J O I N T C E N T E R

AEI-BROOKINGS JOINT CENTER FOR REGULATORY STUDIES

Remedying Undue Discrimination through Open Access Transmission Service and Standard Electricity Market Design

Paul L. Joskow

Regulatory Analysis 03-1

February 2003

Mr. Joskow is Elizabeth and James Killian Professor of Economics and Management and Director, Center for Energy and Environmental Policy Research at the Massachusetts Institute of Technology (MIT), Cambridge, MA. Mr. Joskow submitted these comments before the Federal Energy Regulatory Commission on January 10, 2003. The views expressed here are those of the author and should not be interpreted as representing the views of MIT or any other organization with which he is affiliated. Mr. Joskow benefited greatly from discussions with Steven Stoft (resource adequacy) and from ongoing research with Jean Tirole (transmission investment and transmission rights), though these individuals bear no responsibility for the views expressed here. These comments also draw on my research supported by the MIT Center for Energy and Environmental Policy Research and the Cambridge University-MIT Project on Liberalized Energy Markets. Mr. Joskow can be contacted at 617-253-6664 or pjoskow@mit.edu.

JOINT CENTER

AEI-BROOKINGS JOINT CENTER FOR REGULATORY STUDIES

In response to growing concerns about understanding the impact of regulation on consumers, business, and government, the American Enterprise Institute and the Brookings Institution have established the AEI-Brookings Joint Center for Regulatory Studies. The primary purpose of the center is to hold lawmakers and regulators more accountable by providing thoughtful, objective analysis of existing regulatory programs and new regulatory proposals. The Joint Center builds on AEI's and Brookings's impressive body of work over the past three decades that has evaluated the economic impact of regulation and offered constructive suggestions for implementing reforms to enhance productivity and consumer welfare. The views in Joint Center publications are those of the authors and do not necessarily reflect the views of the staff, council of academic advisers, or fellows.

All AEI-Brookings Joint Center publications can be found at www.aei.brookings.org

TABLE OF CONTENTS

Remedying Undue Discrimination through Open Access Transmission Service and Standard Electricity Market Design

Paul L. Joskow

I. INTRODUCTION

I respectfully submit these Comments in response to the Federal Energy Regulatory Commission's ("Commission") July 31, 2002, Notice of Proposed Rulemaking ("NOPR")[1] regarding Network Access Service and Standard Market Design ("SMD"). I strongly support the Commission's efforts to develop and implement policies designed to break down the remaining barriers to the continued evolution of efficient competitive wholesale electricity markets that span large geographic areas. Such policies are essential so that the ongoing restructuring of the US electricity sector and the expansion of wholesale and retail competition can be successful in bringing long term economic benefits to consumers. There are many features of the proposed Standard Market Design (SMD) that I support since they will help to achieve these goals. They include the basic features of the proposed design for day-ahead and real time markets for energy and ancillary services, including the associated use of Locational Marginal Pricing (LMP), the market power mitigation and market monitoring proposals,[2,3] and many of the related proposals designed to improve the efficiency of wholesale markets and mitigate discrimination in the pricing and use of the interstate transmission system that impede

[1] *Remedying Undue Discrimination through Open Access Transmission Service and Standard Electricity Market Design*, 100 FERC ¶ 61,138, FERC Stats. & Regs., Proposed Regs. ¶ 32,563 (2002).
[2] I participated in a Commission Staff Technical Conference on market power monitoring issues in Washington, D.C. on October 2, 2002.
[3] An exception is the failure of the NOPR to recognize that the allocation of CRRs can enhance market power and the related failure to require that the market power mitigation and monitoring protocols to take this into account. My comments below discuss these CRR-related market power issues briefly.

wholesale and retail competition. I am not submitting Comments on these topics either because I am in general agreement with the thoughtful proposals contained in the NOPR or because remaining issues appear to have already been well developed in the record.

My Comments focus primarily on two sets of issues. First, I offer Comments on the proposed regional long-term resource adequacy obligations. Second, I offer Comments on several components of the proposed framework for stimulating efficient levels of transmission investment. My Comments on the proposed transmission investment framework examine the role of merchant transmission investment, the role of congestion revenue rights (CRRs) in stimulating needed transmission investment, alternative approaches to transmission pricing based on beneficiaries pay principles, market power issues raised by the allocation of CRRs, the proposed assignment of all congestion revenue shortfalls to TOs, the NOPR's conceptualization of "congestion management," and the importance of implementing a program of performance based regulation (PBR) for transmission owners (TOs) and (ideally) independent system operators (SOs). I outline an alternative approach to transmission investment and pricing, including a workable approach to pricing based on beneficiaries pay principles, that I request the Commission to reflect in the Final Rule. My Comments endeavor to provide theoretical and empirical support for and constructive criticisms of these important components of the NOPR. I hope that they will be read from this perspective.

a. Long Term Resource Adequacy

The NOPR properly concludes that energy markets alone are unlikely to provide adequate incentives for long term investment in or continued operation of generating capacity or demand response capabilities consistent with the level of electricity system reliability that consumers have come to expect. However, my analysis leads me to conclude that the reasons for reaching this conclusion are both more extensive and more complex than described in the NOPR. They go beyond imperfections in spot energy and operating reserve markets and the effects of the proposed market power mitigation measures on spot prices and include, among other things, poor incentives for LSEs to enter into forward contracts as a consequence of the unsettled state of retail competition and changing wholesale market rules, and the effects on spot market prices for energy and operating reserves resulting from the ways in which ISOs manage scarcity conditions. Moreover, the NOPR provides little in the way of empirical/factual analysis to support its conclusions. The kinds of sweeping changes being proposed by the Commission in this NOPR require sound supporting empirical analysis as well as sound theoretical analysis. This is especially true with regard to the proposed resource adequacy obligations since there appears to be abundant generating capacity at the present time in most regions of the country. This in turn has led some to conclude that the proposed resource adequacy obligations represent an unjustified bailout for merchant generators who are experiencing financial difficulties as a result of abundant supplies and low wholesale prices. Accordingly, my Comments seek to bolster the record on investment incentives, resource adequacy, and reliability in both the theoretical and the empirical dimensions. I present the results of a study that I performed of the experience in New England regarding the net revenues earned from the spot markets for energy and operating reserves by the marginal generators that operate during a relatively small number of "scarcity hours" each year. I also reference related empirical analyses for PJM and the New York ISO. These studies all support the NOPR's conclusions.

While I agree with the NOPR's conclusions regarding the failure of energy markets alone to provide adequate incentives for investment in generating capacity and demand response capabilities consistent with traditional reliability levels, I have concluded that there are some deficiencies in the particular resource adequacy rules that are proposed in the NOPR. The proposals are especially problematic for those regions where the states have introduced retail competition. And it is in these regions where the market and institutional imperfections that create the need for a resource adequacy requirement are likely to be most severe. Moreover, designing a good set of rules and procedures for resource adequacy obligation is a significant challenge. This challenge is made more difficult by differences between regions in the extent to which retail competition has been introduced, the extent to which distribution utilities retain a traditional obligation to serve through a combination of ownership of generating resources and wholesale power contracts with third parties, and differences in the economic characteristics of the existing fleet of generating plants.

These considerations lead me to conclude the it would be best to give the states and regions flexibility to determine exactly how they design, implement and enforce resource adequacy obligations. Rather than requiring a specific uniform national approach to resource adequacy problems, I suggest that the Commission specify a set of basic principles that acceptable resource adequacy programs must satisfy. My Comments propose and discuss a set of basic principles to guide the development of regional resource adequacy standards and obligations, recognizing that resolving the market and institutional imperfections that create the need for resource adequacy obligations should continue to be a high priority.

b. Transmission Investment, Pricing and CRRs

I agree with the NOPR's conclusion that " [c]ompetitive and reliable regional power markets require adequate transmission infrastructure to allow geographically broad supply choices and minimize the complications created by loop flow." (¶ 335) I also agree that over the last several years there has been inadequate investment in transmission capacity to reduce congestion and to support robust competitive wholesale markets for electricity (¶ 191) and that it is essential to remove any cost recovery impediments that act as barriers to the development of new transmission capacity ((¶ 196). Finally, I agree that it is desirable, <u>to the extent that it is practical and does not create additional barriers to transmission investment,</u> to match cost responsibility to the beneficiaries of transmission upgrades (¶ 197).

However, after careful analysis I have concluded that the framework that is proposed in the NOPR as the foundation for stimulating transmission investment (¶335-351) has very serious deficiencies and will not achieve the Commission's goals. If it is implemented as proposed in the NOPR, it is more likely than not to <u>reduce</u> the pace of investment in new transmission capacity rather than to increase it. This will lead to growing congestion, increasing market power problems, and growing demand for ever more market power mitigation and other potentially costly regulatory interventions.

The proposed approach seems to be based on the assumption that we can rely primarily on "private initiative" to bring forth needed transmission capacity and views "market driven" decisions as the "fundamental mechanism" to provide efficient levels of transmission investment. Thus it appears that the Commission has in mind a regime in which the bulk of future transmission investment will be realized by "merchant transmission projects" that would be supported financially through congestion revenues and the sale of CRRs that reflect the current market value of future congestion revenues. (¶346-347). The framework seems to view alternatives to merchant transmission projects as secondary or tertiary complements to fill modest gaps in transmission investment needs that are not otherwise provided by merchant investors. Regulated incumbent TOs would then play only a supplier of last resort role.

My analysis leads to the conclusion that this vision of future transmission investment needs being met primarily by merchant transmission projects is a fantasy (or perhaps an economist's dream) that fails to incorporate important economic and physical attributes of transmission investments, and is inconsistent with sound economic theory that reflects these attributes. The Commission has taken a leap from the correct observation that merchant transmission investment may be able to provide <u>some</u> of the needed investment in new transmission capacity to the erroneous conclusion that it can be relied upon to provide <u>most</u> of the efficient investments in transmission capacity. The NOPR provides little in the way of theoretical or empirical support for this leap of faith.

On it's face, the proposition that merchant transmission investment supported by revenues from CRRs provides a sound framework to stimulate efficient investment in transmission capacity is <u>inconsistent</u> with the Commission's conclusion that there has been inadequate investment in transmission capacity. I use the capital and operating costs of PJM's transmission network to illustrate why this is the case, demonstrating empirically that if the merchant investment framework had been relied upon to support all of the capital and maintenance costs of PJM's existing transmission infrastructure, congestion revenues would now cover less than 20% of these costs.

My Comments conclude that the NOPR's transmission investment framework is based on a poorly developed and immature theoretical foundation and is inconsistent with the limited empirical evidence that is available. More realistic theoretical analysis and the empirical evidence that does exist suggest that merchant transmission investment will play only a limited role in the future. My Comments explain some of the more serious flaws in the theory upon which the proposals appear to rely and explore the limited international experience with the kinds of merchant transmission projects contemplated in the NOPR. These flaws include: the failure to take economies of scale (or "lumpiness") into account; the failure to recognize that the market and institutional imperfections that the NOPR concludes lead to underinvestment in generating capacity and demand response capabilities will also distort merchant transmission investment incentives; the failure to take account of important firm-specific asset specificity attributes of transmission investments; and a property rights (CRR) definition and allocation framework that is poorly adapted to the physical attributes of real transmission networks. I explain as well why the experience with merchant transmission projects supported by congestion rents in Australia is not very encouraging.

My Comments also explain why the proposed CRR framework cannot provide the perfect hedges against variations in congestion costs envisioned in the NOPR without requiring subsidies by consumers or TOs. Nor does the proposed CRR framework provide efficient incentives for merchant investment in transmission. I suggest alternative approaches for defining financial transmission rights that are likely to have superior properties. However, my overall conclusion about the current state of knowledge is that there are many issues associated with the appropriate definition of property rights for transmission capacity that still need to be resolved, and that the existing theory is too immature and incomplete for it to be responsibly released for "prime time" application as the foundation for a new merchant transmission investment framework.

My Comments go on to explain why these flaws also lead to the conclusion that the proposal to require TOs to bear the full cost of CRR revenue shortcomings is flawed as well. The proposal would subsidize CRRs, may be confiscatory, and distorts investment incentives. Incentives designed to promote good TO maintenance and operating practices are desirable, but the proposed treatment of CRR revenue shortfalls is not a well-designed performance based incentive mechanism.

The erroneous assumption that we can rely primarily on unregulated merchant transmission projects to provide the investments needed to achieve the Commission's transmission infrastructure goals not only leads to a flawed transmission investment framework, but may be leading to other inefficient distortions in the institutional arrangements that govern transmission investment. First, while I agree that a "beneficiary pays" approach should be the foundation for defining the criteria that will determine "who pays" for transmission investments, the implementation of beneficiaries pay principles through project-by-project negotiations between potential beneficiaries of transmission projects is unworkable except in special cases. This approach is likely to continue to be a very serious barrier to efficient investment in the transmission network.

