 1.3, including a congestion rental of 0.225. At this price for Bus 3 and with these loads and flows, the price at Bus 2 is determined by the equilibrium conditions of optimal economic dispatch. The easiest way to verify the equilibrium prices is to assume that an additional 2 MW of power could be supplied at Bus 2. This would allow a reduction of 1 MW at Bus 1 and an increase of 1 MW delivered to Bus 3 that could displace the expensive generation at Bus 3. The flow from 1→2 would reduce by 1 MW, with a corresponding 1 MW increase along 2→3. The flow along 1→3 would still be at the limit of 600 MW, and total losses would be the same. Hence the net savings would be 1 unit at Bus 1 and 1.3 units at Bus 3, for a total of 2.3 units. This implies a price for the incremental generation of 2 MW at Bus 2 of 2.3÷2 = 1.15, which is the equilibrium price. Along with the marginal losses, this total opportunity-cost price implies a Bus 2 congestion price of 0.1125.

By assumption, the cost of the generating plant at Bus 2 is above 1.15, so the plant does not run. Now all the power transmitted is generated at Bus 1, and only 900 MW can be transmitted. The thermal constraint of 600 MW on the line between Bus 1 and Bus 3 is binding. All the users of the grid pay or are paid these prices for the actual dispatch. In addition, the holders of the point-to-point transmission rights receive payments from the pool operators.

The resulting payments are shown in Table II. Hence the owner of the 800 MW right from Bus 1 is paid the congestion rental difference from Bus 1 to Bus 3 of 0.225 for the full 800 MW, requiring a payment from the pool of 180. Likewise, the owner of the 200 MW right is paid the difference in the congestion rental between Bus 3 and Bus 2, or 0.1125, for the full 200 MW right and a payment of 22.5. Both users actually buy power from the pool at Bus 3 for the price of 1.3. For the customer with the 800 MW right to purchase at Bus 1, the payment of 180 is the total value of the congestion price differential between Bus 1 and Bus 3. And for the customer with the right for 200 MW at Bus 2, the payment of 22.5 is the total value of the congestion price differential between Bus 1 and Bus 2.

By purchasing 200 MW at Bus 3 at a price of 1.3 and then applying the transmission right payment of 22.5, the holder of the 200 MW transmission right can in effect purchase 200 MW at the price at Bus 2 and pay only the marginal losses to move the power to Bus 3. Although the actual dispatch is different than the simultaneous allocation of rights, the payments to the right holders provide the guarantee in effect that the right holders can purchase power at the price of power at another location. This holds true even if no power was actually generated at that location, as here for Bus 2. Furthermore, specific performance to actually generate and transmit the 800 MW and 200 MW according to the rights would not be feasible under this economic dispatch. Only by foregoing the advantages of the economic dispatch, and increasing total costs, could the specific plants be used for specific customers. The rights guarantee the economic value of the transmission, but do not determine the actual flows.

