TRANSMISSION INVESTMENT AND COMPETITIVE ELECTRICITY MARKETS

WILLIAM W. HOGAN

April 1998

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

CONTENTS

TRANSMISSION INVESTMENT
AND COMPETITIVE ELECTRICITY MARKETS

William W. Hogan[1]

A short-term electricity market coordinated by a system operator provides a foundation for a competitive electricity market. In this setting, transmission congestion contracts are equivalent to perfectly tradeable physical transmission rights. With such contracts to allocate transmission benefits, it would be possible to rely more on market forces, partly if not completely, to drive transmission expansion.

INTRODUCTION

Electricity market restructuring emphasizes the potential for competition in generation and retail services, with the operation of transmission and distribution wires as a monopoly. Network interactions and economies of scale both complicate the extension of market incentives to investment in the wires business. Two broad approaches suggest themselves for dealing with these network problems in a manner compatible with a competitive generation market.

There could be monopoly management of transmission operations and investment, with incentive pricing for the monopoly. Transmission would be like a large "black box" run by the monopoly that takes on an obligation to provide unlimited transmission service for everyone. With the appropriate price cap or other incentive regulation, the monopoly would make efficient investments or contract with market participants to remove or manage the real transmission limitations. This approach results in a powerful monopoly with the familiar problems of finding the right level and form of incentive regulation. This approach is, for example, found in the

---

[1] Lucius N. Littauer Professor of Public Policy and Administration, John F. Kennedy School of Government, Harvard University, and Senior Advisor, Putnam, Hayes & Bartlett, Inc. This paper draws on work for the Harvard Electricity Policy Group and the Harvard-Japan Project on Energy and the Environment. Many individuals have provided helpful comments, especially Robert Arnold, John Ballance, Jeff Bastian, Ashley Brown, Michael Cadwalader, Judith Cardell, John Chandley, Doug Foy, Hamish Fraser, Geoff Gaebe, Don Garber, Scott Harvey, Stephen Henderson, Carrie Hitt, Jere Jacobi, Paul Joskow, Susan Kaplan, Maria Ilic, Laurence Kirsch, Jim Kritikson, Dale Landgren, William Lindsay, Amory Lovins, Rana Mukerji, Richard O'Neill, Howard Pifer, Susan Pope, Grant Read, Bill Reed, Joseph R. Ribeiro, Brendan Ring, Larry Ruff, Michael Schnitzer, Hoff Stauffer, Irwin Stelzer, Jan Strack, Steve Stoft, Richard Tabors, Julie Voeck, Carter Wall and Assef Zobian.. The author is or has been a consultant on electric market reform and transmission issues for British National Grid Company, General Public Utilities Corporation (working with the "Supporting" PJM Companies), Duquesne Light Company, Electricity Corporation of New Zealand, National Independent Energy Producers, New York Power Pool, New York Utilities Collaborative, Niagara Mohawk Corporation, PJM Office of Interconnection, San Diego Gas & Electric Corporation, Trans Power of New Zealand, Williams Energy Group, and Wisconsin Electric Power Company. The views presented here are not necessarily attributable to any of those mentioned, and any remaining errors are solely the responsibility of the author.

market in England and Wales with the National Grid Company.[2] Although some elements of this monopoly approach may be found in any practical framework for transmission investment, it is not the focus here.

An alternative approach would lean more in the direction of market mechanisms with tradeable transmission property rights.[3] A workable system that provides an equivalent to transmission property rights could be used by market participants to match the actual flow of power or traded in a secondary market. Acquisition of these transmission benefits would provide a market incentive for transmission investment. This approach stands behind policy proposals at the Federal Energy Regulatory Commission (FERC)[4] and influences market restructuring developments around the world. A central problem is in defining a workable notion of property rights for transmission.

A pool-based, short-term electricity market coordinated by a system operator provides a foundation for building a system that includes tradeable transmission property rights in the form of transmission congestion contracts. Coordination through the system operator is unavoidable, and spot-market locational prices define the opportunity costs of transmission that would determine the market value of the transmission rights without requiring physical trading and without restricting the actual use of the system. In this setting, these transmission congestion contracts are equivalent to perfectly tradeable physical rights. Hence, this organization of the market defines a context where it would be possible to rely more on market forces, partly if not completely, to drive transmission expansion. Here we review the main ideas and summarize the requirements of a transmission investment approach with maximum reliance on the competitive market. An outline of the structure of the spot market coordinated by the system operator and supported by a system of transmission congestion contracts sets the stage for an overview of the transmission investment incentives.

ECONOMICS OF A COMPETITIVE ELECTRICITY MARKET

A general framework that encompasses the essential economics of electricity markets provides a point of reference for evaluating market design elements. This framework can be applied to the analysis of the institutional elements needed to support transmission pricing and investment.

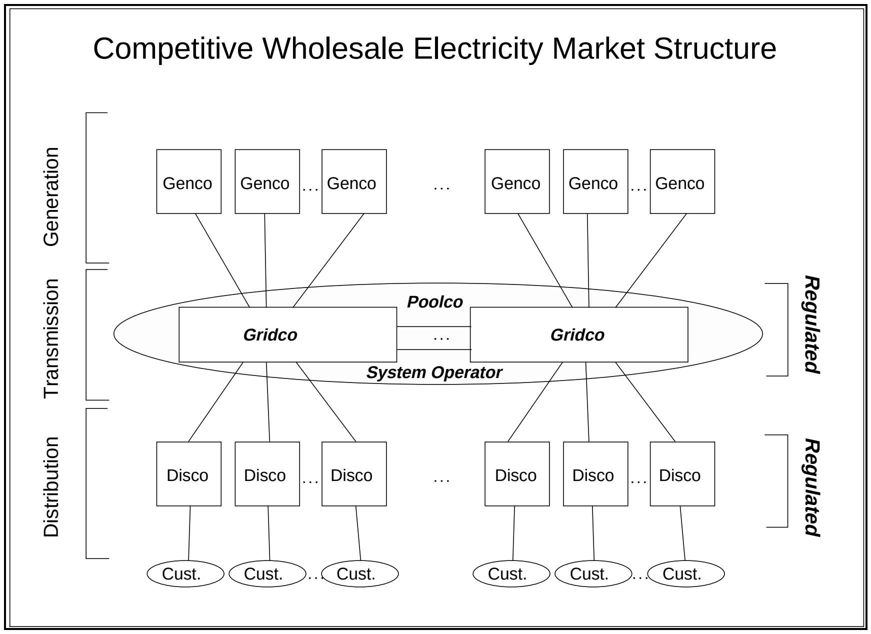
---

[2] David M. Newbery, "Privatization and Liberalization of Network Utilities," European Economic Review, Vol. 41, 1997, pp. 357-383.

[3] James B. Bushnell and Steven E. Stoft, "Grid Investment: Can a Market Do the Job?", The Electricity Journal, January-February, 1996, pp. 74-79. Eric C. Woychik, "Competition in Transmission: It's Coming Sooner or Later," The Electricity Journal, June 1996, pp. 46-58.

[4] Federal Energy Regulatory Commission, "Capacity Reservation Open Access Transmission Tariffs," Notice of Proposed Rulemaking, RM96-11-000, Washington DC, April 24, 1996.

Competitive Market Design



Restructuring of electricity markets typically emphasizes functional unbundling of the vertically integrated system. The usual separation into generation, transmission, and distribution is insufficient. In an electricity market, the transmission wires and the pool dispatch are distinct essential facilities. The special conditions in the electricity system stand as barriers to an efficient, large-scale bilateral market in electricity. A pool-based market coordinated by a system operator helps overcome these barriers.

Reliable operation is a central requirement and constraint for any electricity system. Given the strong and complex interactions in electric networks, current technology with a free-flowing transmission grid dictates the need for a system operator that coordinates use of the transmission system.[5] Control of transmission usage means control of dispatch, which is the principal or only means of adjusting the use of the network. Hence, open access to the transmission grid means open access to the dispatch as well. There must be a system operator coordinating use of the transmission system. That this system operator should also be independent of the existing electric utilities and other market participants is attractive in its simplicity in achieving equal treatment of all market participants. The independent system operator provides an essential service, but does not be compete in the energy market. In the analysis of electricity markets, therefore, a key focus is the design of the interaction between transmission and dispatch, both procedures and pricing, to support a competitive market.

To provide an overview of the operation of an efficient, competitive wholesale electricity market, it is natural to distinguish between the short-run operations coordinated by the system operator and long-run decisions that include investment and contracting. Under the competitive market assumption, market participants are price takers and include the generators and eligible customers. For this discussion, distributors are included as customers in the

---

[5] For further details on power flows and coordination, see Scott M. Harvey, William W. Hogan and Susan L. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1997.

wholesale market, operating at arm's length from generators. The system is much simpler in the very short run when it is possible to give meaningful definition to concepts such as opportunity cost. Once the short-run economics are established, the long-run requirements become more transparent. Close attention to the connection between short- and long-run decisions isolates the special 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 investment decisions have been made. 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 largely 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? The costs associated with delivery of electric energy include many services. The direct fuel cost of generation is only one component. In analyses of energy pricing, there is no uniformity in the treatment of these other, ancillary services.[6] The typical approach formulates an explicit model approximating the full electricity system, computing both a dispatch solution and associated prices for the explicit variables. Everything that is not explicit is treated as an ancillary service, for which the assumption must be that the services will be provided and charged for in some way other than through the explicit prices in the model. Given the complexity of the real electric system, such approximations or simplifications are found in every model, and there is always a boundary between the explicit variables modeled and the implicit variables that will be treated as separate ancillary services. Payment for the ancillary services may be through an average cost uplift applied to all loads. Development of a full description of the interactions with ancillary services is beyond the scope of the present discussion. For the sake of the present discussion, focus on real power and assume that further unbundling would go beyond the point of diminishing returns in the short-run market.