Second, the merchant transmission framework has led the Commission to place far too much weight on promoting "competition" between generation and transmission investments. Indeed, the emphasis that the Commission has placed on this kind of "competition" has created barriers to efficient investment in both generation and transmission, reducing real competition between generators. Transmission and generation do not "compete" in the same way as do, say, Dell, Gateway, HP and IBM in the manufacturing and retail distribution of personal computers and workstations. Rather, transmission investments that reduce congestion enhance the ability of the network to move electricity from location to location, expanding the geographic expanse of competition between generators at different locations on the network. The mantra of "let transmission and generation compete" has become an excuse for suppressing rather than for promoting real competition among generators by frustrating investment in transmission infrastructure.

Instead of assuming that we can rely on merchant transmission to provide the primary governance framework for future transmission investment, the Commission should pursue a transmission investment framework that is better matched to the economic and physical attributes of transmission enhancement projects and has a track record of good performance. The one internationally proven way to stimulate transmission investment is to rely primarily on incumbent regulated transmission network owners operating under sound regional planning guidelines and subject to well-designed performance-based regulatory mechanisms to be the primary vehicle for building, financing and operating transmission facilities. As part of this framework, ex ante criteria, based on a practical set of "beneficiaries pay" principles, would be used to apportion cost responsibility, rather than the unworkable case-by-case negotiation procedure that appears to be a key component of a merchant transmission model.

The approach that I recommend <u>does not</u> mean that merchant transmission projects should not be permitted. Merchant transmission investment should continue be an option available to market participants as long as it is properly reviewed through a regional transmission planning process. The approach that I recommend simply recognizes the practical limits to merchant transmission investment and some of the problems that adopting a framework that relies on it creates. I discuss the most likely attributes of potential merchant transmission investment opportunities and how these proposed projects should be evaluated along with other investment opportunities in a regional planning process.

c. A More Comprehensive View of Congestion Management Is Needed

My Comments also discuss inadequacies in the way the NOPR has conceptualized "congestion management." The NOPR embodies an excessively narrow definition of "congestion management" and improperly equates an ITP's security-constrained bid-based dispatch and the resulting LMPs with a more comprehensive and more appropriate view of congestion management and associated network operation, maintenance, and investment activities that all affect congestion. I support the basic components of the proposed SMD for day-ahead and real time markets, the associated security constrained bid-based dispatch, the computations of the resulting LMPs, and what the NOPR repeatedly refers to as "congestion management." However, it would be much more appropriate to refer to this system as providing a mechanism to "allocate available transmission capacity" efficiently rather than a comprehensive system of "congestion management."

A comprehensive view of "congestion management" should encompasses all actions that can be taken by system operators and transmission owners that can affect congestion and associated congestion costs. Congestion management actions properly encompass maintenance decisions and expenditures, physical operating decisions that are still made (or should be made) by incumbent TOs, and investments, small and large, in the transmission network that are likely to be made most economically by these TOs. While an efficient security constrained dispatch that results in an efficient allocation of a given amount of scarce available transmission capacity is an important part of the congestion management challenge, there are other important aspects of congestion management that can have significant effects on the amount of effective transmission capacity available to allocate through security-constrained dispatch protocols. Many of these aspects of congestion management are today and are likely to continue to be in the hands of incumbent regulated TOs and affected by maintenance, operating and investment decisions they make.

My comments illustrate the differences between a comprehensive view of congestion management and the narrower concept of an efficient allocation of available transmission capacity, using available data for PJM. It is widely acknowledged that PJM has done a very good job using spot market mechanisms along with security-constrained dispatch protocols to allocate efficiently scarce transmission capacity on the PJM network. This is a very good and commendable result. However, there is no evidence that PJM has a similar exemplary record when in comes to actually reducing congestion and managing it in the broader sense described above. Between 1998 and 2001 the transmission constraint hours experience on the PJM system increased by 661%, while transmission congestion charges increased by 500% to 1000%. During the same time period, PJM's operating expenses, including interconnection study expenses, increased by roughly 1000%. Yet transmission investment in PJM has stagnated, as it has elsewhere in the country, falling from an average of $166 million per year during the 1994-1997 period (with a peak of $240 million in 1994) to $94 million per year during the period 1998-2001 (with a low of less than $50 million in 1999). During this period wholesale spot energy costs in PJM have increased as well.[4] These observations are not criticisms of PJM, but rather demonstrate that neither LMP nor ITPs alone or in combination are magic elixirs for managing transmission congestion and stimulating transmission investment in a comprehensive manner.

d. Performance Based Regulation Should Be Reflected in the Final Rule

My conclusions regarding the transmission investment framework and the broader and more appropriate conceptualization of congestion management also lead me to conclude that the NOPR is missing an important component of a comprehensive program to increase investment in and the performance of the nation's transmission infrastructure. Specifically, the NOPR fails to include the development of a comprehensive performance based regulation framework applicable to TOs and (ideally) SOs. The Commission should give a much higher priority to working with ITPs, RTOs, TOs and state regulators to develop and apply good performance based regulation mechanisms that will stimulate a much broader range of beneficial congestion management efforts. Aside from the proposal to rely on a requirement that TOs be responsible for CRR revenue shortfalls, a proposal that my Comments argue has serious deficiencies, the NOPR largely ignores

[4] PJM Interconnection, *State of the Market Report: 2001*, page 33, Table 3, page 36, Table 5, and page 37, Table 6.

these very basic and very important incentive issues. I strongly encourage the Commission aggressively to support the development of performance-based regulatory mechanisms to be applied both to transmission owners and system operators in the Final Rule.[5]

[5] See Comments on these issues that I filed with the Commission over three years ago, available at: http://web.mit.edu/ceepr/www/99010.pdf

II. LONG TERM RESOURCE ADEQUACY (¶457 - ¶455)

The NOPR concludes that spot market prices for energy and operating reserves alone will not stimulate adequate and efficient investment in generating capacity and demand-response capabilities to achieve reliability levels that match consumer preferences/valuation for reliability.[6] Three sets of market and institutional imperfections are identified in the NOPR as the primary factors creating this problem. The NOPR then proposes a long-term regional resource adequacy requirement and associated procedures to implement and enforce it.

I agree with the NOPR's conclusion that, at least in the medium term, spot market prices for energy and operating reserves alone are unlikely to provide adequate incentives to achieve reliability levels that match consumer demand for reliability. I will present some empirical evidence below that reinforces this conclusion. A variety of market and institutional imperfections contribute to this problem. The relevant imperfections include those discussed in the NOPR, as well as additional market and institutional imperfections that are not mentioned but should be taken into account as well. Some type of resource adequacy requirement placed on all LSEs is needed to compensate for these market and institutional imperfections at this time. However, actions designed to remedy these market and institutional imperfections should be given a high priority so that the need for resource adequacy obligations can fade away over time.

[6] Good resource adequacy policies should consider both the quantity of resources available to meet demand reliably as well as the type (attributes) of the resources available to meet demand reliably and economically. For example, a policy should not encourage investment in simple-cycle peaking turbines if investment in an equivalent quantity of CCGT capacity would be more economical overall.

There are aspects of the NOPR's proposed resource adequacy requirements that certainly have considerable merit and, indeed, are very clever (e.g., the requirement that both generation and demand response capabilities be included,[7] transmission deliverability requirements, relying heavily on incentives to induce LSEs to make forward arrangements for resources to match their customers' reliability preferences). Unfortunately, the NOPR's proposed resource adequacy requirements, procedures and enforcement mechanisms also have a number of significant deficiencies that should be remedied in the Final Rule. The proposed resource adequacy rules are particularly poorly adapted to states that have introduced retail competition programs. Yet it is in these states where the market and institutional imperfections that create the under-investment problem are likely to be most severe.

To understand why spot energy and operating reserves markets are unlikely to provide adequate incentives for investment in generating capacity and demand response capabilities it is useful to articulate clearly the attributes of an "ideal" perfectly competitive wholesale energy and operating reserve markets. We can then identify where actual market and institutional conditions depart from this competitive ideal and what the consequences for investment and reliability are likely to be.

[7] It is not clear exactly what demand response capabilities the NOPR contemplates including as a "resource" and how the associated quantities will be measures. Clearly interruptible contracts and related demand control capabilities under the control of an ITP would be included. It is not clear how real time pricing capabilities would be included, however, if at all. Demand-side and supply side "resources" have different attributes and, in particular, measuring the "baseline" level of demand against which demand-side "resource" investments will be measured is a very significant challenge. The Final Rule should include more detailed criteria for qualifying demand response capabilities.

1. Attributes of "Ideal" Perfectly Competitive Spot Wholesale Markets

Efficient competitive energy and operating reserve markets (spot and forward) would have a number of important structural and behavioral attributes. These attributes include:

a. The demand placed on the wholesale market for energy *and* operating reserves should reflect consumers' willingness to pay for reliability (value of reducing energy consumption or value of lost load on the margin) on a minute-to-minute basis at every location on the network.

b. End-use consumers should be able to "see" and respond to short-term variations in prices for energy and operating reserves.

c. Buyers and sellers should have a full range of financial and contractual products available to manage market risks over all relevant contingencies, locations and future time horizons and have the incentive and ability to use these products to manage risks to match their risk preferences.

d. All suppliers of energy and operating reserves should receive competitive market-clearing prices at every moment in time for the services they provide to consumers. These prices should vary from location to location to reflect marginal congestion costs and marginal losses.

e. Under typical operating conditions, these market clearing prices should equal the marginal (opportunity) cost of the last increment of generating capacity that just clears supply and demand at each point in time. In the case of energy, this is the marginal cost of producing a little more or a little less energy. See Figure 1. Inframarginal generating units earn net revenues to cover their fixed costs whenever the market clearing price exceeds their own marginal generation costs. In the case of operating reserves, the efficient price is (roughly) equal to the <u>difference</u> between the price of energy and the marginal cost of the next increment of generation that could supply energy profitably if the price of energy were slightly higher plus any direct costs incurred to provide operating reserves (e.g. costs associated with spinning). This price for operating reserves is equal to the marginal <u>opportunity cost</u> incurred by generators standing in reserve rather than supplying energy. Under typical operating conditions the price of operating reserves will be very small, and far below the price of energy.

f. For a relatively small number of hours each year there will be excess demand at a price that equals to the marginal supply cost of the last increment of generating capacity available on the network to supply energy or operating reserves and prices will rise to a (much) higher level reflecting the value (or value of lost energy or load) that consumers place on consuming less electricity as demand is reduced to match available supplies. See Figure 2. In what follows I will refer to these conditions as "scarcity" conditions. Under competitive scarcity conditions (i.e. in the absence of seller market power), the competitive market clearing price of energy will now generally be much higher than the marginal production cost of supplying the last available increment of energy, reflecting the high opportunity cost (value of lost energy or lost load) that consumers place on reducing consumption by a significant amount on short notice. Furthermore, while the price of operating reserves will continue to be equal to the marginal opportunity cost incurred by generators standing in reserve rather than supplying energy, the opportunity cost of standing in reserve rather than supplying energy will rise significantly as well in response to the higher "scarcity value" of energy. <u>All</u> generating units actually supplying energy and operating reserves in the spot market during scarcity conditions would earn substantial scarcity rents to cover their fixed costs during these conditions. For base load and cycling units, the net revenues they earn during scarcity conditions may account for a significant fraction of the total net revenues they earn throughout the year. For peaking capacity that supplies energy or operating reserves primarily during scarcity conditions, the net revenues they earn during these periods will account for substantially <u>all</u> of the net revenues available to cover their fixed costs (capital, maintenance and operating.)

g. Prices paid by consumers and suppliers at each location should reflect the marginal cost of congestion and of marginal losses consistent with an efficient allocation of scarce transmission capacity. This implies that consumers and suppliers should face locational prices that equal the marginal cost of increasing generation at each location (typical conditions) and when demand is sufficiently high by a higher price reflecting consumers' valuations of energy and reliability at that location (scarcity conditions).

h. Supply and demand should always rationed by price so that there are no involuntary "blackouts" imposed on consumers. See Figure 2.

i. There is no seller market power.

j. There is no buyer market power.

If all of these conditions are satisfied we will get an efficient set of spot market prices for energy and operating reserves at each location, an efficient set of forward prices for energy and operating reserves at each location, the right price incentives will be given to consumers and suppliers at each location to make spot allocation and forward contracting and risk management decisions, and the right long run investment incentives will be provided to suppliers of both generation and demand response services to match consumer valuations of reliability.