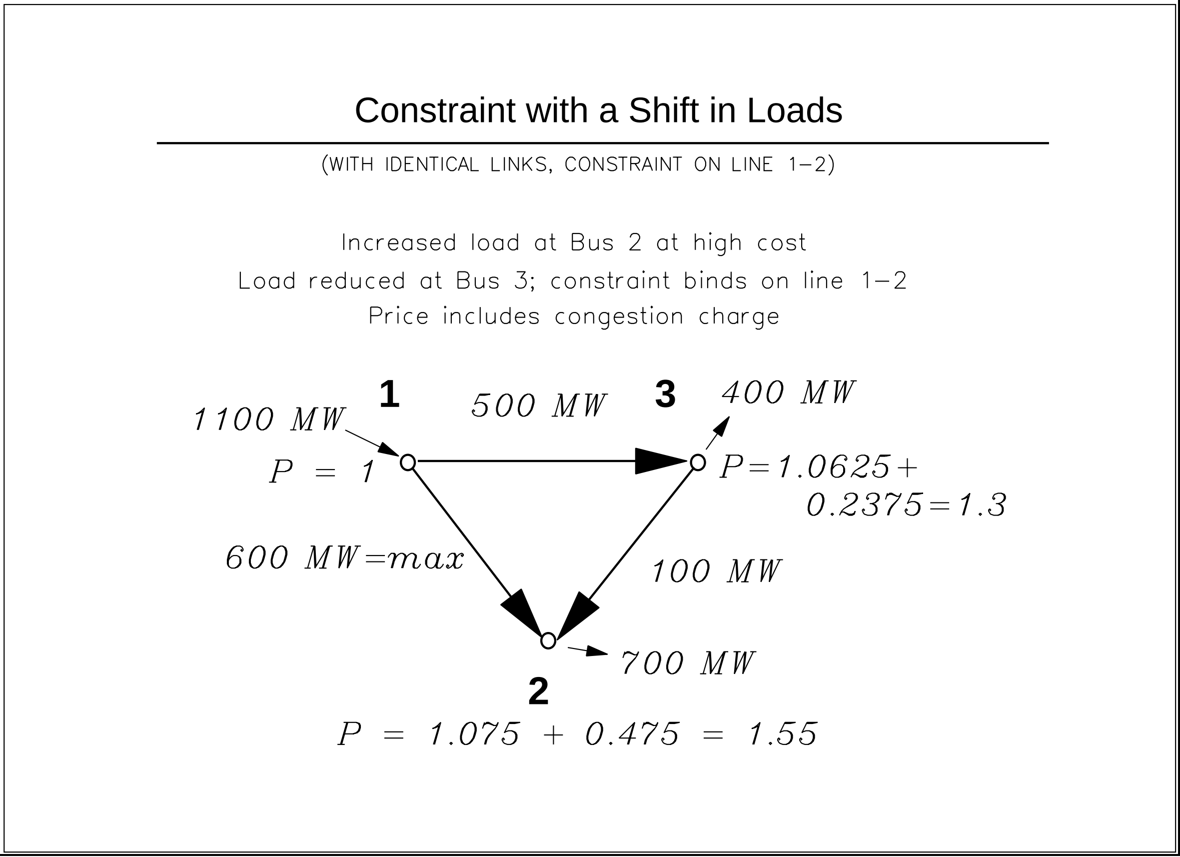
Table II: Congestion with Out-Of-Merit Costs						
Capacity Rights	Q (MW)	Congestion Price Difference	Direct Receipts	Excess Share	Excess Rentals	Net Receipts
1->3	800	0.225	180	80.00%	0	180
2->3	200	0.1125	22.5	20.00%	0	22.5
Total			202.5		0	202.5

The example in Figure 11 finds the pool dispatcher collecting congestion rentals from the actual users and paying the same rentals to the owners of transmission rights. Because the same transmission constraint limits both the actual dispatch and the initial allocation of transmission rights, there are no excess congestion rentals. All the congestion revenue collected is required to compensate the holders of the point-to-point rights.

An alternative case is shown in Figure 12. In this case the economics of load and dispatch have changed significantly. Power still costs 1.3 at Bus 3, but now the net load is reduced to 400 MW. There is a big net load at Bus 2, and the equilibrium power cost there is at 1.55. The relatively cheap generation at Bus 1 is used at the level of 1100 MW, which causes a shift in the flows. Now the dispatcher has no problem with a thermal limit on the line between Bus 1 and Bus 3, but the line between Bus 1 and Bus 2 has reached a similar thermal limit at 600 MW. This transmission constraint induces the indicated bus prices and congestion rentals.

Again the pool pays or is paid the short-run prices for power at each of the locations. And again the pool makes payments to the holders of the point-to-point rights. The summary of the various payments appears in Table III. For the customer with the right for 800 MW between Bus 1 and Bus 3, the congestion differential is 0.2375 and the total payment from the

Figure 12



pool is 190. This allows the right holder to purchase 800 MW at Bus 3--with some of that power necessarily generated by plants located at Bus 3--pay the price of 1.3, and after netting out the payment of 190 from the pool, effectively purchase the 800 MW at Bus 1 and pay only marginal losses to move the power to Bus 3.