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. The collection of generator costs stacks 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 that are sensitive to price, and higher prices produce

---

[6] E Hirst and B. Kirby, “Creating Competitive Markets for Ancillary Services,” ORNL/CON-448, Oak Ridge National Laboratory, OakRidge, TN, October 1997.

lower demands. Generators and customers do not act unilaterally; they provide information to the dispatcher to be used in a decision process that will determine which plants will run at any given half hour. Power pools 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. The system operator controls operation of the system to achieve the efficient match of supply and demand.

This efficient central dispatch can be made compatible with the market outcome. The fundamental principle 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. However, marginal cost implicitly determines the least-cost dispatch, and marginal cost is the standard determinant of competitive market pricing.

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. Typically, 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. 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 would accept to run the plant in the given half hour. And these bids serve as the guide for 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 determining 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. Hence, with enough competitors and no collusion, the short-run central dispatch market model can elicit bids from buyers and sellers. The system operator can treat these bids as the supply and demand 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 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 simple settlement process.

Transmission Congestion. This overview of the short-run market model is by now familiar and found in operation in many countries. However, this introductory overview conceals a critical detail that would be relevant for transmission pricing. Not 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.

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. 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 different marginal costs and different prices, depending on location, but the basic market model and its operation in the short-run would be preserved.

Transmission congestion has a related effect. 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. Power will flow over the transmission line from the low cost to the high cost location. 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 could be used, and some of the cheap plants would be "constrained off." In this case, the demand would be met by higher cost plants that, absent the constraint, would not run, but due to transmission congestion would be then "constrained on." The marginal cost in the two locations differs because of transmission congestion. The marginal cost of power at the low cost location is no greater than the cost of the cheapest constrained-off plant; otherwise the plant would run. Similarly, the marginal cost at the high cost location 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.[7]

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 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 locations. 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

---

[7] Losses are not always negligible and could be treated explicitly. However, it would add nothing to the discussion here.

by more than 100% across different locations.

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 includes a set of prices, one for each location. Economic dispatch would still be the least-cost equilibrium. Generators would still bid as before, with the bid understood to be the minimum acceptable price at their location. Customers would bid also, with dispatchable demand and the bid setting the maximum price that would be paid at the customer's location. The economic dispatch process would produce the corresponding prices at each location, incorporating the combined effect of generation, losses and congestion. In terms of their own supply and demand, everyone would see a single price, which is the locational marginal cost based 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. However, the corresponding transmission price would be the difference between the prices at the two locations.

This same framework lends itself easily to extensions including bilateral transactions. If market participants wish to schedule transmission between two locations, the opportunity cost of the transmission is just this transmission price of the difference between spot prices at the two locations. This short-run transmission usage pricing, therefore, is efficient and non-discriminatory. In addition, the same principles could apply in a multi-settlement framework, with day-ahead scheduling and real-time dispatch. These extensions could be important in practice, but would not fundamentally change the outline of the structure of electricity markets.

This short-run competitive market with bidding and centralized dispatch is consistent with economic 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 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 added to the cost of ancillary services and applied to all load. The system operator coordinates the dispatch and provides the information for settlement payments, with regulatory oversight to guarantee comparable service through open access to the pool run by the system operator through a bid-based economic dispatch.

With efficient pricing, users have the incentive to respond to the requirements of reliable operation. Absent such price incentives, the system operator would be required to restrict choice and limit the market, in order to give the system operator enough control to counteract the perverse incentives that would be created by prices that did not reflect the marginal costs of dispatch. A competitive market with choice and customer flexibility depends on getting the usage pricing right.[8]

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.

Even without transmission congestion constraints, the spot market price can be volatile. 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 a 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. It is not even in the interest of the generators or the customers to restrict their dispatch and forego the benefits of the most economic use of the available generation. The short-term dispatch decisions by the system operator 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

---

[8] William W. Hogan, "Independent System Operator: Pricing and Flexibility in a Competitive Electricity Market," Harvard-Japan Project on Energy and the Environment, Harvard University, March 1998.

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 the deviation from 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 the system operator, 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. Familiar in other commodity industries, this is known in the electricity market as a "contract for differences."

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 system operator other than for the continuing short-run market transactions. But through the interaction with system operator, 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, system operator need take no notice of the contracts, and have no knowledge of the terms. Such contracts for differences have become common in restructured electricity markets.

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.

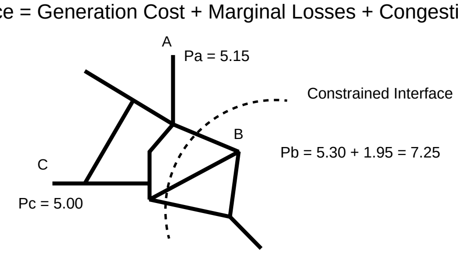
Transmission congestion in the short-run market raises another related and significant matter for the system operator. In the presence of congestion, revenues collected from customers will substantially exceed the payments to generators. The difference is the congestion rent that accrues because of constraints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

A convenient solution for both problems – providing a price hedge against locational congestion differentials and removing the adverse incentive for the system operator – is to re-distribute the congestion revenue through a system of long-run transmission congestion 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 congestion contract that provides compensation for differences in prices, in this case for differences in the congestion costs between different locations across the network.

As illustrated in the figure, a transmission congestion contract defines a hedge for differences in locational prices. Everyone who uses the transmission system pays according to the spot-market locational prices. For example, generation at A receives 5.15 cents. Load at B pays 7.25 cents, with the price including marginal losses and congestion. Transmission between C and B would pay 2.25 cents. Included in this transmission charge would be a congestion payment of 1.95 cents. A transmission congestion contract between C and B would result in a payment of the 1.95 cents to the holder of the contract.[9] Hence, if the participant actually transported the power, the transmission congestion contract would just balance

NETWORK TRANSMISSION CONGESTION CONTRACTS

Bus Price = Generation Cost + Marginal Losses + Congestion Costs



Pcb = Pb - Pc = Marginal Losses + Congestion Costs = 0.3 + 1.95 = 2.25

- DEFINE TRANSMISSION CONGESTION CONTRACTS BETWEEN LOCATIONS.
- FOR SIMPLICITY, TREAT LOSSES AS OPERATING COSTS.
- RECEIVE CONGESTION PAYMENTS FROM ACTUAL USERS; MAKE CONGESTION PAYMENTS TO HOLDERS OF CONGESTION CONTRACTS.
- TRANSMISSION CONGESTION CONTRACTS PROVIDE PROTECTION AGAINST CHANGING LOCATIONAL DIFFERENCES.

---

[9] For an expanded discussion of transmission congestion contracts, see Scott M. Harvey, William W. Hogan and Susan L. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1997.

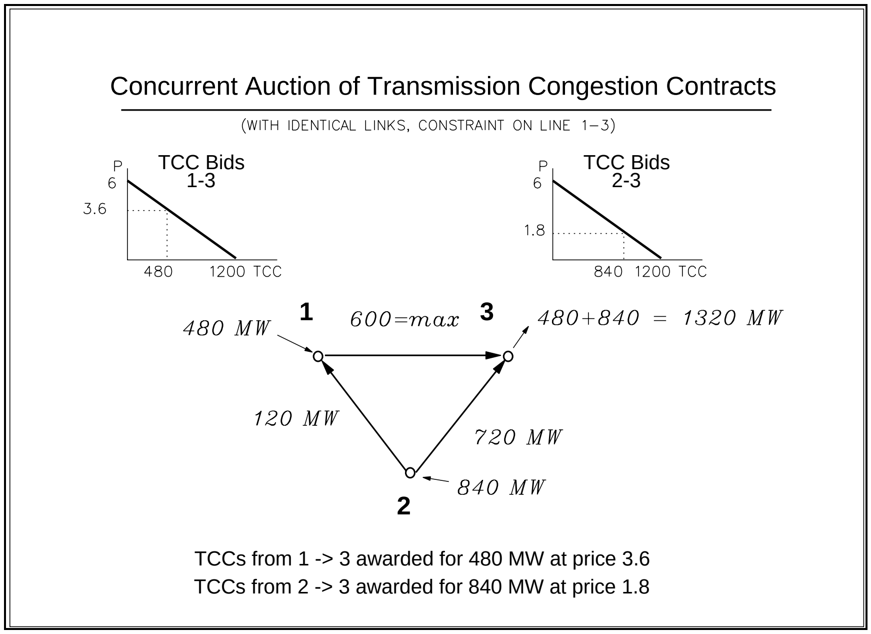
the congestion charge for the quantity covered by the contract. And if the holder of the transmission congestion contract does not transport power, the result is the same as selling the transmission right to the actual user.