FIGURE 1

PERFECTLY COMPETITIVE WHOLESALE SPOT ELECTRICITY MARKET "NORMAL" CONDITIONS

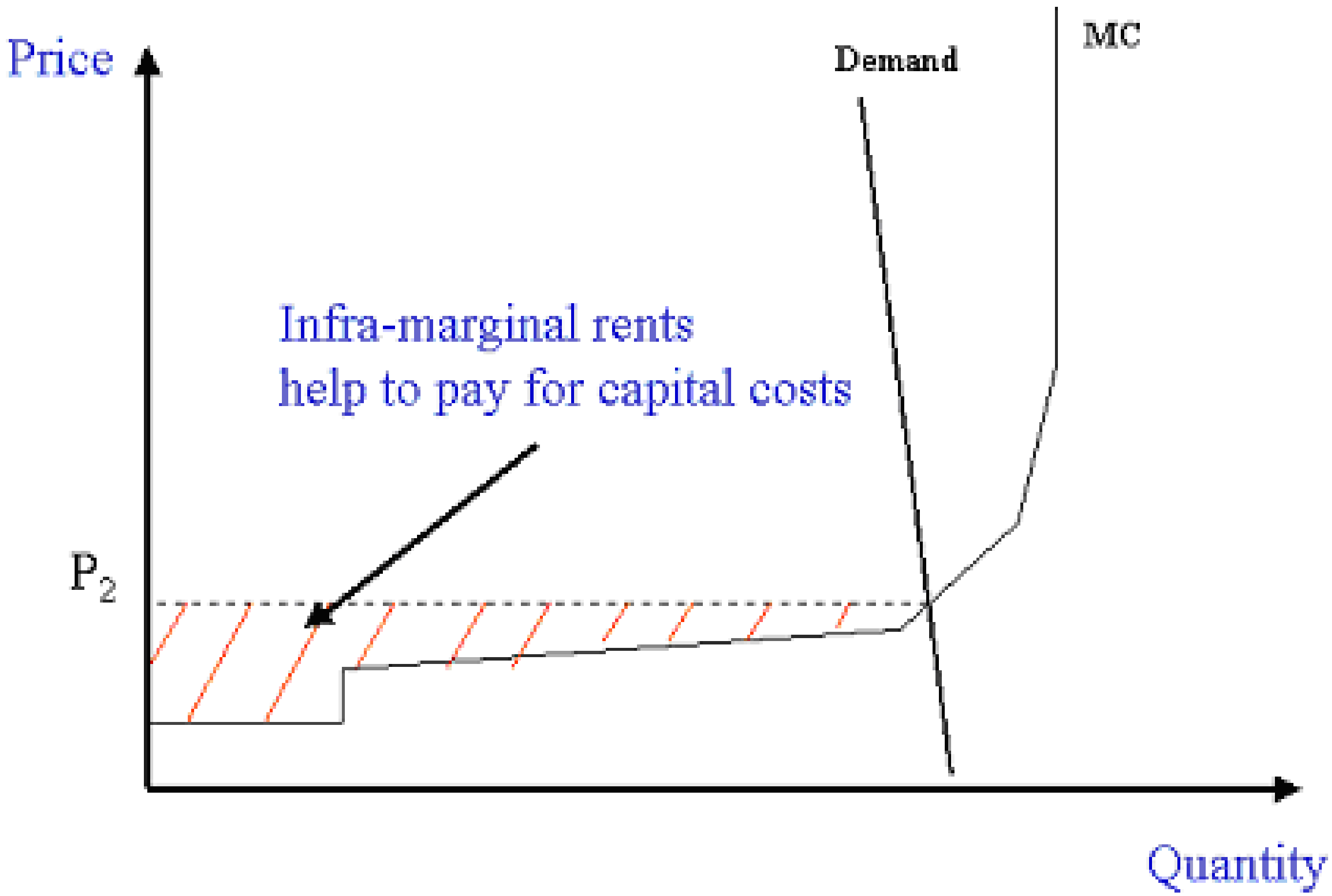
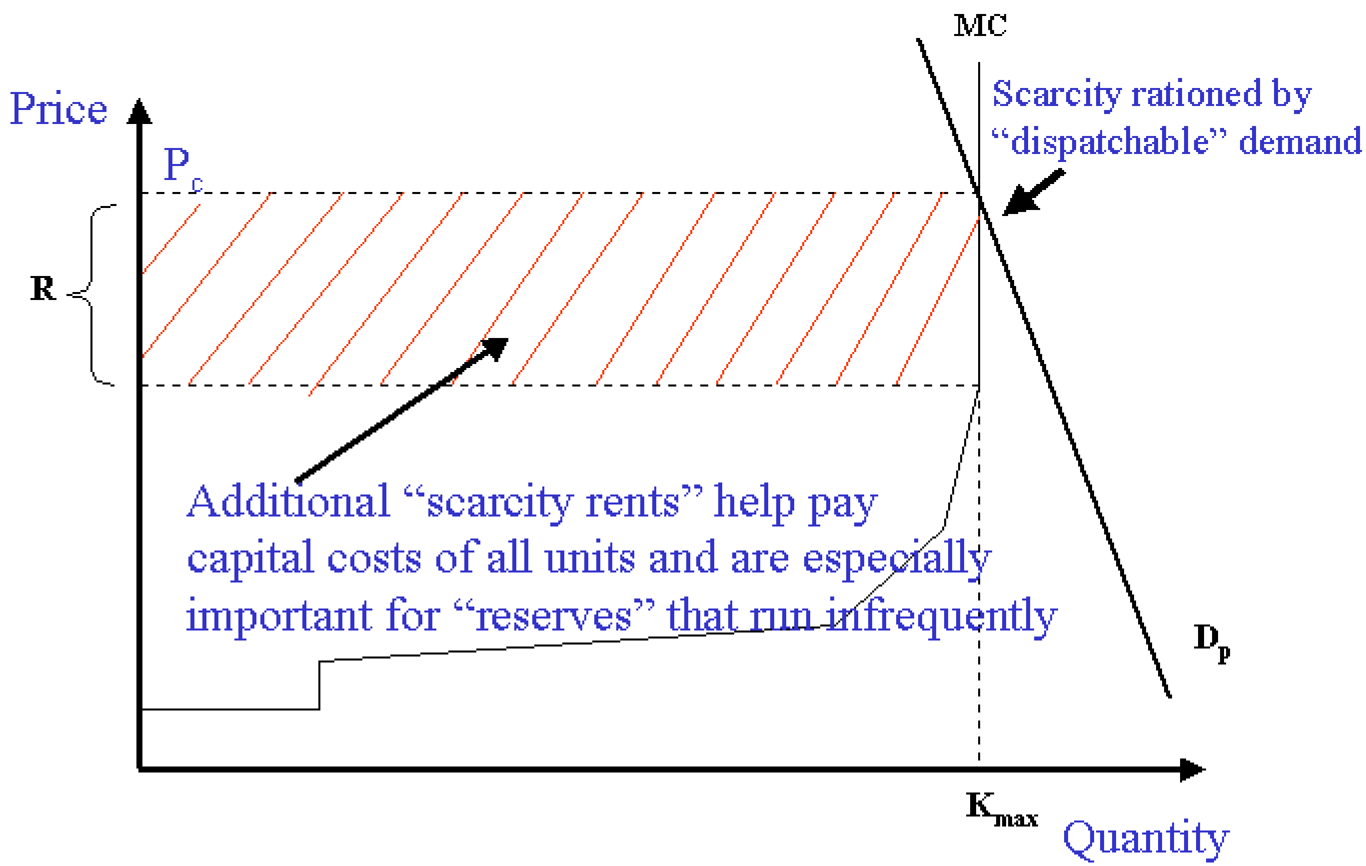


FIGURE 2

RATIONING SCARCE CAPACITY



2. Market Imperfections[8]

Even after the introduction of the improved market design features embodied in the proposed SMD, there are likely to be significant departures from these idealized conditions. These market imperfections will distort both short run allocation decisions by consumers and suppliers as well as long run investment incentives faced by generators and suppliers of demand-response capabilities. The most important market and institutional imperfections are:

a. Consumer demand for energy and reliability are not well represented in wholesale spot markets. Due to metering costs, communications and consumer response limitations, and the slow diffusion of both, consumers do not "see" all relevant spot prices for energy and operating reserves and cannot respond effectively to variations in them. These imperfections severely limit the ability of market mechanisms properly to reflect consumer valuations for alternative levels of reliability and for investors on the supply and demand sides to respond efficiently to them.

b. The limited amount of real time demand response in the wholesale market leads to spot market demand that is extremely inelastic. Especially during high demand periods as capacity constraints are approached, this creates significant opportunities for suppliers to exercise unilateral market power leading to supra-competitive prices even with a relatively unconcentrated distribution of suppliers.[9]

c. Scarce generating capacity is not price-rationed during true scarcity conditions. Reliance on "out-of-market" supply-side and demand-side resources to manage operating reserves deficiencies leads to spot prices for energy and operating reserves that may be too low during these conditions. The costs of these scarcity management tools are not reflected in spot market prices and may be spread in charges to

[8] Some commentators have responded to the NOPR's resource adequacy proposals by assuming that there are no market or institutional imperfections that need to be taken into account. Such comments are necessarily tautological and irrelevant; assume that all markets and institutions operate perfectly and, by definition, there are no problems! Comments that simply avoid dealing with the issues created by market and institutional imperfections contained in the NOPR, or that just assume them away, are not useful.

[9] Paul L. Joskow and Edward Kahn, "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000", *The Energy Journal*, Vol 23, No 4, (2002), pp. 1-35.

consumers over many non-reserve deficiency hours through an uplift charge.

d. As system operators manage operating reserve deficiencies the reliability of the system may deteriorate and "random" blackouts may be necessary. These reductions in effective service quality are generally "shared" across the network rather than allocated based on consumer valuations and the associated social costs are not accurately reflected in market prices. This creates incentives for "free-riding" which in turn leads to underinvestment in generating capacity and demand-response programs.

e. Immature, incomplete and illiquid forward markets for risk hedging/contracting arrangements undervalue rare events and make it difficult for consumers and suppliers to manage long-term risks efficiently. This in turn, reduces the ability of investors in new generating capacity to hedge market risks and increases their financing costs above what they would be if consumer and supplier risk preferences could be better matched.

f. Ambiguities in retail procurement responsibilities, competitive retail market imperfections and regulatory opportunism and uncertainty affects contracting incentives and behavior and leads to too much short- term forward contracting and too little long term contracting. This undermines the development of liquid forward markets for energy and operating reserves which in turn, reduces the ability of investors in new generating capacity to hedge market risks and increases their financing costs above what they would be if consumer and supplier risk preferences could be better matched.

In theory, these imperfections in spot <u>and</u> forward markets for energy and operating reserves could lead to too little <u>or</u> too much investment in generating capacity and associated operating reserves. Inelastic demand and market power lead to supracompetitive prices and to incentives for over-investment in generating capacity.[10] The other market imperfections generally lead to under-investment in generating capacity and demand response programs.

[10] See, for example, Richard Green and David Newbery, "Competition in the British Electricity Spot Market, *Journal of Political Economy*, 100(5), 929-53, 1992, examine the inefficiencies associated with excessive entry stimulate by supra-competitive prices resulting from the exercise of market power.