Similarly, the customer with the right for 200 MW from Bus 2 to Bus 3 can purchase 200 MW at Bus 3 at the price of 1.3. However, this price is lower than the price at Bus 2, and the difference in congestion rentals is now negative, at -0.2375. Under the terms of the point-to point right, this customer must make an added payment of 47.5 to the pool. When coupled with the purchase of 200 MW at Bus 3, this is equivalent to purchasing 200 MW at Bus 2 at 1.55 and then moving to Bus 3 paying only the marginal losses (in this case the marginal losses also would be negative between Bus 2 and Bus 3). The final effect is as promised under the transmission right of the customer at Bus 3 to purchase 200 MW at Bus 2.

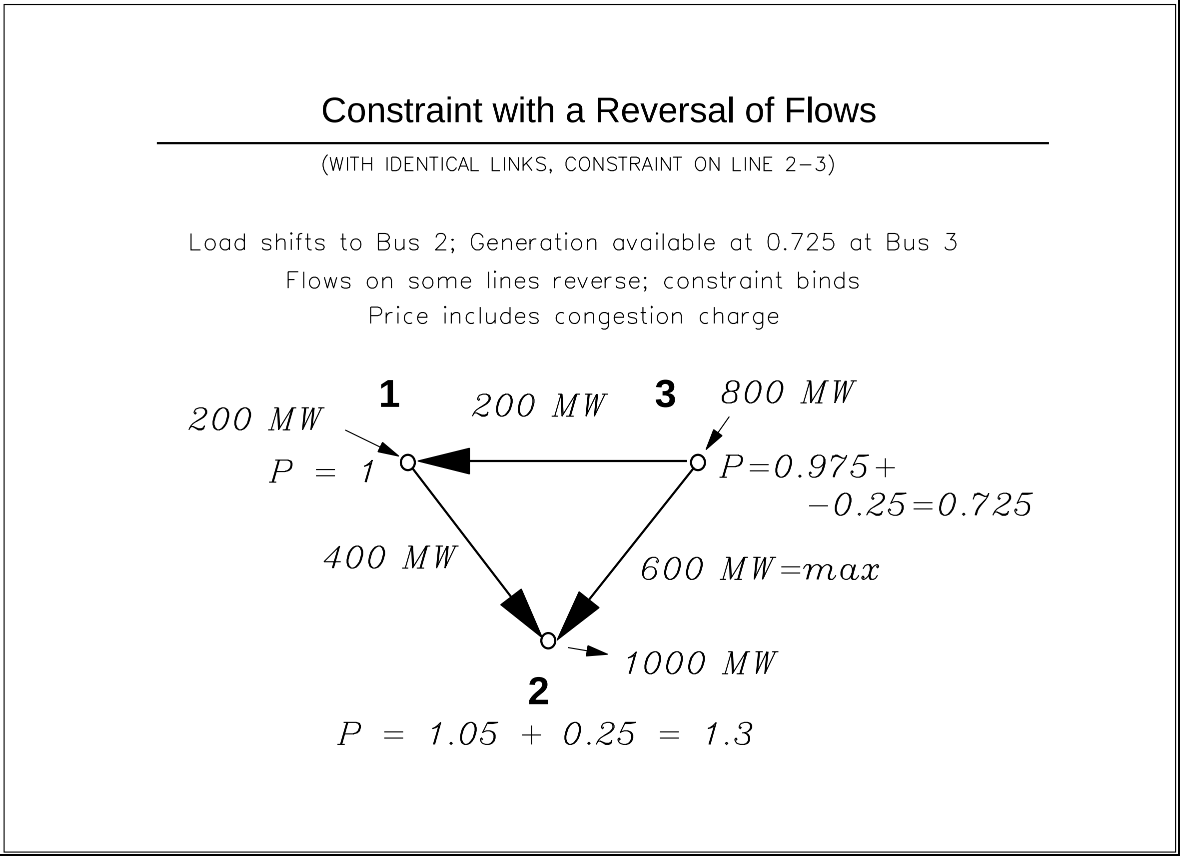
Table III: Constraint with a Shift in Loads

Capacity Rights	Q (MW)	Congestion Price Difference	Direct Receipts	Excess Share	Excess Rentals	Net Receipts
1->3	800	0.2375	190	80.00%	228	418
2->3	200	-0.2375	-47.5	20.00%	57	9.5
Total			142.5		285	427.5