The transmission congestion contract for compensation would exist for a particular quantity between two locations. A generator at location C might obtain a transmission congestion contract for 100 MW between the generator's location and the customer at location B. The right provide by the contract would not be for specific movement of power but rather for payment of the congestion rental. Hence, ignoring losses, if a transmission constraint caused prices to rise to 6.95 cents at the customer's location, but remain at five cents at the generator's location, the 1.95 cents difference would be the congestion rental. The customer would pay the pool 6.95 cents for the power actually taken. The pool would in turn pay the generator five cents for the power supplied in the short-run market. As the holder of the transmission congestion contract, the generator would receive 1.95 cents for each of the 100 MW covered under the transmission congestion contract. The generation and the load would see the right incentives on the margin. The transmission congestion contract 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 congestion contract, the generator would have no revenue to compensate the customer for the difference in the prices at their two locations. The transmission congestion contract completes the package described as a "contract network."

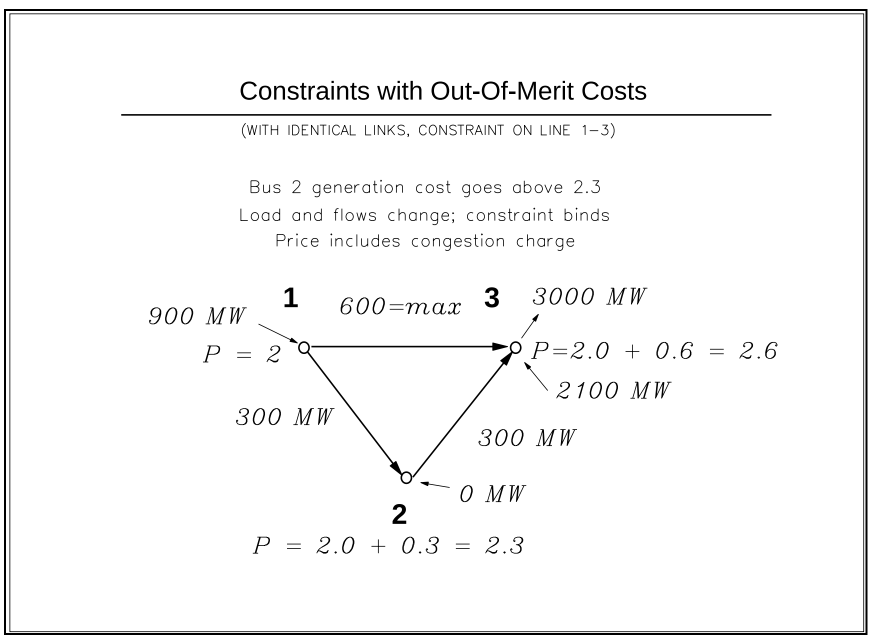
Note that the transmission congestion contract is defined as a point-to-point comparison of locational prices. There is no reference to the individual transmission links or the path by which the power may move between two points. This is an important distinction that separates the transmission congestion contract from "link-based" rights that otherwise confront difficult anomalies and perverse incentives. For example, it is entirely possible to construct examples of valuable transmission links that have no net power flow, making them appear worthless for many definitions of link-based rights. But this condition would not impact the definition or the value of transmission congestion contracts.

There are alternative interpretations of contract network rights defined as transmission congestion contracts, with various advantages for implementation and interpretation. As above, the contract could be defined as the right to the difference in congestion costs; the contract holder receives the difference in congestion costs between two buses for a fixed quantity of power. This is equivalent to the ability to purchase at a distant location; the contract holder at B can purchase a fixed quantity of power at B but pay the price applicable at a distant location, C. Or we could describe the transmission congestion contract as providing for dispatch with no congestion payment; the transmission congestion contract holder can inject and remove a fixed quantity of power without any congestion payment.

The allocation of transmission congestion contracts among market participants could arise in many ways. For a market in transition during a restructuring process, the allocation might be intended to reflect explicit or implicit historical rights. For new transmission, there could be a negotiation process to select and award one of the many possible combinations of



feasible contracts. In addition, some or all of the transmission congestion contracts for the grid could be defined and awarded through an open auction. The collective bids would define demand schedules for contracts. The concurrent auction would respect the transmission system constraints to assure simultaneous feasibility. For example, the accompanying figure illustrates a hypothetical auction with two demand curves for transmission congestion contracts on a simple network with three locations. The optimal award that maximizes the benefit[10] of the contracts includes 480 MW from bus 1 to bus 3 and 840 MW from bus 2 to bus 3.



With spot market locational prices, the transmission congestion contracts provide price protection. Even with changing load patterns, the congestion revenues collected by the system operator will be at least enough to cover the obligations for all the contracts. Consider a typical dispatch with the transmission constraint binding. Suppose the solution is as in the accompanying figure, with a total load of 3000 MW at

10 The benefit is defined as the consumer welfare, the area under the demand curves. Note this is not the same as the value of the transmission congestion contracts at the market clearing price. The model is the simplified DC-Load model.

bus 3. The generation at bus 3 accounts for 2100 MW, with the balance of 900 MW coming from bus 1. The assumed market clearing prices at buses 1 and 3 are 2 cents and 2.6 cents, respectively. From these prices, and the knowledge that there is only one binding transmission constraint, we can compute the implied equilibrium prices at all other buses.[11] In this case bus 2 has a price of 2.3 cents. Note that the market equilibrium for this dispatch is quite different from the pattern of flows envisioned in the allocation of the transmission congestion contracts. However, the different pattern of use, and the associated spot prices, present no difficulty for the system operator. As along as the allocation of the transmission congestion contracts defines a set of inputs and outputs that would be simultaneously feasible, then the revenue collected from the spot prices for congestion will be sufficient to compensate the holders of the contracts for the obligations at the same set of spot prices.[12]

The resulting calculation is illustrated in the accompanying table. The payment obligations for the inputs and outputs at the spot prices define a set of net payments for the actual loads. The payment obligations for the transmission congestion contracts equal the net revenues.

System Operator Revenues			
	Quantity	Price	$
Bus 1	900	2	($1,800)
Bus 2	0	2.3	$0
Bus 3	2100	2.6	($5,460)
Bus 3	-3000	2.6	$7,800
TCC 1-3	480	0.6	($288)
TCC 2-3	840	0.3	($252)
Net Total			$0

In effect, the system operator collects the congestion rents from the users of the system and distributes these same rents to the holders of the transmission congestion contracts. In this case, with the simplified assumptions and the same binding constraint, the payments exactly balance. In general, excess congestion rents may remain after paying all obligations under the transmission congestion contracts. These excess rentals should not remain with the system

[11] William W. Hogan E. Grant Read, and Brendan J. Ring, "Using Mathematical Programming for Electricity Spot Pricing," International Transactions in Operational Research, Vol. 3, No. 3/4, 1996, pp. 209-221.

[12] For a further discussion of power flows and spot pricing, see William W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, 1992, p. 214.

operator, but could be distributed according to some sharing formula to those paying the fixed costs of the existing grid or along with the payments under transmission congestion contracts.[13] Many variants are possible, allowing great flexibility in developing and trading contracts. The contract network can allow great commercial flexibility while respecting the reality of the actual network in determining the locational prices.

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 would be far less obvious. There will be many possible transmission combinations between different locations. There is no single definition of transmission grid capacity, and it is only meaningful to ask if the configuration of allocated transmission flows is feasible. However, the net result would be 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 system operator coordinates the short-run market to provide economic dispatch. The system operator collects and pays according to the short-run marginal price at each location, and the system operator distributes the congestion rentals to the holders of transmission congestion contracts. Generators and customers make separate bilateral arrangements for generation contracts. Unlike with the generation contracts, the system operator's participation in coordinating administration of the transmission congestion contracts is necessary because of the network interactions, which make it impossible to link specific customers paying congestion costs with specific customer receiving congestion compensation. If a simple feasibility test is imposed on the transmission congestion contracts awarded to customers, the aggregate congestion payments received by the system operator will fund the congestion payment obligations under the transmission congestion contracts. Still, the congestion prices paid and received will be highly variable and load dependent. Only the system operator will have the necessary information to determine these changing prices, but the information will be readily available embedded in the spot market locational prices. The transmission congestion contracts define payment obligations that guarantee protection from changes in the congestion rentals.