3. Institutional Imperfections

Whether it is too much or too little investment depends in part on the relative magnitudes of these effects in practice and in part on other institutional arrangements that affect spot market prices and the structure, behavior and performance of forward markets. The institutional arrangements of particular importance are:

a. Market power mitigation mechanisms

The NOPR proposes a variety of general and locational price mitigation measures to respond to potential market power problems in spot markets for energy and operating reserves. These mitigation measures include general bid caps (e.g. $1000/Mwh) applicable to all prices, location specific bid caps (e.g. marginal cost plus 10%), and other bid mitigation and supply obligation (must offer) measures.[11] Unfortunately, the supply and demand conditions which should lead to high spot market prices in a well functioning *competitive* wholesale market (i.e. when there is true competitive "scarcity") are also the conditions when *market power* problems are likely to be most sever (as capacity constraints are approached in the presence of inelastic demand, suppliers' unilateral incentives and ability to increase prices above competitive levels, perhaps by creating contrived scarcity, increase). Accordingly, even the best-designed mitigation measures will inevitably "clip" some high prices that truly reflect competitive supply scarcity and consumer valuations for energy and reliability as they endeavor to constrain high prices that reflect market power. The NOPR reflects the judgment that, on balance, these

[11] The NOPR proposes to require that under certain "non-competitive conditions" (e.g. local market power problems caused by congestion) generators be required to offer all available energy (must-offer requirement) to the system operator subject to a pre-specified bid cap. FERC Docket No. RM01-12-000, Notice of Proposed Rulemaking, July 31, 2002, ¶ 409. It also invites ITPs to propose additional mitigation measures that could apply under certain conditions where market power would be a significant problem, id. at ¶ 415. Finally, the NOPR provides for a regional "safety net bid cap" that would apply to the day-ahead and real time markets under all conditions, id. at ¶ 433.

mitigation measures will lead to prices that are too low during extreme conditions (e.g. reserve deficiency conditions) to attract sufficient investment in peaking capacity and demand response capabilities available during very high demand contingencies to match consumer preferences for reliability.

b. Discretionary behavior by ISOs/RTOs/ITPs during true scarcity conditions

Because electricity cannot be stored, electric power systems must rely on a combination of "reserve capacity" and short-term demand response to achieve the levels of reliability that consumers have become accustomed to. Reserve capacity is capacity that is needed to provide energy and operating reserves for a relatively small number of hours each year when demand is at its highest levels. This capacity stands idle for the vast majority of hours during the year, and is "in the money" for only a few hours each year when it is needed to provide energy or operating reserves to balance supply and demand. That is, the marginal generation at the top the generation supply stack that is available to clear the market under extremely high demand conditions is expected to supply energy or operating reserves for relatively few hours in a typical year. Owners of this type of peaking or reserve capacity rely on revenues they receive during these few hours of the year to recover their capital costs and their fixed operation and maintenance costs. These marginal suppliers of energy and operating reserves that provide the last increment of generating capacity required to achieve any particular reliability level depend almost completely on revenues produced during what I defined previously as "scarcity conditions." As I will discuss further presently, the profitability of investment in this marginal capacity, and the profitability of suppliers' continuing to make existing marginal capacity available to the system during extreme conditions (rather than retiring it), depends primarily on prices realized during scarcity conditions.

The level of prices for energy and operating reserves realized during these scarcity conditions also depend critically on the ways in which system operators respond to reserve deficiencies and how these responses are reflected in spot market prices for energy and operating reserves. Small changes in system operators' behavior can have large effects on the "scarcity rents" earned during these hours and, in turn, large effects on the profitability of investing in and making available the marginal capacity that has traditionally cleared the market under these conditions. There are three separate issues effecting investment incentives that emerge here. First, to the extent the system operators manage reserve deficiencies (true competitive scarcity) using "out-of-market" measures that are not reflected in spot market prices, spot market prices will be too low. Second, bid mitigation mechanisms are likely to become binding constraints during reserve deficiency conditions and may also depress spot market prices too much during these conditions. Third, the mere prospect that the discretionary behavior of system operators can have significant effects on the profitability of this marginal capacity raises classical opportunism problems. It is now widely recognized that opportunism problems lead to under-investment and that credible long-term contracts or vertical integration are efficient institutional responses to opportunism problems.[12]

Taken together, these problems will not be so easy to fix quickly. Moreover, sharply restricting the system operator's discretion to manage reserve deficiencies in real time can be very costly and increase rather than reduce reliability problems. All things

---

[12] Jean Tirole, *The Theory of Industrial Organization*, MIT Press, 1988, Paul L. Joskow, "Contract Duration and Relationship Specific Investment, *American Economic Review*, pp. 168-185, 1987.

considered, I agree with the NOPR's conclusion that the net effect of all of these market and institutional imperfections (several of which are not identified in the NOPR) is more likely than not to lead to underinvestment in generating capacity and demand response capabilities. However, whether this conclusion is valid, and how important the imperfections are likely to be in practice, are empirical issues and should be resolved based on empirical evidence, not just theoretical assertions. Few real markets satisfy all of the textbook conditions of perfect competition and are characterized by at least some market imperfections. Yet, we do not establish resource adequacy or other economic regulatory requirements in most markets just because there are some imperfections. This is the case because regulatory rules and obligations can create costs that exceed their benefits. Accordingly, an empirical assessment of the significance of these market and institutional imperfections in electricity should guide the use, structure and duration of any resource adequacy requirements and associated enforcement procedures. In this regard, electricity's unusual characteristics (e.g. no storage, inelastic demand, continuous balancing of supply and demand, physical network constraints, and the high costs of unplanned outages) suggest that the costs of having a given amount of "excess" resource capability are likely to be much lower than the costs of having an equivalent amount of resource deficiency.

4. Empirical Evidence on Net Revenues During Scarcity Conditions

This section presents a method to calculate the "scarcity rents" that are earned by the marginal generators that just clear the market when there is true competitive "scarcity." The next section applies this method to measure the scarcity rents produced from spot energy and operating reserve markets operated by ISO-New England during the period 1999-2002 (through November 27, 2002). That is, the method measures scarcity rents under conditions where available generating capacity must be "rationed" to balance supply and demand and to maintain the network's frequency, voltage and stability targets because available capacity to supply energy and the minimum level of operating reserves pursuant to bilateral contracts and ITP operated spot markets has been exhausted. The rationing could be done by allowing prices to rise sufficiently above the marginal operating cost of the highest cost generating unit available to supply energy and the minimum level of operating reserves until price sensitive demand falls to a level equal to the capacity available. The necessary rationing may also be accomplished by calling pre-negotiated interruptible contracts, by calling on emergency generators under pre-negotiated contracts (e.g. at hospitals and universities), by drawing operating reserves down below normal levels, increasing the risk of outages, by voltage reductions, or by rolling blackouts. The methods chosen to manage operating reserve deficiencies, and how the associated costs are recovered from consumers, can have significant effects on spot prices during reserve deficiency hours and on the scarcity rents earned by the marginal generators that run only during these hours.

In the U.S., traditional reliability levels imply that capacity "scarcity" as defined here occurs infrequently, typically when hourly demand is at its highest annual levels. Involuntary load shedding occurs even more infrequently.[13] So, significant scarcity rents accrue to marginal generators or demand response resources during a very small fraction

[13] I have seen various interpretations of the classical "one day in 10 years" reliability criterion. The most frequent interpretation I have been given is "one hour of unserved energy in 10 years." At the other extreme it is sometimes interpreted at "24 hours of unserved energy in 10 years." These interpretations would have very different implications for optimal reserve margins, though more modest impacts on total costs. Steven Stoft, *Power System Economics*, IEEE Press, pp. 163-164.

of the hours in a typical year. However, there are capital costs associated with investing in capacity that will be available to meet demand during these few hours, as well as annual fixed O&M costs associated with staffing and maintaining such a generating unit, property taxes, and any costs associated with meeting various regulatory (including environmental) requirements. There are also costs associated with starting a unit up to supply energy or available as operating reserves. For simplicity, I will refer to these costs as "fixed costs." Generating units that are expected to operate at very low capacity factors typically have relatively low fixed costs and relatively high marginal generating costs. Let C_K be the annualized fixed cost per Mw-year (including the amortization of investments in this capacity where relevant --- see below)) and MC_E the marginal operating costs per Mwh <u>of the last (highest operating cost) generating unit in the merit order available to provide energy or operating reserves</u>. Let P_s be the market price of electricity during "scarcity hours" and H_s the expected number of scarcity hours per year. The probability that "scarcity" conditions will exist is then given by $H_s/8760$. The condition for investors in the marginal unit of capacity to just break even given any given probability of scarcity ($H_s/8760$) is then given by:

(1) $$C_K = (P_s - MC_E)H_s = R$$

where R equals the annual expected scarcity rents that are available to cover the fixed costs of the "last unit of capacity" available to supply energy or operating reserves.

The "optimal" amount of generating capacity should reflect as well the valuation that consumers place on reliability and (ultimately) on being curtailed during scarcity conditions. Let V be the <u>hourly</u> opportunity cost that the <u>marginal</u> consumer incurs by consuming a little less during scarcity conditions, reflecting any associated reductions in network quality, the increased likelihood of being curtailed, and actual curtailments.[14] Then the efficient level of investment will be defined by equating the marginal cost of the last unit of generating capacity to the marginal consumers' expected cost of being in scarcity conditions:

(2) $$C_K = H_s*V$$

If we know V and C_K then we could derive the optimal H_s, the optimal probability of being in scarcity conditions (H_s/8760) and the optimal quantity of generating capacity and demand response capability consistent with this probability. The higher is V, the lower is the optimal H_s and the higher is the optimal amount of reserve capacity (and vice versa).

In the next section I discuss the methods that I used to estimate the values for R (and H_s) implied by hourly energy and operating reserve prices observed in ISO-New England's energy and ancillary services markets during the four-year period 1999-2002 (through November 27, 2002). The analysis assumes that a $1000/Mwh price cap has been in effect during this entire period.[15] I compare these scarcity rents to alternative

[14] For simplicity, this presentation is a little different from the traditional presentation which focus only on the cost of lost energy when load is curtailed. In reality, measuring V is very difficult. It varies from consumer to consumer, with the severity of scarcity conditions, and with the methods used to ration demand when curtailments are required. See generally Steven Stoft, *Power System Economics*, IEEE Press, Wiley Interscience, 2002, Chapter 2-5.

[15] There were five hours in May 2000 when spot energy prices exceeded the cap and the analysis reduces those prices to $1000/Mwh. However, the conclusions that flow from the analysis would not be changed if the actual prices realized in these hours had been used instead.

measures of C_K. The analysis shows that R has been much lower than C_K in New England over the last four years.

This result can lead to one of two conclusions. One conclusion would be that too little revenue is produced in the spot energy and ancillary services markets to support enough generating capacity to meet traditional reliability levels. An equilibrium consistent with traditional reliability levels would then require more revenue to flow to marginal generators either by allowing energy and operating reserve prices to rise when the price caps are now binding or by creating the opportunity for generators to earn revenues from sales of capacity to meet administratively determined capacity obligations. Another conclusion could be that traditional reliability levels are too high and that the probability of operating reserve deficiencies is too low (i.e. that condition (2) is not satisfied because H_s is too low). By reducing the system's generating capacity, holding demand constant, the number of hours of operating reserve deficiencies H_s would increase and the equilibrium condition defined in equation (1) would eventually be restored with a higher level of reserve deficiencies, high prices, less generating capacity, and more involuntary outages. I favor the first conclusion, and provide some evidence below to support it, but this is a judgment that the Commission must make to support the need for resource adequacy obligations.[16]

[16] A similar analysis has been done for the New York ISO and comes to similar conclusions. See David S. Patton, *Summer 2002 Review of New York Electricity Markets*, October 15, 2002, pp. 25, 42-64. This kind of analysis could also be performed for PJM. I do not have the necessary data or resources to perform these calculations for PJM. I hope that the methods presented here provide a framework for performing similar analyses in other regions. Data obtained from the very comprehensive PJM Interconnection annual *Report On the State of the Market* for 1999-2001 can be used to provide an upper bound estimate of R for PJM as well. While the value of R produced from spot energy prices appears to be higher in PJM than in New England, the upper bound value for R calculated in this way is still well below the value of C_K established in PJM. So, the basic story is the same in all three ISOs --- spot energy markets alone do not produce adequate scarcity rents to cover the long run incremental costs of the marginal generators and, at least in New England, may be too low to keep older generators on line that may be necessary to maintain

5. Scarcity Rents in New England

Most simple discussions of competitive "scarcity conditions" implicitly assume that this is the level of supply/demand where the lights will go out if supply is reduced by 1 Mw.[17] In fact, this is not an accurate characterization of how electric power networks are operated. "Scarcity conditions" are triggered when system operators find that they have an operating reserve deficiency that cannot be satisfied by buying more energy or operating reserves through ordinary organized spot market mechanisms.[18] This in turn typically triggers the SO's implementation of a set of "operating reserve conservation" actions to reduce demand or augment supply using out-of-market instruments. Only as a last resort --- and very infrequently --- has it been necessary to implement rolling blackouts with traditional reliability criteria and associated generating reserve margins. The calculations that I present here reflect this "operating reserve deficiency" protocol framework.

These calculations are performed for the hourly spot market in ISO-New England in the following way. First, I identified all hours when the ISO declared an operating reserve deficiency. Operating reserve deficiencies trigger NEPOOL Operating Procedure

---

traditional reliability levels. Of course during the 1999-2002 time period, there were capacity obligations in all three ISOs. These obligations create a market value for capacity when it is scarce. The revenues available to marginal generators (as well as to infra-marginal generators) from sales of capacity to LSEs and wholesale market intermediaries effectively increase scarcity rents. In PJM, the data available suggest that the combination of scarcity rents available to marginal generators from sales of energy and operating reserves plus the revenues available from sales of capacity is roughly sufficient to cover target estimates of C_K.