In this case of a shift in loads, with a new transmission constraint binding, the pool can make all the necessary payments to the holders of the point-to-point rights, but the payments out amount to only 190 - 47.5 = 142.5. However, the total for the congestion rentals paid by the users of the grid is 700*0.475 + 400*0.2375 = 427.5. There remain excess congestion rentals of 285. Assuming that the fixed charge payments are proportional to the transmission capacity rights, one way to dispose of these excess congestion rentals would be to pay them out to the right holders in the ratio of 80 to 20. Hence the transmission right holder from Bus 1 would receive an additional payment of 228, for total receipts of 418. Likewise, the transmission right holder from Bus 2 would receive 57 from the excess congestion rentals for total receipts of 9.5.

In Figure 12 there is a shift in load and economics, and one of the transmission right holders, for power from Bus 2, is required to make addition congestion payments to the pool. With enough of a change in the loads and transmission flows, it is possible that everyone with a transmission right holds them in the reverse direction, and in this case the payments under the

Figure 13



sharing of excess congestion rentals take on added importance. For example, consider the conditions depicted in Figure 13. Here the economics of the dispatch and load have changed even more dramatically compared to the initial allocation of transmission rights. Now there is low price at Bus 3 and a net input of 800 MW, and the higher price is at Bus 2 with a net load of 1000 MW. The flows on the links from Bus 3 are now reversed.

The prices at the buses include a positive congestion component at Bus 2 and a negative congestion impact at Bus 3, all relative to Bus 1. Once again, the users of the grid pay or are paid according to these short-run marginal cost prices. The pool collects the payments and, in turn, makes the necessary payments to the holders of the transmission rights. In this case, both the customers with rights to Bus 1 and those with rights to Bus 2 face negative congestion rent differentials. Hence the customer with rights of 800 MW from Bus 1 sees a differential of -0.25, and makes a total additional payment to the pool of 200. With the purchase of 800 MW at Bus 3 at the price of 0.725, these combined payments are equivalent to a purchase of 800 MW at Bus 1 and then moving to Bus 3 at the cost of marginal losses.

For the customer with a 200 MW right to Bus 2, the congestion price difference is -0.5, and the direct payment to the pool is 100. These payments from the right holders to the pool add to the total congestion rentals collected by the pool from the actual users of the grid, who pay -800*(-0.25) + 1000*0.25 = 450. In all, as summarized in Table IV, there are 750 units of excess congestion rentals. As before, these congestion rentals could be distributed according to the share in the fixed cost allocation. In the present example, this would provide a payment of 600 to the customer with rights of 800 MW from Bus 1, for net receipts of 400; and 150 for the customer with rights of 200 MW at Bus 2, for net receipts of 50.

Table IV: Constraint with a Reversal in Loads

Capacity Rights	Q (MW)	Congestion Price Difference	Direct Receipts	Excess Share	Excess Rentals	Net Receipts
1->3	800	-0.25	-200	80.00%	600	400
2->3	200	-0.5	-100	20.00%	150	50
Total			-300		750	450

In all cases, the net effect of economic dispatch, marginal cost pricing, and assignment of transmission rights is to collect congestion costs from the actual users of the grid and pay the congestion costs to those who bear the burden of the fixed charges. The transmission rights based on price differences compromise two forms. First, point-to-point rights can be offered which provide the economic equivalent of a customer at one location always having the effective right to buy delivered power at the cost at a distant location plus the marginal losses. Second, to the extent that there are excess congestion rentals, these rentals can be distributed according to some agreed formula. In aggregate, the congestion rentals paid are always adequate to honor the point-to-point rights, and sometimes there can be additional rents that could be dispersed according to a sharing formula. In some instances, the congestion payments under the point-to-point rights can be negative, but only when the economics of the dispatch have switched to provide the right holder, who has access to cheap generation, the money from its operating margin through the pool dispatch that can fund the congestion payments. The pool dispatcher and operator of the settlements system is taking no financial risk in providing these price guarantees, and the actual dispatch is not constrained by the transmission rights. The dispatcher always has the freedom to provide the most economical generation possible given the current costs, bids, and system constraints.

-end-