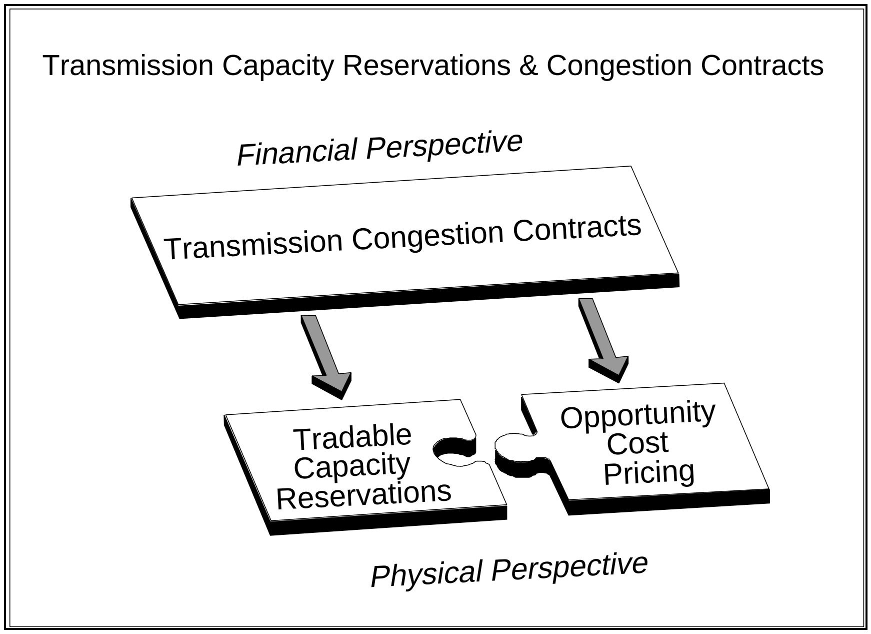
Equivalent Transmission Interpretations. The transmission congestion contract can be recognized as equivalent to an advantageous form of point-to-point “physical” transmission right. Were it possible to define usage of the transmission system in terms of physical rights, it would be desirable that these rights have two features. First, they could not be withheld from the market to prevent others from using the transmission grid. Second, they would be perfectly tradable in a secondary market that would support full reconfiguration of the patterns of network use, at no transaction cost. This is impossible with any known system of physical transmission rights that parcel up the transmission grid. However, in a competitive electricity market with a bid-based economic dispatch, transmission congestion contracts are equivalent to just such perfectly tradable transmission rights.

With a competitive market and tradable physical transmission capacity reservations,

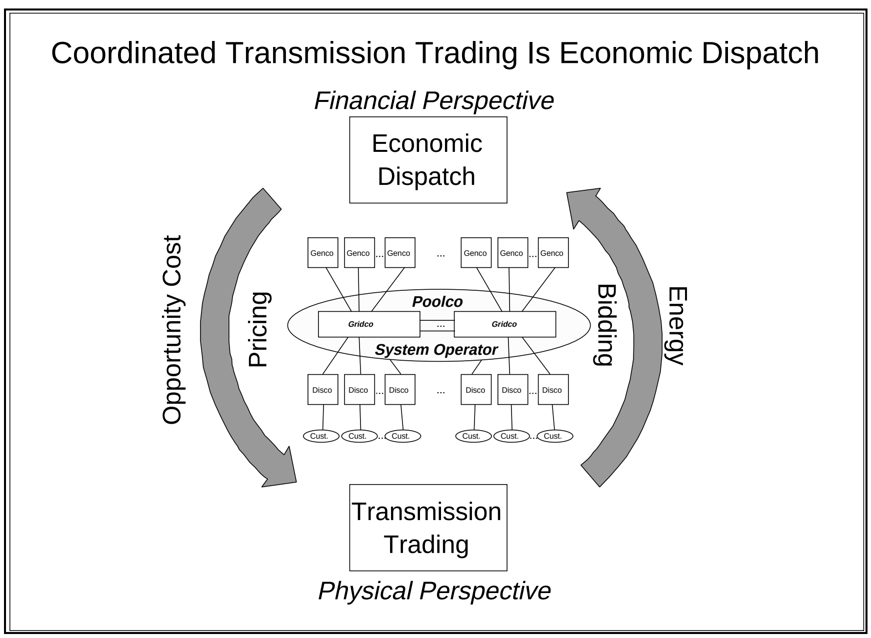
---

[13] The allocation of excess congestion rentals is part of the specification of the transmission congestion contracts. Although important in practice, it would have little impact on the present analysis.

any use of the system not matched by a reservation would be settled at opportunity cost prices determined by the final dispatch or actual use of the system. Trading of physical transmission capacity reservations must be coordinated through the system operator. Under competitive conditions, the price that would be paid for transmission capacity would be the opportunity cost, which is the same as the difference in the locational prices. If there were no transaction costs, therefore, this physical perspective would be indistinguishable from the financial perspective of transmission congestion contracts based on the same opportunity cost.



The physical perspective may be more intuitive. The financial perspective with transmission congestion contracts as perfectly tradeable rights is easier to implement and has lower transaction costs. Hence, we can describe transmission congestion contracts either as financial contracts for congestion rents or as perfectly tradable point-to-point physical transmission rights. If the capacity for transmission congestion contracts has been fully allocated, then the system operator will be simply a conduit for the distribution of the congestion rentals. The operator would have no incentive to increase congestion rentals: any increase in congestion payments would flow only to the holders of the transmission congestion contracts. The problem of supervising the dispatch monopoly would be greatly reduced. And through a combination of generation contracts and transmission congestion

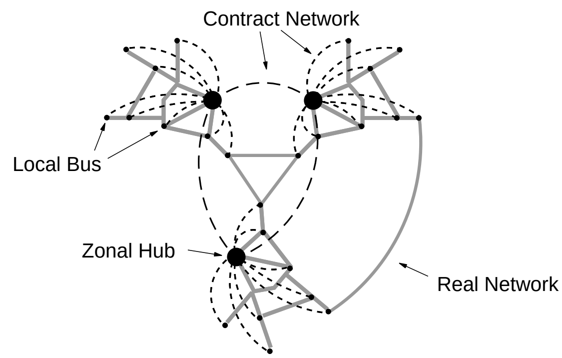


contracts, participants in the electricity market can arrange price hedges that could provide the economic equivalent of a long-term contract for specific power delivered to a specific customer.

Further to the application of these ideas, locational marginal cost pricing lends itself to a natural decomposition. For example, even with loops in a network, market information could be transformed easily into a hub-and-spoke framework with locational price differences on a spoke defining the cost of moving to and from the local hub, and then between hubs. This would simplify without distorting the locational prices. As shown in the figure, a contract network could develop that would be different from the real network, without affecting the meaning or interpretation of the locational prices.[14]

Contract Network Connects with Real Network

Determine Locational Prices for Real Network; Implement Transmission Congestion Contracts and Trading on Contract Network



With market hubs, the participants would see the simplification of having a few locations that capture most of the price differences of long-distance transmission. Contracts could develop relative to the hubs.[15] The rest of the sometimes important difference in locational prices would appear in the cost of moving power to and from the local hub. Commercial connections in the network could follow a configuration convenient for contracting and trading. The separation of physical and financial flows would allow this flexibility.

The creation or elimination of hubs would require no intervention by regulators or the system operator. New hubs could arise as the market requires, or disappear when not important. A hub is simply a special node within a zone. The system operator still would work with the locational prices, but the market would decide on the degree of simplification needed. However, everyone would still be responsible for the opportunity cost of moving power to and from the local hub. There would be locational prices and this would avoid the substantial incentive

---

[14] For further details on long-run transmission congestion contracts, see Scott M. Harvey, William W. Hogan and Susan L. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1997.

[15] For a similar argument, see S. Walton and R. Tabors, "Zonal Transmission Pricing: Methodology and Preliminary Results form the WSCC," Electricity Journal, November 1996, pp. 40.

problems of averaging prices. The hub-and-spoke approach appears to give most of the benefits attributed to zones without the costs, and it implies that the system operator works within a locational pricing framework.

A version of this hub-and-spoke system appeared in the discussions among market participants in the Pennsylvania-New jersey-Maryland (PJM) system in January 1998. Although the market participants could create hubs without the participation of the system operator, there was a popular request that the system operator identify and post hub prices. In order to reduce the price volatility that might be present in selecting a single location as a hub, the system operator responded to the request to create hubs consisting of a fixed-weight average of a number of underlying locations. Hence, purchases and sales at the hub are equivalent to a portfolio of purchases and sales at the underlying locations, with the portfolio composition in terms of the fixed weights. Likewise, transmission between any location and the hub is the equivalent of transmission between the given location and the portfolio of locations that make up the hub. The PJM system operator began posting prices for such hubs in the preliminary price reporting system that began operation in February 1998.[16] This system became operational in the beginning of April 1998. If the market participants find these hubs convenient, then contracts would be written using the composite hub prices as references, and futures contract could be constructed relative to the same spot-market hub prices. However, nothing limits the markets to the hubs selected by the system operator, and anyone is free to utilize another hub and organize trades relative to this hub price.