[17] This discussion assumes that the proposed market power mitigation mechanisms are successful, that prices for energy and operating reserves are competitive, and the market power does not lead to contrived scarcity. To the extent that scarcity conditions reflect the exercise of market power in New England during the period studied here, my estimates would overestimate the true competitive scarcity rents produced by the spot energy and operating reserve markets in New England under competitive conditions. The same is true for the data for PJM and NY described in the previous footnote.

[18]There is no reason in principal why a system operator should not be able to respond to projected reserve deficiencies by making forward (e.g. two-day ahead) commitments if that is a lower cost option. However, the proposed market rules do not provide for forward contracting by system operators. If system operators had the right financial incentives if would make sense to expand their contracting options.

4.[19] NEPOOL Operating Procedure 4 (Op-4) has 16 action steps of increasing severity. For example, Action 11 allows 30-minute reserves to go the zero. Action 12 begins the implementation of voltage reductions. Op-4 (or at least some steps in Op-4) seems to me to be a reasonable definition of "scarcity" when we should expect competitive market prices to rise far above the marginal operating cost of the last generator available to supply energy and operating reserves.

During these scarcity conditions, marginal generators can earn revenues in one of two ways. They may be called to supply energy and are paid for the energy supplied. Or they may be providing operating reserves and are paid for the operating reserves they supply. <u>These payments are not cumulative at a given point in time.</u> A generator (or in theory demand) is paid for one or the other at any moment in time. As previously noted, for generators supplying energy, the "scarcity rent" is the difference between the price they are paid and their marginal supply costs. For generators supplying operating reserves, the "scarcity rent" is no higher than the payment they receive for operating reserves. As discussed above, if energy and operating reserve markets are integrated efficiently, there is also a "textbook" relationship between the price of energy and the price of operating reserves during scarcity conditions. Specifically, the price of operating reserves should be roughly <u>equal</u> to the price of energy minus the marginal operating cost of the units providing operating reserves. That is, the price of operating reserves is equal to the "opportunity cost" incurred by generators supplying operating reserves rather than energy.

For all Op-4 hours during the period 1999 through November 27, 2002, I obtained the price of energy and the price of 10-minute operating reserves. When the price of

[19] http://www.iso-ne.com/operating_prodecures/Op4Fin.doc/ . accessed 11/27/02.

energy exceeded $1000, I set it to the $1000 price cap that was implemented after May 2000.[20] When the price of 10-minute spinning reserves and 10-minute non-spinning reserves were different (as was often the case during Op-4 conditions, especially in 1999), I took the highest of the two prices. There was only one hour when the operating reserve price exceeded $1000 and the price was set to $1000 for that hour.[21] To calculate the "scarcity rents" associated with supplying energy during Op-4 conditions, I assumed that the marginal costs of supplying energy from the marginal generator was either $50/Mwh or $100/Mwh (it doesn't matter much). This range should bracket the true marginal generating costs and any associated start-up, no-load and ramping costs for these units given variations in gas prices during this time period. I took the operating reserve revenues without making an adjustment for any operating costs incurred to supply operating reserves and, as a result, my method probably slightly overestimates the scarcity rents accruing to suppliers of operating reserves during scarcity conditions. I then aggregate the data for each year to calculate values for the "scarcity rents" per Mw-year available from supplying either energy or operating reserves (or any combination of the two) during scarcity conditions.

The results are reported in Table 1. The average scarcity rents from supplying either energy or operating reserves during OP-4 conditions earned by marginal generators is about $10,000/Mw-Year. The scarcity rents generated from selling energy and operating reserves during scarcity conditions (Op-4) are, on average, almost identical (as

[20] There are only five hours during this period (in May 2000) that are "trimmed" in this way, but the effect on scarcity rents associated with energy supplies (though strangely not operating reserves) is substantial. See the footnote to Table 1.

[21] The revenue effect is relatively small.

theory suggests they should be).[22] There is significant volatility from year to year in the rents earned, however. On average there are on average 46 hours per year when Op-4 is in effect and 32 hours per year when Op-4 step 11 is in effect. There is significant volatility in the annual number of operating reserve deficiency hours as well. On average there were only six hours per year when the price cap was binding, again with considerable year-to-year variation.[23] This suggests that the $1000 price caps are unlikely to be the primary source of the revenue deficiencies (more on this below).[24] There are other factors, at least partially related to the reliance on out-of-market instruments to manage reserve deficiencies, that are depressing spot prices during reserve deficiency conditions associated with the mechanisms used by the system operators to manage reserve deficiencies.

The $10,000/Mw-Yr average value estimated for scarcity rents in New England during this period can be compared with the fixed costs (capital amortization and fixed O&M) of a new combustion turbine that might be built to provide the systems "reserve capacity." This cost would be roughly $60,000 - $70,000/Mw-Yr in New England. Clearly, the scarcity rents are far below what would be necessary to attract a CT to the market to be available to supply operating reserves and energy only during scarcity conditions.

---

[22] However, the relationship between energy prices and operating reserve prices on an hourly basis vary from theoretical predictions, especially in 1999 when the operating reserve prices are often very strange.

[23] One must wonder if 1999 is just an unusual year, with the ISO and market participants learning how to operate within the new New England market arrangements. There are many more Op-4 hours than in other years, but only one hour when the energy price exceeded $1000/Mwh (and as I understand it no price caps were in effect). The scarcity rents are much higher than in other years.

[24] David Patton's analysis of the New York ISO cited above also suggests that the reliance on out-of-market mechanisms and associated discretionary behavior by the New York ISO during reserve deficiency hours plays a much more important role than do the price caps.

One might argue that this is the wrong comparison, since there are many other hours when these generators can earn scarcity rents. If this were true then the $10,000/Mw/Yr value is an underestimate of the true quasi-rents available to cover capital costs. However, I have examined all hours when the market price for energy exceeded $100/Mwh during this period and find that about 80% of the scarcity rents are earned during Op-4 conditions. Once they are in effect, the proposed SMD market power mitigation mechanisms should further reduce the rents produced for the marginal generators I focus on here outside of reserve deficiency hours by moving spot market prices closer to the textbook marginal cost in those hours. [25]

Another possible objection to this comparison would be that the total costs of a new CT is not the relevant benchmark for New England. Because New England has a lot of old conventional oil, gas and coal fueled steam-turbine generating capacity, the market clearing prices reflect their relatively high heat rates (say 11,000 Btu/kWh) during many hours of the year. CCGTs with much lower heat rates (say 7500 Btu/kWh) are attracted to the market and earn rents to cover their capital costs on the "spark spread" associated with the difference between their heat rates and the heat rates of the generators that clear the market, as well as from the scarcity rents I have identified. Under this scenario, CCGTs are inframarginal, but push older conventional steam plants higher up in the merit order. These old plants then can provide operating reserves during tight supply situations. In this scenario, the scarcity rents identified must be high enough to cover the fixed-O&M costs of the existing generators that will provide this reserve capacity so that they find it profitable to stay open and available to provide operating reserves. I am told

[25] James Bushnell and Celeste Saravia, "An Empirical Assessment of the Competitiveness of the New England Electricity Market," February 2002.

that the annual fixed O&Ms of an older fossil steam units is in the range of $20,000 to $35,000/Mw/Yr.[26] The scarcity rents that I have measured for New England are not high enough to compensate for these annual fixed costs either and absent an additional source of revenues these plants would simply be mothballed or retired permanently.

A third objection could be that the system is too reliable and that the shortfall in spot market revenues reflects excess capacity relative to consumer valuations of consuming more or less on the marginal during scarcity conditions. Resolving this question requires making assumptions about the appropriate value for V, a number that is very difficult to measure.[27] We can obtain some insight into this explanation by solving for the implied value of V in equation (2) above. If C_k is $60,000/Mw-year and H_s is 46 hours, then the implied value of V in equation (2) is about $1,300 per Mwh. If C_k is $30,000/Mw-year, the implied value of V is about $650/Mwh. If we focus instead on the Op-4 Action 11 hours (32 hours on average) then the implied values of V are $1,875/ Mwh and $937/Mwh respectively. While these implied values for V are below the limited number of estimates of the value of lost energy used in other countries (e.g. England and Wales during the 1990s, Australia today) to set price caps, the numbers are not directly comparable. Recall that V in equation (2) is defined as the marginal consumer's opportunity cost of consuming more or less <u>averaged over all reserve deficiency hours</u> and not just during the tiny number of hours when load is actually curtailed.[28] We would expect the implied value of V as defined here to be below the

---

[26] These older plants also typically face costly environmental mitigation obligations if they continue to operate and these costs should be factored in as well.

[27] Steven Stoft, <u>op. cit.</u>, Chapter 2-5.

[28] Australia now uses a value of lost load of about $5,800/Mwh ($AU 10,000/Mwh). The value of V as defined here should be lower since it is an average value for all reserve deficiency hours.

value consumers place on consuming more or less energy during the very small number of hours they are actually subject to significant curtailments. Accordingly, the implied values of V as defined here and the prevailing levels of reliability do not seem to be out of line with the limited evidence on consumer valuations.

It is often argued that the $1000 price caps are the primary problem leading to prices that are too low in spot markets for energy and operating reserves. If the relevant value of C_K is $65,000 and the value of R is $10,000 per Mw-Year than there is a shortfall of $55,000/Mw-Year. If there are six hours per year when the $1000/Mwh price cap is binding, based on historical experience, then the cap would have to be raised to about $10,000/Mwh (and be hit during these hours) to generate adequate revenue. This price exceeds all estimates of the marginal value of lost energy that I have seen used to set cap prices in other countries. Moreover, the implied values of V associated with current levels or reliability calculated above ($650/Mwh - $1,875/Mwh) average about $1,300/Mwh, which is not far out of line with the current price cap in effect in the Northeastern ISOs. <u>If</u> the price of energy actually rose to a binding price cap during scarcity conditions, as it should in theory, a price cap in the $650 - $1,875 range would yield sufficient scarcity rents to keep marginal generators needed to maintain current levels of reliability in business or to attract investment in additional capacity as the case may be. Moreover, an increase in the price cap of this order of magnitude (to $10,000/Mwh) would dramatically increase incentives for exercising market power, leading to much higher prices in hours when there is not true scarcity and more hours of contrived scarcity. It would be bad regulatory and economic policy to solve problems created by net revenue deficiencies during scarcity conditions by facilitating the exercise of market power in other hours. So, the primary problem here is not the price cap itself but that it is not hit during most reserve deficiency hours.

The conclusion that I draw from this analysis is that the spot hourly energy and ancillary services markets in New England have not provided scarcity rents that are nearly sufficient to make it profitable for reserve "peaking" capacity to enter the market through new investment or to continue operating consistent with conventional levels of reliability. These results are consistent with those contained in related studies done for PJM and the New York ISO. Whether or not there is too much or too little reliability is a more difficult question to answer definitively. However, these calculations reinforce the NOPR's conclusions that spot energy and operating reserve markets alone are unlikely to provide adequate incentives to bring forth enough generating capacity to maintain traditional reliability levels. [29]

Of course generators in New England, New York and PJM have another revenues stream: capacity payments associated with load serving entities capacity obligations.[30] And there is no shortage of generating capacity in New England, PJM or New York

[29] Data available from the PJM Market Monitor's annual "State of the Market" report provides information that can help us to define an upper bound on the measure of scarcity rents that I have produced for New England. These PJM reports calculate the net revenues earned from spot energy sales for units with different marginal supply costs. The values calculated for units with marginal costs greater than $50 and $100 respectively are upper bounds for the values that would emerge by applying the same methods to PJM as I applied to New England. They are upper bounds, because they include all hours during each year and not just scarcity hours and reflect rents earned in other hours when there may be some market power. It is evident that the energy market rents for high heat-rate units appear to be much higher in PJM than in ISO-NE. Nevertheless, even in this case, the average rents earned from the energy market are roughly 50% of the PJM target effective annualized capacity cost of about $63,000/Mw-year.

[30] In addition, as I have already discussed, the New England market frequently is cleared on the margin with generation from the large quantity of existing older oil/gas/coal fueled generating capacity with relatively high heat rates. CCGT capacity coming into the market could earn net revenues to cover capital costs during many "non-scarcity" hours from the spark-spread representing the difference between the heat rate of the old steam units that clear the market and define the competitive spot market price and the lower heat rates of the CCGTs. Accordingly, CCGT capacity expands more quickly than demand grows, the older steam capacity will be pushed higher up in the merit order and can contribute to reliability as long as these units can earn enough in scarcity rents to cover their fixed O&M costs and the costs of required environmental mitigation investments.