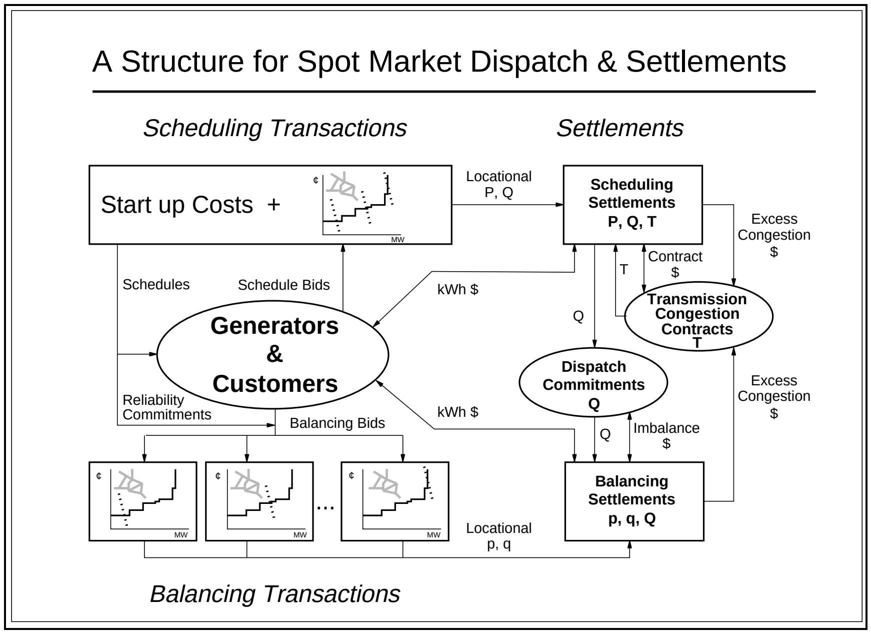
Scheduling and Balancing. Implementation of the short-run dispatch market could take many forms. In principle, the coordinated dispatch might be left to only the final hour, with all other commitments and trades developed through decentralized transactions, notifying the system operator of the schedules only in time to complete the final balancing. At the other extreme, with long lead times for changing the configuration of generation or the patterns of loads, it might be preferred to include unit commitment decisions over weeks or even longer periods.

In practice, most countries or regions adopt or recommended a procedure that falls somewhere in between these alternatives. The system operator accepts bids and nominations for scheduled dispatch, say for a day ahead, and determines an appropriate market clearing equilibrium and associated payment settlements. This schedule then defines a set of commitments for delivering and taking power in the short-run dispatch. In the event, the actual dispatch will differ from the scheduled commitments, and appropriate balancing settlements would be arranged.

These connected scheduling and balancing settlements present no major difficulties, but there are a few points that need clarification to maintain consistent payments and incentives. The basic schematic appears in the accompanying figure. Participants in the market submit scheduling bids for the day ahead for both supply and demand. These bids may include start up

---

[16] The Internet address is "www.pjm.com/ferc/filings/19971231/lmp_info.html".



costs, ramping rates and any of range of plant and load characteristics.[17] The system operator utilizes all this information to define a least-cost dispatch over the day that matches scheduled load and generation. The results is a set of market clearing prices (P) and quantities (Q) that define the schedule. In addition, the system operator arranges for any necessary reliability commitments, such as for spinning and standby reserves to meet any expected deviations in load from the day ahead schedules.

The equilibrium prices and quantities differ by location and represent immediate firm commitments. Based on these commitments, payments would be made through a settlements process. Conceptually these settlements could occur the minute the schedule is determined; in practice, the settlements would occur after the fact. However, to preserve the consistency of the incentives and payments, the settlements must be based on the scheduling prices and quantities. These settlements will include payments under long-term transmission congestion contracts shown by the symbol "T" in the schematic. Holders of transmission congestion contracts would receive or pay the appropriate amounts of congestion cost differentials between locations for the contracted quantity under the transmission congestion contract. Under some fixed sharing rule, the transmission congestion contract holders, or the presumptive parties responsible for paying the fixed charges of the transmission grid, would also share in any excess congestion payments after settlement of the locational differences.

The schedule and the associated dispatch commitments (Q) would provide the reference point for the actual dispatch. In principle, bids in the scheduling market could be revised to create balancing bids for increments and decrements against the dispatch commitments. Again the system operator would find the least-cost dispatch, hour by hour, based on the actual conditions and the final balancing bids. The result would be an actual dispatch with associated equilibrium prices (p) and quantities (q). These prices and quantities would differ to a degree from the schedules, with the "imbalances" (q-Q) settled at the balancing price of "p". In this balancing settlements, the dispatch commitments are conceptually similar to the transmission

[17] NGC Settlements Limited, An Introduction to the Pool Rules, Electricity Pool of England and Wales, April 1993.

congestion contracts that apply in the scheduling settlements. And just as for the transmission congestion contracts, after settling all the imbalances at the market clearing price "p", there could be some excess congestion payments that would be disbursed to the users and not kept by the system operator. In part, this excess would be used to reduce user payments for overhead and ancillary services, not shown in the schematic, or rebated again to those who bear the fixed costs of the grid.

The precise treatment of the excess congestion rentals is not important, other than to disburse them to the users and not the system operator in a way that creates no incentives for an inefficient dispatch. What is important, however, is to settle the scheduling markets and balancing markets with their own internally consistent prices (P,p) and quantities (Q,q). This is required, for example, to avoid the initial over-the-day gaming problems created in England and Wales by settling based on scheduled prices (P) and actual quantities (q).

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 there are 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 generators 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. The system operator makes no guarantees as to the price at the location. The system operator 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. The complexity and reach of these contracts would be 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 for the generators to defer investing in new plant until sufficient long-term contracts with customers can be arranged to cover a sufficient portion of the required investment. The generation contracts could be with one or more customers and might involve a mix of fixed charges coupled with the obligations to compensate for price differences relative to the spot-market price. But the customer and generator would ultimately buy and sell power at their location at the half-hourly price.

If either party expects significant transmission congestion, then a transmission congestion contract would be indicated. If transmission congestion contracts are for sale between the two points, then a contract can be obtained from the holder(s) of existing rights. Or new investment can create new capacity that would support additional transmission congestion contracts. The system operator would participate in the process only to verify that the newly created transmission congestion contracts would be feasible and consistent with the obligation to preserve any existing set of transmission congestion contracts 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 set of transmission congestion contracts for compensation–while protecting the existing rights–and the test is independent of the actual loads that may develop. Hence, incremental investments in the grid would be possible anywhere without requiring that everyone connected to the grid participate in the negotiations or agree to the allocation of the new transmission congestion contracts.

Evaluating the Market Design

The basic outline of a competitive market preserves certain features of the traditional system, such as reliability and economic dispatch. However, it changes other elements that are either unnecessary with opportunity cost pricing or are left to the choices of the market participants. A review of the prominent features sets the stage for a further discussion of transmission pricing and expansion issues.

Security-Constrained Economic Dispatch. Security-constrained economic dispatch is preserved in the competitive market based on coordinated dispatch with customer bids. The system operator 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, implying 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 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 the system dispatch and the lines. 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 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 design, 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, arises in the non-competitive, regulated portion of the market. Here the obligation to serve takes on a special and restricted meaning 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 line owners, the obligation is to connect customers and negotiate in good faith any requested enhancements of "wires" capacity with appropriate creation of new transmission congestion contracts. The regulated segments have no obligation to review or be responsible for resulting investments or their subsequent economic consequences.

Security Concerns and Capacity Reserves. Secure economic dispatch coordinated through the system operator 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 and the cost of losses. In this sense, evolution of the grid would be 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. 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.

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 in the short run. 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 long-run required 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 system operator.

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 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. Given the complexities of network interactions, such analyses imply requirements for long-run forecasts of both the timing and the location of supply and demand for power.

Someone may prepare such forecasts, 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 system operator 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. In the absence of market failure, there is no need for the regulator to prepare or endorse forecasts of future opportunity costs. These 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 contracts or invest to create new contracts 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.

Two-Part Prices. There are substantial economies of scale in transmission expansion.[18] The total

---

[18] Ross Baldick and Edward Kahn, "Network Costs and the Regulation of Wholesale Competition in Electric Power," Journal of Regulatory Economics, Vol. 5, 1993, pp. 37-384. The problems created by economies of scale are addressed in the context of transmission pricing in the New Zealand by E. G. Read and D. P. M. Sell, "Pricing and Operation of Transmission Services: Short Run Aspects," Report to Trans Power, Canterbury University and Arthur Young, New Zealand, October 1988; E. G. Read, "Pricing of Transmission Services: Long Run Aspects," Report to Trans Power, Canterbury University, New Zealand, October 1988; E. G. Read and D. P. M. Sell, "A Framework for Transmission Pricing," Report to Trans Power, Arthur Young, New Zealand, December 1988.

congestion revenue from transmission usage charges, therefore, generally would be less than the cost of the grid. This means that transmission pricing cannot rely solely on congestion payments to recover the full costs of the existing transmission grid. The incentive of avoiding future congestion payments may prompt transmission investment, but the ex post congestion payments could be lower after the investment. Hence, the competitive market model outline here relies on pricing according to two-part prices including a fixed and variable charge. This approach contrasts with the alternative of a one-part transmission charge with a single price based on usage. Application of two-part prices is especially apt in the case of transmission, where it is 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 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 price. 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.

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 workably 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 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 congestion contracts. Although in principle this too could evolve as an arms-length market without participation by the system operator, other information economies may require participation by the transmission entities. In particular, the physical "terms-of-trade" between two transmission congestion contracts are far from obvious. It may well be that 100 MW between locations A and B could be exchanged for 200 MW between C and D, or 150 MW between E and F. The network interactions are nonlinear, and only the central dispatch operator has the information required to know which exchanges are feasible.

Other Features. The list of additional features could be expanded much further, and the pricing model compared. Problems of free-riders remain in any system that includes the inescapable economies of scale of transmission expansion. If there is market power in generation or load, then the dispatch and transmission pricing model can be manipulated by the dominant player; however, the problems here are no more severe than in any other system that includes market power. To a limited extent, the use of locational prices and transmission congestion contracts can help mitigate market power and reduce incentives for inefficient investment. In addition, the competitive model does nothing in itself to internalize environmental externalities or facilitate other social policies.

ALLOCATING TRANSMISSION BENEFITS AND COSTS

The economies of scale and scope in transmission construction present a number of issues in long-run transmission pricing. Transmission investments and upgrades tend to come in large increments that can have a significant effect on market prices. The facilities are capital intensive, require long investment lead times, and are long lived. Any change in the configuration of the transmission grid can alter the extensive network interactions, thereby affecting virtually all users of the transmission grid. Furthermore, most of the costs for the existing transmission system are sunk. Hence, there is a marked difference between the problems of determining the allocation of the costs for the existing system and the task of defining the allocation for system expansion.

The incentive for transmission investment would come from those who did not already have transmission congestion contracts. These market participants would be relying on the spot market. When the anticipated congestion cost and losses become large enough, there would be an economic motivation for investing in transmission. However, the motivation for the investment assumes, at a minimum, that those making the investment would then reduce the cost of the losses and avoid the corresponding congestion payments.

The discussion of the economics of competitive electricity markets emphasizes the importance of defining and allocating benefits of the transmission grid. If there is no method to allocate the benefits of transmission, the strong network interactions exacerbate the difficulty of allocating costs. Without some form of transmission rights, there could be substantial incentives to shift costs to others (i.e., free-riding), it would be difficult to define a standard of performance for transmission providers, and the normal operation of market principles would founder because of the lack of property rights. By contrast, where it is possible to define the equivalent of property rights, free-riding and these difficult allocation problems could be substantially reduced, if not eliminated.

Consider the allocation of the costs for short radial connections from plant to a nearby transmission grid. In the simplest of cases, when connecting a single generator or a single customer, the benefits of the connection seem clear. The benefits would accrue to the facility connected to the grid, and the costs would naturally be allocated to the same facility. The close match between benefits and costs would allow for investment charges that would be consistent with operation of a market. It is for this reason that transmission pricing often treats such investments separately.

When we turn from simple radial connections to the main transmission grid, however, the obvious allocation of costs and benefits is no longer available. This is closely related to the complicating fact of network interactions that make it difficult or impossible to define the total capacity of a network. The simple reality is that in any sufficiently interconnected network, it is not possible to define the capacity of the network without also defining the pattern of usage. Except for truly radial systems, transmission congestion can change the capacity of the grid depending on the configuration of load and generation.

This difficulty in defining the capacity of the grid without specifying the pattern of use does not mean that no definition of capacity or property rights would be available. For example, it is always possible to test easily to determine if a particular pattern of usage of the transmission grid would be within the security constrained dispatch limits dictated by the requirements of reliable operation. The system operator makes just such judgments on a regular basis, with an affirmative conclusion implicit in every dispatch. Hence, it would be legitimate to define any feasible pattern of system use as within the capacity of the system, and the associated input levels could be defined as transmission rights. With these hypothetical property rights, we would be closer to a method for allocating the benefits of transmission grid to support the allocation of the costs.

One difficulty of this approach in terms of physical property rights would be in the constant need to redefine and reconfigure transmission usage to conform to the actual load patterns. In effect, we would like those who have paid for transmission property rights to be able to sell those rights in a secondary market to support the actual dispatch. Were this possible, many of the most perverse incentives and logical puzzles in transmission pricing would be removed. With well-defined, tradable property rights, the transmission regime could operate more according to market principles. However, a literal system of physical rights traded to match actual use would be impractical as a means to control system dispatch.

This is where the transmission congestion contracts fit into the long-run picture. A transmission congestion contract provides payment equal to the congestion price differential between two locations. These financial contracts are equivalent to perfectly tradable transmission rights. Hence, these transmission congestion contracts can be the embodiment of the benefits of transmission investment. Ownership of the transmission congestion contract can be defined as the benefit that comes with paying for transmission investment. The transmission congestion contract provides a perfect hedge against changes in the transmission usage charges arising in the form of congestion costs. The transmission congestion contract internalizes the complicated network interactions, and allows the market to operate with simple point-to-point arrangements that act like property rights.

In some cases, the transmission congestion contracts alone could support market decisions for transmission investment. Faced with transmission congestion and persistent differences in locational prices, market participants could come forward to negotiate a contract for transmission expansion. In exchange for a long-term agreement to pay the costs of the expansion, they would receive a set of transmission congestion contracts that would reflect the incremental transfer capability created by the investment. In general, there would be no unique set of potential transmission congestion contracts, as many patterns of increased flow could be accommodated by the incremental investment. However, the participants would agree on the preferred collection of new transmission congestion contracts, and the assignment of both the costs and the benefits would be established by contract before the investment was made. The system operator would participate to certify the feasibility of the incremental transmission congestion contracts in combination with any other existing rights. The transmission provider would build the transmission enhancement, the customer would pay the investment costs as dictated by the contract, and the system operator would administer the transmission congestion contracts providing the hedge against congestion costs.

The transmission congestion contracts could also provide the definition of the future obligations of the transmission provider. Absent some later change in the explicit contracts, the obligation of the transmission provider would be to maintain the feasibility of the collection of transmission congestion contracts. This could be achieved either by literally maintaining the wires and other facilities to preserve the feasibility of the power flows as specified in the contracts, or by assuming the financial responsibility to make up any payment deficits under the transmission congestion contracts. Hence, the set of transmission congestion contracts could serve as the cornerstone for an incentive pricing scheme for the transmission provider.

If there were no market power present, the value of the opportunity costs under the transmission congestion contracts would define the opportunity costs of any reduction of system capacity. This would set a framework for any payment obligations by the transmission provider. In the presence of market power in generation or load, including the important special case of a radial connection for a single entity, the transmission congestion contract obligation could be adapted to specify obligations of the transmission provider. For example, in case of the loss of a radial line, there would be no observed market clearing price at the plant to define the opportunity cost of the reduced capacity caused by a line outage. However, in the context of a spot market, a pattern of bids from previous periods might establish a workable estimate of the incremental cost of the plant. The difference relative to the grid price at the point of connection could be used as the estimate of the opportunity cost in the rare event of line outage. In this case, the opportunity cost estimate would serve as the implied obligation under the transmission congestion contract. The typical contract might include some limited periods or maximum payments to account for normal outages and maintenance, with any excess payments resting in whole or in part as the responsibility of the transmission provider.

In circumstances with multiple owners of the transmission grid and substantial parallel flows, there would be a requirement to define a sharing mechanism of responsibility among the transmission providers. In the case of a significantly interconnected system, the capacity on one part of the system could affect the capacity everywhere, and not all transmission providers would be able to control these impacts. This allocation of responsibility would be similar to the joint cost allocation problem, for which a number of practical solutions would be available.

This contract network approach with spot market locational prices and transmission congestion contracts is a "conforming" pricing proposal under the FERC criteria. The transmission grid owner receives embedded costs for the existing grid or investment costs for new increments. The payment for the existing grid or new grid investments can be cost based according to traditional regulation, or allowed to adjust to market prices. Those paying the regulated fixed charges for transmission receive the transmission congestion contracts.

The spot market participants pay locational prices which include transmission opportunity costs. The difference in locational spot prices defines the opportunity cost of short-term transmission usage. Everyone has access to use the grid and pays the transmission opportunity cost. The transmission congestion contract holders receive the congestion payments. Through the settlement payments, the system operator collects the congestion payments from the system users and disburses the congestion payments to the holders of the transmission congestion contracts. The system operator keeps nothing; the grid owner receives the regulated fixed charges; the transmission congestion contract holder receives compensation for out-of-merit dispatch costs; and the transmission user pays the true opportunity cost.

NETWORK EXTERNALITIES AND INVESTMENT INCENTIVES

Complete reliance on market incentives for transmission investment would be unlikely as a practical matter and is subject to a number of theoretical deficiencies. The tradeoff will be between imperfect markets and imperfect regulation. The existence of a viable system of long-term transmission congestion contracts changes the balance, however, to rely more on market forces.

Economies of Scale and Free-Rider Effects

The locational opportunity cost pricing system would not pay for the grid. The fixed charges of the grid reflect investments that, in principle, have been made to avoid paying the opportunity costs of congestion, losses, and more expensive generation. Typically, because of economies of scale and scope, the fixed charges would be greater than the congestion opportunity costs after the investment was made. It follows that fixed cost recovery for grid investments would be through a combination of access charges and long-term contracts. The transmission congestion contracts would be the embodiment of the long-term guarantee that would be obtained in exchange for the fixed charge payments. Transmission congestion contracts would provide a way to untangle the network and convert short-run opportunity cost prices into long-term transmission arrangements.