(except in New York City where investment in new generating capacity faces additional challenges) and in most other regions of the country (California and the Southwest being the primary exceptions) there is surplus capacity projected for several years into the future.[31] Indeed, inadequate investment in transmission capacity is a much more serious problem that limits imports of abundant regional supplies of generating capacity into New York City, Long Island, portions of Southern Connecticut and out of Maine, Rhode Island, and Southeastern Massachusetts to serve demand in other parts of the region.

To the extent that there is a regional generation resource adequacy problem here it is a problem associated with new investment that should come into the market several years from now to meet growing demand and the possibility that a large amount of older existing generating capacity will be retired prematurely because prices are not high enough during operating reserve deficiency hours to cover the fixed O&M costs and any avoidable costs of meeting tighter environmental regulations that this existing capacity faces.[32] It is these potential problems upon which any resource adequacy rules should focus, not the financial problems faced by some merchant generators who are suffering from low market prices due to the current abundance of generating capacity that has been added in the last three years. However, despite, the current supply/demand balance, these issues associated with long run resource adequacy should be confronted now, before there are resource shortages, rather than in the midst of a supply crisis when rationale decision making becomes more difficult, and recognizing that it takes a few years

[31] North American Electric Reliability Council, *Reliability of Bulk Power Supply in North America: Summer 2002*, May 2002.

[32] In this regard it should be recognized that the cost of capital applicable to merchant generation in the future is likely to be significantly higher than it was during the recent generation investment boom.

between the time a project goes into the siting and construction pipeline and when it comes out the other end and can begin supplying electricity.

TABLE 1

SCARCITY RENTS IN ISO-NEW ENGLAND[33]

Year	OP-4 Rents Energy MC = $50 ($/Mw-Year)	Op-4 Rents Energy MC =$100 ($/Mw-Year)	Op-4 Rents Operating Reserves ($/Mw-Year)	Op-4 Hours All	Op-4 Hours Step 11	Price Cap Binding Hours
2002	$ 5,070	$ 4,153	$ 4,723	21	21	3
2001	$15,818	$14,147	$11,411	41	37	15
2000[34]	$ 6,528	$ 4,241	$ 4,894	25	14	5
1999	$18,874	$14,741	$19,839	98	55	1
Average	$11,573	$ 9,574	$10,217	46	32	6

[33] Computation procedures are discussed in the text.

[34] There were five hours where energy prices exceeded the $1000 price cap in May 2000 before the caps were imposed. For four of these hours the average price was $6,000/Mwh. If we include the actual revenues earned during these five hours rather then capping them at $1000 the values for 2000 $/Mw/Yr would be $28,349 (MC = $50/Mwh) and $27,362 (MC=100). There was only one hour when operating reserve prices exceeded the $1000 price cap. The operating reserves revenues were $7,294/Mw/Yr in 2000 without imposing the $1000/Mwh cap.

6. Policy Responses

The NOPR has identified a real potential problem that is consistent both with the theoretical effects of the identified market and institutional imperfections and with the empirical evidence. The question then is what are the best policies to remedy the problems that have been identified? Obviously, the long run policy goal should be to fix the market and institutional imperfections that create the problem and this should be a continuing goal. However, this will necessarily take a considerable amount of time and some of the imperfections identified above may be very difficult to fix even in the long run. Accordingly, I believe that some type of resource adequacy obligation placed on LSEs is necessary. This being said, it should also be recognized that designing a good resource adequacy obligation and associated implementation and enforcement mechanisms is a very challenging task. And we do not want to create a set of resource adequacy rules whose cost exceed the benefits that consumers will receive. The nature and magnitude of the problems is also likely to vary from region to region. These considerations lead me to conclude that it would be best for the Commission to specify a set of principles that satisfactory resource adequacy programs should meet, but leave it to the states and regions to design specific programs that satisfy these criteria in light of the economic and institutional conditions they each face. I suggest that the Final Rule reflect the following basic principles for the design of acceptable resource adequacy programs:

a. Resource adequacy obligations should apply to all LSEs, including competitive retail suppliers, municipal and cooperative utilities. Otherwise, the rules will further encourage free riding by LSEs that are not required to adhere to them.

b. A satisfactory resource adequacy obligation policy must include generation and credible demand-response programs as eligible "resources." This will help to facilitate the development of an active demand side in wholesale markets, one of the primary sources of market imperfection.

c. Generation resources eligible for meeting an LSE's resource adequacy obligation must meet a transmission deliverability obligation as is now the case in PJM. It makes no sense to allow an LSE, say located in Connecticut, to meet its resource obligations with capacity, say located in Maine, if there are transmission constraints that make it impossible to export the generation from Maine and deliver it in Connecticut during the time periods when it is needed to meet peak demands. Moreover, it is my view that that a transmission deliverability requirements is likely to play a much more important role in stimulating transmission investment than is the availability of additional CRRs created by these investments.

d. The program should focus on providing incentives for all LSEs to meet resource adequacy obligations with resources that are available to supply energy or operating reserves or equivalent demand response during reserve deficiency hours when generating capacity "reserves" and demand-response are valuable. "Capacity" that cannot supply energy or operating reserves when they are needed to meet demand is worthless from a resource adequacy perspective.

e. The program must be compatible with the evolving and uncertain state of retail competition in the states that have or may implement it. States without retail competition are likely to already place resource adequacy obligations (explicitly or implicitly) on their LSEs as they always have and to have a regulatory framework that supports long term arrangements for resources. In the absence of retail competition, LSEs can forecast demand for their retail customers and acquire resources as they have done for a century. They can meet their obligations to their retail customers by building generating capacity or contracting in the wholesale market, depending on state regulatory rules and procedures. This is not the case in states with retail competition or in states where retail competition is likely to be introduced in the foreseeable future. The evolution of retail competition in those states that have implemented it has proceeded less smoothly than many had hoped and continues to be an uncertain work in progress.[35] A large fraction of customers in many retail competition states have not switched to competitive retailers, though they can switch on short notice, and continue to be served on default service tariffs of various types, many of which will expire in the next few years. Other retail customers have switched back and forth between competitive retailers and

[35] Presentation by Paul L. Joskow "Retail Competition in the United States," July 2002. http://econ-www.mit.edu/faculty/pjoskow/files/JOSKOW-R1.pdf.

utility default service tariffs. Competitive LSEs have no way to measure what their retail loads will be three years from now. Indeed, the competitive LSEs that are hoped to be operating in many states three years from now may not even be in the market today. Similarly, the uncertainties associated with the pace and direction of retail competition, default service obligations, and applicable cost recovery rules make it very difficult for distribution company-LSEs to forecast the loads they will serve or to make long-term resource commitments without taking on unreasonable financial risks.

f. The resource adequacy program should not create or rely on a traditional centralized regional "all source" resource planning program that defines LSEs' obligations and how they should procure generation and demand-response resources. The failures of these programs implemented in many states during the 1980s, and the high costs that they imposed on consumers, were a primary stimulus for introducing retail and wholesale competition. Returning to this central planning model would represent a major defeat for the evolution of well functioning competitive wholesale and retail markets.

g. The ITPs that the Commission envisions <u>should not</u> be responsible for procuring reserve capacity pursuant to long-term contracts and allocating the associated costs to LSEs in their regions. It is an extremely bad idea to allow entities that are not responsible for the financial implications of their decisions to negotiate contracts or make financial commitments on behalf of third parties. This would lead to even worse resource acquisition decisions than were made through centralized planning processes in the past. The affect of relying on ITP procurement and long term contracting is bound to lead to the wrong kinds of capacity being acquired at an excessive price. This approach would be a step backward from the commitment to creating competitive wholesale markets

h. RTOs/ITPs can play a useful role by providing a stakeholder forum for the development of a non-binding long-term *indicative* resource plan that puts together the best available information about future demand growth, generation investment, demand response programs, and potential resource development opportunities. This information would be very useful to market participants as they make investment decisions on both the supply and demand sides.

i. There should be an organized <u>voluntary</u> capacity *market* (which is distinct from a capacity *obligation*) where contracts with qualified generating capacity or demand response capabilities can be traded among LSEs and between LSEs and suppliers. The ability to trade capacity entitlements in competitive and transparent capacity markets will be especially helpful to LSEs in retail competition states to meet their obligations efficiently. If there are any price caps applied to these markets they should be much higher than the caps applied to spot market energy and operating reserves since LSEs can contract forward for resources to mitigate market power. These caps should be defined consistent with the small number of reserve deficiency hours that are expected to occur.

j. In this regard, it is important to distinguish between the attributes of a *resource adequacy obligation* program and any organized voluntary *capacity markets* that may be put in place to help LSEs meet their resource obligations. These are two distinct but related institutions that are often discussed as if they were one. Some of the criticisms of resource adequacy obligations have actually been criticisms of problems with the voluntary capacity markets that have been created in conjunction with them rather than with capacity obligations per se. We want both a well-designed set of resource adequacy obligations <u>and</u> a well-designed voluntary organized market in which buyers and sellers can exchange resources to facilitate efforts to meet these obligations efficiently.

k. The program must have credible enforcement mechanisms that are consistent with providing incentives to LSEs to meet their resource adequacy obligations as well as being compatible with the state of retail competition programs. The penalties for failing to meet resource obligations during operating reserve deficiency hours should be high and increase with the severity of the emergency. Providing for penalties that increase with the severity of a reserve deficiency would be consistent with the introduction of demand functions for operating reserves.

l. The attributes of the resource adequacy obligation protocols should not conflict with efforts to remedy the market and institutional imperfections that lead spot energy and operating reserve markets to fail to provide appropriate incentives to achieve reliability levels consistent with consumer preferences and valuations for reliability and price volatility.

m. The resource adequacy obligation protocols should not be influenced directly by the current financial difficulties faced by merchant generating companies. These suppliers bet their money and they took on the risks of market price fluctuations. Shifting the risks associated with changing market conditions from consumers to suppliers is one of the primary benefits of competition. Now that markets have turned against them, suppliers should not be looking to consumers for a bailout. Resource adequacy rules should not be a corporate welfare programs. Of course, a good resource adequacy program will necessarily benefit both existing and new generators to the extent that resources are scarce and have a positive market value during scarcity conditions.

n. Market monitors should be required to monitor the performance of the resource adequacy program and the progress of market design changes and market maturation and remedy the underlying market imperfections. Periodically, they should offer a recommendation as to whether or not the costs of market imperfections have been remedied sufficiently that the resource adequacy obligations can be relaxed or withdrawn.

7. Assessment of the Proposed Resource Adequacy Rules

As I understand them, the proposed resource adequacy rules and procedures have the following key components:

a. They would require all LSE's to demonstrate to the relevant ITP that they have forward (e.g. three years ahead) contracts for qualifying generating capacity and/or demand response capability to meet their forecast peak demand plus a reserve margin. Qualifying capacity resources would have to meet a transmission deliverability requirement. Regions would be given flexibility to define what the reserve requirement would be, subject to a minimum specified in the SMD. The ITP would verify that the LSE has met its forward resource adequacy obligation.

b. These obligations would be enforced primarily by creating incentives for LSEs to meet their resource oblications that would kick in when reserve deficiencies occur and through the enforcement of criteria (e.g. deliverability) for resources to qualify to meet the resource adequacy obligation. LSEs that failed to meet their forward resource adequacy obligations would be assessed a penalty for energy and operating reserves taken from the spot markets for energy and operating reserves during operating reserve deficiencies. This penalty might rise as the seriousness of the reserve deficiency increases. Second, if the reserve deficiency becomes so severe that load must be shed, LSEs that had not met their resource adequacy obligations would have their loads shed first.

There are several aspects of this NOPR's resource adequacy obligation proposal that are very sensible, even clever. The Final Rule should retain the portions of the proposal that would require (or at least encourage) the establishment of regional (or ITP) resource adequacy obligations applicable to all LSEs, allow both generation and demand response programs to satisfy resource adequacy obligations and the requirement that generation used to meet resource adequacy obligations satisfy transmission deliverability requirements.