This happy resolution of the puzzle of transmission expansion and pricing through voluntary market forces alone is subject to at least two other important caveats. First, there still may be market failures even with the definition of a workable set of equivalent property rights. For example, with many small market participants, each benefiting a little from a large transmission investment, the temptation to free-ride on the economies of scale and scope may create a kind of prisoner's dilemma. Everyone would be better off sharing in the investment, but the temptation to free ride and avoid paying for the expense may overcome any ability to form a consortium or negotiate a contract. It may be that the investment could not go forward in a timely manner, at the right scale, or at all, without some entity that can mandate payment of the costs.[19] In this case, however, the task should be simplified by the ability to simultaneously allocate the benefits in the form of a share of the transmission congestion contracts. The market could take care of many, perhaps most, investments, and the regulatory option would be easier to implement when needed.

Second, operation of voluntary market forces would have little sway in the allocation of the costs for an existing transmission grid that already provides open access. The costs are sunk, and typically the sunk costs of the wires exceed the transmission congestion opportunity costs of using the grid. This is due, in large part, to the effects of the economies of scale. Hence, given the choice of paying the sunk costs but avoiding the congestion costs, versus avoiding the sunk costs while using the system and paying the continuing cost of congestion, most users would prefer the latter. If the sunk costs are to be recovered in prospective payments, therefore, there must be some form of requirement to pay these costs as a condition for using the grid. The resulting access charges would be the functional equivalent of the contract payments for new investment.

The need for access charges would not dictate the form of the charges. From an efficiency perspective, the preference would be to divorce the charge from usage levels, such as through a meter connection charge. In practice, the access charge might be collected proportional to usage, but in this case it should be at the last point of connection in order to minimize perverse incentives for inefficient bypass to avoid paying for sunk costs.

Assignment of the access charges could be simplified by the simultaneous allocation of the benefits of the existing grid through the award of transmission congestion contracts. One approach might be to award the initial ownership of the contract along with a long-term obligation to pay the access charges. Another possibility would be to auction the transmission congestion contracts and apply the revenues to reduce the required payment for the transmission grid. The remaining costs would be collected as access charges for all users of the grid. In effect, these access charges would pay for the assets of the grid in a system allowing open access for full use of the system. Everyone who used the system would pay the transmission usage charge for congestion, the second part of a two-part tariff, and the holders of transmission congestion contracts would have the transmission hedge.

---

[19] This situation appears to be what is described often as investments for reliability. However, with price responsive demand and security constrained economic dispatch, there is in principle no difference in reliability. The only difference created by the investment would be in the economic benefits of the actual dispatch.

Allocation of embedded costs and transmission rights for the existing system presents a major transitional challenge. However, this challenge must be met. To support a competitive market, and to meet the test of comparability, the old implicit allocations of rights must be made explicit. There will be many issues to resolve in this transition, and many ways to allocate costs and benefits without distorting the market. Looking forward, however, the market framework would condition system expansion decisions.

Grid expansion and pricing would continue to present a need for regulatory oversight, but the existence of workable transmission congestion contracts would substantially simplify transmission investment decisions. The users of the system who are buying and selling electricity without a complete hedge through transmission congestion contracts would face the short-term market clearing price. In the face of transmission congestion, the locational prices provide the proper incentive for investment in transmission facilities. Investments should be made when justified by the savings in congestion costs. Those who are prepared to make the investment would obtain the associated transmission congestion contracts. The role of regulators, therefore, would be to review requests for transmission expansion, examine the compatibility with the companion request for new transmission contracts, and ensure an open process for all to join in developing combined transmission investments recognizing the interactions in the network. The regulator would be responsible for enforcing a requirement for existing transmission facility owners to support expansions and reinforcements at a traditional regulated cost that recovered the incremental investment, and then to assign the corresponding transmission contracts. If no coalition of grid users were able to agree to pay for a grid expansion that appears to be beneficial for the system as a whole, any interested party could propose a project and an allocation of its costs among those grid users who would benefit. Regulatory procedures, similar to those used now, would determine whether the project should go forward and how its costs should be allocated to those expected to benefit from the effect on future locational pool prices, with the payers granted rights to compensation to assure that future congestion does not rob them of the benefits they are paying for.[20]

The transmission congestion contracts, once created, would no longer need any special regulation. Although investments in the transmission grid would by lumpy and would require the cooperation of the owners of existing facilities, the transmission congestion contracts would be divisible and freely tradable in a secondary market. This secondary market would provide a ready source of transmission hedges that would serve as an alternative to system expansion. The price of the transmission contracts should never rise above the long-term expected congestion opportunity costs or the cost of incremental expansion of the grid. In this way, the unregulated market for transmission congestion contracts would emulate the broad outlines of FERC's pricing policy. Transmission contracts would be obtained at the lesser of opportunity costs or incremental costs. Holders of existing transmission rights, converted into the appropriate transmission congestion contracts, would pay embedded costs but not opportunity costs. Those using the transmission grid without holding transmission congestion contracts would pay

---

[20] D. Garber, W. Hogan, L. Ruff, "An Efficient Electricity Market: Using a Pool to Support *Real* Competition," The Electricity Journal, Vol. 7, No. 7, September 1994, pp. 48-60.

opportunity costs but not any embedded costs other than the costs of any stranded assets that would be collected from all users.[21] Most important of all, the long-term transmission market could be more like a market, relying as much as possible on the incentives and forces of competition, limiting the role of planning and regulation to address the unavoidable interactions in the transmission grid. Investment decisions would be made at the initiative and with the agreement of those required to bear the cost.

Inefficient Transmission Expansion

Investment in the transmission grid should create new economic capacity. However, it is well known that different pricing regimes can create incentives for "uneconomic" transmission investment in the sense of reducing total welfare.[22] The simple example of a radial transmission line between two points would illustrate the point. Suppose that the line connects a low cost region to a high cost region. Under market pricing rules, generators in the high cost region would benefit if the capacity of the line were reduced. In an electric grid, a conceptually simple way to reduce the transfer capacity would be to add a weak parallel line between the two regions. Power flow would now split between the two parallel lines. Flow on the original parallel line would then have to be reduced to keep the parallel flow on the new line below its low capacity. It might be worth it to the high cost generators to build such a line, lower transfers, and raise prices in the high cost region. Overall economic efficiency would be reduced, but the profit to the high cost generator could justify the expenditure for the transmission investment.

Most of the examples of such perverse transmission investment incentives rely on this property of parallel flows. This simple perverse example would be easy to dispose of in practice. The regulator might prohibit such degradations of capacity. However, in a real network, with complex interactions, the same effect may be present but less obvious. The analysis to reveal a truly uneconomic investment might be both more difficult and controversial. Hence, we cannot rule out the possibility that real investment incentives would be influenced by wealth transfers that dominate efficiency losses.

It would be desirable, therefore, to have a system of property rights that reduced or eliminated the incentives for such perverse investments. In the case of a contract network as described above, the investment incentives would be affected by the existence of the transmission congestion contracts. As demonstrated by Bushnell and Stoft, under certain restrictions, the existence of transmission congestion contracts may even fully internalize the efficiency effects of investment decisions.

---

21 These stranded assets might include transmission sunk costs that could not be recouped through a sale of transmission rights.

22 For a review of the literature on these incentives, see James B. Bushnell and Steven E. Stoft, "Electric Grid Investment Under a Contract Network Regime," Journal of Regulatory Economics, Vol. 10, 1996, pp. 61-79.

The allocation of transmission congestion contracts under a feasibility rule arises naturally from the interpretation of the transmission congestion contracts as described above. It is clear that simultaneous feasibility of the transmission congestion contracts is necessary to guarantee that the revenues collected by the system operator are sufficient to pay the obligations under the transmission congestion contracts. Furthermore, under certain minimal conditions, the same feasibility condition would be sufficient to guarantee that the revenues from a market equilibrium in the spot market would be sufficient to guarantee this revenue adequacy.[23] However, in addition, the allocation of transmission congestion contracts under the same feasibility rule mitigates incentives for inefficient transmission investment.

The definition of the simultaneous feasibility test and investment feasibility rule can be stated as:

Feasibility Test: The aggregate of all transmission congestion contracts defines a set of net power injections in the grid. The set of contracts is feasible if these injections and their associated power flows satisfy all the system constraints.

Feasibility Rule: The grid expansion investor selects a set of new transmission congestion contracts with the restriction that both the new and the old transmission congestion contracts will be simultaneously feasible after the system expansion.

Application of this rule would permit a great deal of flexibility. For example, if necessary, the incremental transmission congestion contracts could provide counterflow to eliminate some of the existing contracts and thereby change the pattern of use. This would be equivalent to purchasing the existing contracts, or simple assuming the responsibility to meet the financial obligations under these same contracts.