The proposed resource adequacy rules also have a number of significant deficiencies. These include:

a. The requirement to enter into forward contracts several years before the power is to be delivered is not well adapted to states with retail competition programs and could serve as an additional barrier to the success of these programs. As I have already discussed, LSEs in these states are unable to predict accurately what their loads will be three or more years from now and competing retail suppliers that may be serving some of the customers now served by distribution utilities may not even be active in the region today. This requirement would effectively require utility distributor LSEs with default service obligations to take on contractual obligations that may end up being stranded costs if market prices fall in the future and their retail customers switch to competitive retail LSEs. At the same time, the requirement may serve as a barrier to entry to competitive LSEs since they may enter the market or expand their presence after the date at which resource obligations must be firmed up. Of course, they could go to the incumbent utility LSEs to acquire their "qualified" resources but this could raise difficult negotiating challenges. At least in retail competition states, applying forward contracting obligations in a fair and efficient way would be very difficult at the present time. It would be better to provide powerful incentives to enter into forward contracts for qualified resources well in advance of the date they may be needed, but to allow LSEs to make their own decisions by weighing the costs and benefits of contracting for resources sooner rather than later.[36]

b. In any event, as I read the NOPR, the forward contracting requirement is very easy to evade. An LSE can avoid buying energy and operating reserves out of the spot markets by entering into short-term (2-day ahead!) bilateral contracts with generators instead of buying in the ITP's spot markets, thereby eliminating any obligation to pay penalties for the failure to cover its resource adequacy obligations years in advance. And this is likely to be a profitable strategy both for LSEs that have not fulfilled their forward contract obligations and for generators. This is the case because proposed rules create a wedge between the price that generators receive if they sell in the ITP's spot

[36] The proposed forward contracting obligations are also likely to lead to costly and inefficient regional centralized all-source resource planning and procurement programs or to ITP procurement of forward contracts for resources and the distribution of the associated costs among LSEs. This would be an unfortunate movement away from competition and represent a solution that is more costly than the problem it is designed to cure.

market during reserve deficiencies and the spot market price plus the penalty that LSEs would otherwise be required to pay if they purchase in the ITP's spot market. Both the LSEs and the generators will be better off by bypassing the ITP's spot market.[37] As a result, the spot market purchase penalty components of the proposed enforcement mechanism for the multi-year ahead forward contracting obligations would be ineffective. It would be better to focus on improving the incentives to contract forward rather than applying a hard requirement that can be easily evaded.

c. Nor is the proposed "priority curtailment" rule, likely to be workable in retail competition states. In addition to the issues already noted, competitive LSEs will have retail customers spread throughout a distribution company's service territory. Except for the very largest customers, it is unlikely to be physically possible to curtail only the customers of a specific competitive LSE that has failed to meet its resource adequacy obligations in a timely fashion. Moreover, as a practical matter, even if LSE loads could be selectively curtailed in this way, the selective curtailment of specific LSE loads is unlikely to be politically credible during serious sustained load curtailment conditions.[38]

We have experience with capacity obligation and capacity markets in PJM, New York and New England. We should build on this experience, learning from the successes and the mistakes of the last several years. While the energy market alone has not provided adequate revenue to support the fixed costs of new peaking capacity in PJM, the market value of capacity traded in its capacity markets made up for almost all of the deficiency in 2001,[39] in 1999,[40] and accounted for a very large fraction of the scarcity rents realized in 2000.[41] So despite the problems experienced with the PJM ICAP market, it seems to have done its job as far as providing an additional revenue stream to

[37] Bilateral contract energy prices would still be quite high under these conditions reflecting the tight supply situation. However, if the energy market is competitive, the bilateral contract prices would equal the expected spot market prices for energy and operating reserves (which are subject to a price cap) during the expected reserve deficiency conditions. In this way, LSEs could avoid paying the penalties even if they failed to contract forward to meet their resources adequacy obligation two or three years in advance as required by the proposed rule.

[38] If the Town of Concord's (MA) municipal light department were to fail to meet it's forward contracting obligations, is it credible that the New England ISO could meet a modest aggregate load curtailment need by turning off all of the power in Concord on a very hot July 4 when hundreds of thousands of people are in Concord to celebrate the anniversary of the shot heard 'round the world? I doubt it.

generators.[42] Accordingly, it seems to me that the focus of any rule should be on reforming the ICAP obligations and associated procedures and markets that exist in the Northeastern ISOs, rather than simply rejecting them and adopting a new system with which we have no experience.

At the same time, more attention needs to be devoted to an examination of ITP procedures during reserve deficiency hours to improve spot market incentives. Many reserve deficiency management costs on both the supply and demand sides that are now "out-of-market" should, as far as is reasonably possible, be reflected in spot market prices for energy and operating reserves in these ours. Bid mitigation procedures which depress prices too much during reserve deficiency hours also need further evaluation and possibly reforms.

[39] *PJM Interconnection State of the Market Report 2001*, page 18.
[40] *PJM Interconnection State of the Market Report 1999*, page 9.
[41] *PJM Interconnection State of the Market Report 2000*, page 15.
[42] Obviously, vertically integrated LSEs which have retained their generating capacity to meet their retail default service obligations, might not separately record revenues and expenses from the sale and repurchase of capacity under their retail service transition agreements unless they are provided for specifically in these agreements.

III. TRANSMISSION INVESTMENT, PARTICIPANT FUNDING AND CRRs[43]

1. Summary of Conclusions on Proposed Transmission Investment Framework

I agree with The NOPR's conclusion that " [c]ompetitive and reliable regional power markets require adequate transmission infrastructure to allow geographically broad supply choices and minimize the complications created by loop flow." (¶ 335) I also agree that over the last several years there has been inadequate investment in transmission capacity to reduce congestion and to support robust competitive wholesale markets for electricity (¶ 191) and that it is essential to remove any cost recovery impediments that act as barriers to the development of new transmission capacity ((¶ 196). Finally, I agree that it is desirable, to the extent that it is practical and does not create additional barriers to transmission investment, to match cost responsibility to the beneficiaries of transmission upgrades (¶ 197).

However, the framework that is proposed in the NOPR as the foundation for stimulating transmission investment (¶335-351) has very serious deficiencies and will not achieve the Commission's goals. If it is implemented as proposed in the NOPR, it is more likely than not to reduce the pace of investment in new transmission capacity rather than to increase it. This will lead to growing congestion, increasing market power problems, and growing demand for ever more locational market power mitigation[44] and

[43] This section draws heavily on ongoing research with Jean Tirole. However, only I am responsible for the views presented here.

[44]The NOPR proposes to require that under certain "non-competitive conditions" (e.g. local market power problems caused by congestion) generators be required to offer all available energy (must-offer requirement) to the system operator subject to a pre-specified bid cap. FERC Docket No. RM01-12-000, Notice of Proposed Rulemaking, July 31, 2002, ¶ 409. It also invites ITPs to propose additional mitigation measures that could apply under certain conditions where market power would be a significant problem, id.

other regulatory interventions that will frustrate the spread of wholesale and retail competition and associated restructuring initiatives to more states.

The NOPR's approach seems to be based on the assumption that we can rely primarily on "private initiative" to bring forth needed transmission capacity and views "market driven" decisions as the "fundamental mechanism" to provide efficient levels of transmission investment. Thus it appears that the Commission has in mind "merchant transmission projects" that would be supported financially through congestion revenues (the difference in nodal prices) and the sale of CRRs that reflect the market value of congestion revenues as the foundation of its transmission investment framework. (¶346-347). This framework appears to view alternatives to merchant transmission projects as secondary or tertiary complements to fill modest gaps in transmission investment needs that are not otherwise provided by merchant investors. If merchant transmission investment falls short of meeting identified needs the next step would be to seek investment through an ITP managed RFP/competitive bidding process. Regulated incumbent TOs would then play only a supplier of last resort role. In all cases, investment would be mediated through a regional transmission planning process, with merchant projects permitted to go forward first as long as they satisfied feasibility constraints and did not infringe on existing rights.

I will focus my Comments here on the assumption that merchant transmission investment is likely to make a significant contribution to efficient transmission investment needs. In my view, this assumption, and the vision of a transmission

---

at ¶ 415. Finally, the NOPR provides for a regional "safety net bid cap" that would apply to the day-ahead and real time markets under all conditions, id. at ¶ 433.

framework dominated by merchant transmission projects, is a fantasy (or perhaps an economist's dream) that fails to take account of important economic and physical attributes of transmission investments, and is inconsistent with sound economic theory and available empirical evidence. The Commission has taken a leap from the correct observation that merchant transmission investment may be able to provide <u>some</u> of the needed investment in new transmission capacity to the erroneous conclusion that it can be relied upon to provide <u>most</u> of the efficient investments in transmission capacity. There is little, if any, theoretical or empirical support for this leap of faith contained in the NOPR. In my view, the NOPRs transmission investment framework is based on a poorly developed and immature theoretical foundation and has little if any empirical support. The proper theoretical analysis and the limited empirical evidence that does exist suggest that merchant transmission will play only a limited role in meeting the nation's transmission investment needs. Adopting an untried transmission investment framework with such an unsatisfactory theoretical or empirical foundation would be very risky and carry potential costs that far exceed the expected benefits compared to readily available alternatives.

In Comments that I submitted to the Commission over three years ago I concluded that:[45]

> "My optimism about relying primarily on private market-based initiatives has waned with the experience with restructuring in the US and other countries over the past few years. *Indeed, proceeding under the assumption that at the present time "the market" will provide needed transmission network enhancements is the road to ruin.* There is abundant evidence that market forces are drawing tens of thousands of megawatts *of new generating capacity* into the system. There is no evidence that market forces are drawing significant entrepreneurial investments

[45] Comments of Professor Paul L. Joskow, Notice of Proposed Rulemaking, Regional Transmission Groups, Docket No. RM99-2-000, August 16, 1999.

> into new transmission capacity. While third parties should be given the opportunity to propose market-based private initiatives to expand transmission capacity, incumbent transmission owners, in the context of a sound RTO/ISO planning process, must be relied upon to play a central role in expanding the transmission system. Increases in transmission capacity that their initiatives create, and the associated transmission rights that conform to the protocols being applied in their regions, could then be auctioned off to market participants with the proceeds used to help to defray the costs of the transmission network.[46] …
>
> These observations do not mean that third-parties should be precluded from making proposals for transmission upgrades for consideration by transmission owners, RTOs, and regulators. I simply would not assume that we can depend on these market-based initiatives at the present time to produce the most cost-effective enhancements to transmission networks necessary to meet reasonable economic and reliability goals. The transmission owners operating through a sound RTO/ISO transmission planning process should be expected to be the primary, but not necessarily the exclusive, source of network enhancement initiatives."

Three and one-half years of additional experience has convinced me that these conclusions are even more valid today than they were over three years ago. The problems associated with stimulating transmission investment to support the successful development of competitive wholesale and retail markets for electricity that I identified in 1999 have, unfortunately, been realized. Transmission investment continues to lag and inadequate transmission infrastructure is increasingly a barrier to creating robust regional competitive wholesale electricity markets. After all of the talk about merchant transmission investment we have exactly one merchant transmission project, with attributes quite different from the theoretical foundations supporting the merchant model,[47] that is nearing completion and operation in the U.S.[48] There are two other

---

[46] Paul L. Joskow, op. cit. 1999. "It is important that the regulatory framework assure that the transmission owner does not profit directly by increasing the value of these rights so that it does not have an incentive to increase congestion."

[47] This is the Cross Sound Cable DC link connecting Southern Connecticut with Long Island. While the details of the arrangement remain confidential, it appears that the link is supported by a long term contract with one municipal monopoly distribution company that faces serious resource constraints, has de facto exclusive use of the link to contribute to its energy and reliability needs, and can recover the costs of the

merchant transmission projects that have been completed in Australia, one of which sought regulatory protections almost immediately after commencing operations (more on Australia below). This is a rather thin reed upon which to base a transmission investment policy framework affecting a critical infrastructure sector for the entire country. Indeed, the Commission's continued flirtations with the merchant transmission model and a variety of additional conceptual flaws that flow from it (e.g. the focus on rules that force transmission and generation "to compete" with one another) have increased the barriers to adequate and timely investment in transmission infrastructure and have contributed to the growing costs of transmission congestion and the increased disaffection in many states with expanding wholesale and retail electricity competition and implementing supporting restructuring initiatives.

Instead of relying on merchant transmission to provide the primary governance framework for transmission investment, the Commission should pursue a regional transmission investment framework that is better matched to the economic realities of transmission investment and has a track record of good performance. The one internationally proven way to stimulate transmission investment is to rely primarily on incumbent regulated transmission owners (TO) operating under sound regional planning guidelines and subject to well-designed performance-based regulatory mechanisms to be the primary vehicle for building, financing and operating transmission facilities. As part of this framework, ex ante criteria, based on a practical set "beneficiaries pay" principles,

---

long term transmission contract from its captive retail customers. As noted below, I do not consider the proposed Path 15 upgrade to be a merchant project.