The feasibility test is (relatively) easy to perform and is a familiar step in any transmission analysis. However, there is an important simplification here compared to the standard analysis of transmission investment. In the traditional analysis of regulated markets, with no response to price changes, it has been necessary to examine the feasibility of any set of transmission flows under a large number of possible patterns of load and system usage. This is one reason that available transmission capacity is defined often with a range of values, or with substantial caveats about the sensitivity to the pattern of future load. In the present case, however, there is no commitment or requirement for feasibility of the set of transmission

---

[23] Scott M. Harvey, William W. Hogan and Susan L. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1997.

congestion contracts along with other uses of the system.[24] The only requirement is internal consistency of the set of transmission congestion contracts. Hence, there is only one load flow to analyze, and only for the loads associated with the contracts.

As regards to investment incentives, an obvious point is that any exposure to future costs or inefficient investment falls on those who are relying solely on the spot market. Future investments in the grid cannot reduce the welfare of aggregate use according to the existing transmission congestion contracts. By definition, it would always be possible to use the system in the manner defined by the pattern of transmission congestion contracts. Therefore, the holders of the transmission congestion contracts could either dispatch their load and generation according to the transmission congestion contracts, or be compensated such that they would be no worse off than using the system in this manner. With private benefits, therefore, all the exposure to rent transfers would be limited to the spot market participation.

This already takes us a long way towards a market based system. Long-term contracts could, in principle, immunize the parties to the spot-market variations. Exploitation of rent transfers and inefficient investment, if it remains at all, would be only in the spot market. Although any inefficiency would be detrimental, there is likely to be much less concern with designing full protection for spot-market participants who have chosen not to protect themselves through long-term contracts.

Bushnell and Stoft have gone further to show that under certain conditions, there would be an additional reduction in incentives to make investments in inefficient transmission. For example, if transmission congestion contracts initially match dispatch in the aggregate, and new transmission congestion contracts are allocated for transmission investment under the feasibility rule, then the increase in social welfare for any transmission investment will be at least as large as the ex post value of the new contracts.[25] Hence, investors in transmission only, who are profiting only from the new transmission congestion contracts they create, would have no incentive to make inefficient transmission investments. This moves part way to the best possible outcome, but it does not remove the possibility that a combination of rent transfers and profits on generation might outweigh the efficiency loss from perverse transmission investments.

Under more restrictive assumptions, we could improve on this situation. If transmission congestion contracts match dispatch individually, not just in the aggregate, then the allocation of transmission congestion contracts under the feasibility rule ensures that no one can

---

[24] This is true in the case of a complete system. In the case of regional, interconnected systems, the result would not apply to the individual regions, but only to the aggregate.

[25] James B. Bushnell and Steven E. Stoft, "Improving Private Incentives for Electric Grid Investment," Resource and Energy Economics, Vol. 19, 1997, pp. 85-108.

benefit from a network investment that reduces social welfare.[26] In effect, this condition is a limiting case of the point made above. With the dispatch perfectly matching the individual transmission congestion contracts, the spot market net of transmission congestion contract coverage is reduced to zero volume. Any investment and movement in the spot market would impose the costs on the individual making the investment, thereby internalizing the efficiency effects.

As summarized by Bushnell and Stoft:

> "Any evaluation of investment incentives provided by transmission property rights must take into account the rules for allocating rights to the new property produced by grid modifications. Measuring the new property can be difficult. In fact, some 'expansions' of the network can increase congestion and in effect destroy *existing* property.
>
> [Transmission congestion contracts], when allocated according to the feasibility rule, can, under some conditions, account for this destruction of property and penalize those responsible. This could pave the way for a market that properly accounts for the complex externalities of the electric grid."[27]

In practice, of course, the actual dispatch for changing market conditions will never fully match the allocation of transmission congestion contracts. A more complete analysis of the investment incentives, therefore, would depend on an expected value calculation. Bushnell and Stoft conjecture that any deviations from efficiency would be small. It is clear that the existence of transmission congestion contracts, or any system of perfectly tradeable physical rights, would improve the investment incentives, even while not completely internalizing all the network externalities. And for those who want to protect a long-term contract from future inefficient investment, the feasibility rule provides the mechanism whereby the guarantee could be available. Any inefficiency would be isolated with loads relying solely on the spot market.

CONTRACT NETWORK MODELS

The key elements of the contract network approach – bid-based economic dispatch, locational pricing, transmission congestion contracts, and market-based grid expansion rules – have been implemented or are in the final stages in many electric systems. Prominent examples

---

[26] James B. Bushnell and Steven E. Stoft, "Electric Grid Investment Under a Contract Network Regime," Journal of Regulatory Economics, Vol. 10, 1996, pp. 61-79. For a related analysis in the context of generator market power, see Yves Smeers and Wei Jing-Yuan, "Transmission Contracts May also Hinder Detrimental Network Investments in Oligopolistic Electricity Markets," CORE, October 28, 1997.

[27] James B. Bushnell and Steven E. Stoft, "Grid Investment: Can a Market Do the Job?", The Electricity Journal, January-February, 1996, p. 79. (emphasis in original)

include:[28]

PJM	(Pennsylvania-New Jersey-Maryland Interconnection). FERC approved this contract network approach with all the essential elements. Implemented at the beginning of April 1998, with bid-based economic dispatch, locational pricing, transmission congestion contracts, and support for market-based grid expansion rules.
NYPP	(New York Power Pool). At the end of 1997, submitted for FERC approval a complete contract network system with bid-based economic dispatch, nodal pricing, transmission congestion contracts, and market-based grid expansion rules.
New Zealand	At the end of 1997, bid-based economic dispatch, with nodal pricing in place; transmission congestion contracts and market-based grid expansion rules under consideration or about to be implemented.
Norway	At the end of 1997, bid-based economic dispatch and locational pricing were in place, with transmission congestion contracts and market-based grid expansion rules under review.[29]
WEPEX	(California) At the end of March 1998, a restricted bid-based economic dispatch and approximate locational pricing through zones became operational. Market-based transmission expansion a declared policy, but transmission congestion contract implementation deferred.
Australia	The planned National Energy Market structure includes bid-based economic dispatch with approximate locational pricing through zones. Proposals for inter-regional financial hedging contracts, similar to transmission congestion contracts, under consideration. Transmission investment rules include opportunities for market participants to propose expansions.

The theory for dealing with competitive markets in the face of substantial economies of scale and scope, and the practical implementation of the contract network approach, present numerous challenges. The theory outlined above exploits the price-taking assumption of the competitive market model. In the presence of market power in generation, however, there would

[28] For further discussion of the development of independent system operators, see William W. Hogan, "Independent System Operator: Pricing and Flexibility in a Competitive Electricity Market," Harvard-Japan Project on Energy and the Environment, Harvard University, March 1998.

[29] Paul Dawson and Graham Shuttleworth, "Efficient Transmission Pricing in Norway and Sweden," A Report for NVE, National Economic Research Associates, London, February 1997.

be no special exemption for transmission. Inefficient investment in the transmission grid could be used to create market power in generation by imposing new bottlenecks. Transmission congestion contracts could mitigate but are not likely to eliminate this form of market power.[30]

Continuing oversight of transmission owners will focus on providing incentives for making and maintaining transmission investments. In the United States, Section 211 of the Energy Policy Act of 1992 requires grid owners to connect new users and expand the grid as necessary. Enforcement for true open access will not be easy, but it would be facilitated by the availability of the contract network model. The advance agreement to pay for transmission investment would eliminate one of the objections of transmission providers by reallocating the risks to market participants who can enjoy market returns. The incentives for maintenance and equipment replacement to preserve the existing capacity of the grid have not been given enough attention, but would also be facilitated by an assignment of responsibility for preserving the feasibility of the set of transmission congestion contracts.

The planning role and incentives for the system operator would also be affected. Analysis if grid expansion opportunities will at some stage involve the independent system operator. The timing, degree of participation, and incentives for the system operator are subject of controversy. The role of the system operator would be especially important in dealing with the free-rider problems where efficient transmission investments are being overlooked by the market participants. However, the more effective the transmission congestion contracts are in supporting market decisions, the less will be the need for such non-market interventions. In addition, the ability to allocate the transmission benefits as well as the costs would simplify if not eliminate such central investment decisions.

CONCLUSION

A pool-based, short-term electricity market coordinated by a system operator provides a foundation for building a system that includes tradeable transmission property rights in the form of transmission congestion contracts. Coordination through the system operator is unavoidable, and spot-market locational prices define the opportunity costs of transmission that would determine the market value of the transmission rights without requiring physical trading and without restricting the actual use of the system. In this setting, these transmission congestion contracts are equivalent to perfectly tradeable physical rights. Hence, this organization of the market defines a context where it would be possible to rely more on market forces, partly if not completely, to drive transmission expansion. In the case of market failure, transmission investment would still require some regulatory supervision, but these case would be simplified through the use of transmission congestion contracts to allocate the transmission benefits along with the transmission costs.

---

[30] For an analysis of a particular game theoretic model examining the interaction of market power in generation and transmission congestion contracts, see Steven E. Stoft, "How Financial Transmission Rights Can Curb Market Power," Office of Economic Policy, federal Energy Regulatory Commission, March 25, 1998.