[48] The Cross Sound Cable project apparently continues to face environmental compliance hurdles in Connecticut before it can commence regular commercial operation. "Connecticut Regulators Reject Cross Sound's Proposal for Burying Cable," *Electric Transmission Week,*" January 6, 2003.

would be used to apportion cost responsibility, rather than the unworkable case by case negotiation process that appears to be a key component of a merchant transmission model. This implies, in turn, that it is important for the Commission to include in the Final Rule a requirement that ITPs work with TOs to develop performance based regulatory (PBR) mechanisms applicable to all TOs, whether or not they are independent transmission companies (ITCs). I will also discuss below why the Commission's one modest effort to include an incentive provision in the NOPR, by placing the full burden of congestion revenue shortfalls on TOs, is flawed and is likely to be confiscatory as written. The approach to transmission planning and investment that the Commission accepted in its recent order regarding RTO West is a major step in the right direction that is consistent with the framework that I support since it gives primary responsibility to incumbent TOs to upgrade and expand RTO network transmission facilities mediated through a regional planning process.[49]

I recognize that in the U.S., unlike most other countries that have moved to competitive wholesale and retail markets, vertical integration between transmission, generation, and marketing businesses complicates the challenge of stimulating efficient transmission investment by TOs. However, at least at this point in time, vertical integration is not the primary barrier to investment in transmission capacity. If it were, we would see significant differences in transmission investment behavior between regions where most of the generation has been divested (e.g. New York, New England, California) and those where vertical integration continues to be the norm. I have seen no evidence that transmission investment is more robust in the former areas than the latter. Rather, transmission investment has lagged everywhere.

[49] *Avista Corp., Bonneville Power Admin, et. Al.* 100 FERC ¶ 61,274, at 209-234 (2002).

Moreover, a growing number of Independent Transmission Companies (ITCs) are emerging in the U.S. and, with continued encouragement and incentives from the Commission, I expect to see many more in the near future. The industry may look very different a few years from now than it does today and the Final Rules should anticipate and encourage these changes. Even in the presence of vertical integration, ITP oversight, market monitoring, and properly designed incentive regulation mechanisms should be able to mitigate the most serious potential problems arising from vertical integration. If the Commission has evidence that vertical integration is a continuing barrier to transmission investment and that it does not have the regulatory tools to mitigate it, then I suggest that the Commission redouble its efforts to provide incentives to create ITCs rather than turn to an inferior merchant transmission model that cannot be expected to work well in practice.

The approach that I recommend <u>does not</u> mean that merchant transmission projects should not be permitted if they are proposed voluntarily and are consistent with the relevant regional expansion plan. Rather, it simply recognizes the practical limits to merchant transmission investment and some of the problems that adopting a framework that relies on it creates. I believe that the most likely opportunities for merchant transmission investment (as defined more precisely below) will be in situations where one or more of the following conditions are satisfied: (a) there are very large expected <u>long term</u> differences in prices between two regions that are, as a result, effectively separate economic markets, (b) where there are significant constraints on the expansion of economical generation within the relevant import constrained region, (c) where generation supplies are expected to be relatively cheap and abundant in the relevant exporting region over the long run, (d) where there is a large buyer in the importing region or a large seller in the exporting region which can internalize the benefits of merchant link and is willing to enter into a long term contract for the bulk of the merchant transmission capacity and (e) where a self-contained controllable transmission link (DC) of a size that represents a relatively small fraction of demand in the importing region is well-adapted to connecting the two market areas. The Cross Sound Cable project, the Basslink project in Australia, and the proposed New Jersey to New York City Harbor Cable project are consistent with these attributes. As I will discuss presently, the bulk of the transmission investment needs that have been identified by ISO planning processes are not.

2. Definitions: Merchant Investment, Participant Funding, Rolled-in Pricing

Let me start by defining a variety of terms that are being used in the discussion of transmission investment. I believe that there is a lot of confusion in this discussion and that clear definitions are required to have a constructive dialogue on the issues.

a. Merchant transmission investment:

These are investment projects whose costs and financial performance are fully underwritten by the owners of the project. That is, capital and operating costs associated with merchant projects cannot be recovered in a FERC regulated tariff. Instead, the owners of a merchant project recover their investment, maintenance and operating costs through the sale of transmission service at market-based rates to LSEs, generators, retailers, wholesale marketers and other financial intermediaries.[50] Of course, to the

[50] I do not consider the proposed upgrade of Path 15 in California to be a merchant project. The participants have requested and the Commission has approved cost-of-service rate treatment for this joint public-private project. "The Cost of the project would be eligible for recovery as part of the Project Proponents .… Transmission Revenue Requirement (TRR) through the ISO High Voltage Access Charge

extent that the transmission service (including CRRs) is acquired from merchant suppliers by a regulated LSE (i.e. a distribution company), the associated costs must ultimately be recovered from retail consumers through a regulated retail tariff. (It is not clear to me how this will be accomplished if the Commission excludes these costs from it's regulated tariffs and preempts state regulation of retail transmission service, but I will leave this to the lawyers to figure out.)

Within the basic SMD market framework proposed in the NOPR, merchant transmission investment will be forthcoming if and only if the expected spot congestion revenues or revenues from sales of CRRs reflecting the expected value of future spot congestion revenues is greater than or equal to the expected capital and operating costs of the project. So, the fundamental question about the viability of merchant transmission is whether or not an efficient level of transmission investment will be forthcoming if investors must rely on congestion revenues and more generally the market value of CRRs to compensate them for the capital and operating costs of the project.

b. Participant funding:

Participant funding is one of the more mysterious and confusing phrases being used in the discussion of transmission investment. It is especially confusing because in competitive retail and wholesale markets, everyone is a market "participant" directly or indirectly and the effects of transmission investment are widely dispersed over time, over the participants who gain and lose from the investments, and over the nature and

…." Memorandum from Armando J. Perez to ISO Board of Governors, dated June 20, 2002. http://www.caiso.com/docs/2002/06/20/2002062017315212815.pdf. It should be noted as well that the incumbent TOs (WAPA and PG&E) are necessary participants in the project because the enhancements require the upgrading and use of their existing facilities and transmission corridors. http://www.caiso.com/docs/2002/06/24/2002062415320629050.pdf.

magnitudes of the benefits received. If participant funding is simply another phrase to describe merchant investment where funding is arranged through project-by-project negotiations with market participants and intermediaries, we can avoid confusion by dropping the participant funding concept entirely and simply refer to merchant investment. Otherwise, the term participant funding must be defined more precisely than it has been so far.

I believe that participant funding is best conceptualized as defining the criteria that will be used to determine "who pays" for transmission investments whether or not the payments come (fully or partially) through regulated transmission tariffs. The basic concept advanced in the NOPR that seeks to assign payment obligations to those who benefit from a transmission investment makes very good sense. The critical question that must be addressed is whether the magnitude of the benefits and their distribution to specific groups of beneficiaries will be negotiated on a case-by-case basis for each proposed transmission project or whether a set of reasonable ex ante cost responsibility criteria will be used (perhaps as a default) to apply to all transmission investments. As I will discuss, the ex ante specification of criteria for apportioning cost responsibility will be much more successful in bringing forth needed transmission investment than will case-by-case negotiations, which I believe are unworkable except in special cases.

c. Regulated Transmission Investments, Rolled-in Pricing and Participant Funding

These are transmission investments whose capital and operating costs are included in a TO's regulated transmission tariff. However, just because a transmission investment is included in a regulated transmission tariff does not and should not mean that "rolled in" pricing must be used to recover all of the associated costs from the retail customers of the LSEs in the TO's area where the new transmission facilities happen to be located. There is no reason why reasonable ex ante criteria for apportioning cost responsibility based on "beneficiaries pay" principles cannot be applied to regulated transmission investments. Of course, there will be many circumstances where the primary beneficiaries of a transmission project are the retail customers of the LSEs in a TOs area and it is appropriate in such cases to include the costs in the TOs regulated license-plate rates applicable to LSEs in its area.

3. Alternative Approaches to Participant Funding

a. Case-by-Case Negotiations Is Unworkable and A Barrier to Transmission Investment

In my view, the successful implementation of the "beneficiary pays" principle based on a negotiated case-by-case or project-by-project basis will be extremely difficult. Relying on case-by-case resolution of participant funding obligations is likely to lead to an extended process of debate and haggling about who should pay for what and when. The ultimate result of this approach will be that transmission investment will be delayed or retarded by free riders who would prefer not to pay and by market participants who would prefer to continue to benefit from congestion.

Why will a case-by-case negotiated apportionment of cost responsibility be so difficult? There are several important reasons. The benefits created by transmission investment will typically accrue to a variety of market participants and are likely to be widely dispersed among many market participants. These beneficiaries may be distributed across multiple TOs and ITPs. Any negotiation with so many potential beneficiaries will be time consuming and involve significant transactions costs. Moreover, it will be very difficult to identify who the relevant beneficiaries are who should be at the negotiation table. The distribution of benefits can vary widely over time as supply and demand conditions change so that identifying who benefits and by how much is very difficult.[51] Since transmission investments are long-lived, identifying the distribution of beneficiaries and quantifying the benefits over the life of a particular project is especially difficult at the project development stage. In addition, many transmission investments provide both economic and reliability benefits so that these investments are characterized by joint costs and the associated challenges of cost allocation.

The presence of "lumpiness," economies of scale, and economies of scope associated with some transmission investments further complicates the problems of bringing beneficiaries to the table to negotiate payment responsibilities due to free rider problems.[52] When a transmission investment of any significant size is completed it will lead to lower prices for consumers in the constrained import zone and to higher prices for suppliers. These benefits are received by consumers and suppliers at these locations whether or not they agree to help to pay for the investment --- there is no feasible way to exclude all consumers in a constrained import area from receiving the benefits of lower

[51] The word "time" here should be construed very broadly. Within a given year, a transmission project may provide benefits to different sets of market participants. For example, during the Spring hydro runoff in the Northwest, energy flows from North to South. During the winter months it may flow in the other direction and cheap generation in the Southwest allows water to be stored in the Northwest. Over longer periods of time, the use of transmission enhancements will change as well. For example, the DC link between France and England was originally designed to exploit diversity in demand patterns between France and England, with exchanges in both directions expected to be the norm. During the 1990s, the line was used almost entirely to import power from France to England. During recent European cold snaps, however, the line has been used to supply emergency power to France from England to maintain reliability when generating plants in France suddenly went down. Moreover, the basic regulatory and market framework within which this project operates has changed over time. It was built as a regulated cost-of-service project. European Union regulations recently required that the English Channel interconnector to rely entirely on competitive auction revenues to cover its costs.

[52] James Bushnell and Steven Stoft, *Grid Investment: Can a Market Do the Job?*, The Electricity Journal, January/February 1996, Volume 9, Number 1 at p. 77.

prices whether they pay for the transmission investment voluntarily or not. Every one of these "beneficiaries" (as well as customers that are not located within the importing area but benefit in some other way from the investment) has the incentive to count on others to pay for the project --- to "free ride" --- since they cannot be excluded from receiving the benefits once it is built. Efforts to apply sophisticated participant valuation aggregation procedures to allocate cost responsibilities for major new transmission lines in other countries have worked poorly, with consumer interests receiving especially poor representation.[53]

The benefits created by transmission investments will also take a variety of forms: lower wholesale and retail prices, higher profits, increased reliability, all of which will accrue to many different market participants. Indeed, the sharp distinction between "reliability" and "economic enhancement" investments contemplated in the NOPR is difficult to square with reality. Many transmission investments provide both economic (lower wholesale prices) and reliability benefits. Moreover, the distribution of economic and reliability benefits can vary from hour to hour, season to season, year to year, as supply and demand conditions change and, from the perspective of an investor in a project with a life of several decades, are highly uncertain.

b. Simple Clear *Ex Ante* Cost Responsibility Criteria are Needed

A significant barrier to transmission investment today is the lack of clear, simple, and specific ex ante criteria that specify who must pay for transmission investments, consistent with the economic attributes of these investments (more on this presently) and the complex and uncertain distribution of benefits that accrue from these investments

[53] Omar O. Chisari et. al., "High-tension electricity network expansion in Argentina: decision mechanisms and willingness-to-pay revelation," *Energy Economics*, 23, 2001, pp. 697-715.

over time and space. The Commission's focus on merchant investment, case-by-case negotiations on participant funding obligations and on transmission and generation "competing," have created further confusion and disincentives to invest. ISOs are finding it necessary to respond to this gridlock, and the growing problems created by inadequate transmission investment, by defining virtually all investments identified in regional transmission plans as "reliability investments."[54] Accordingly, the Commission should give a high priority to defining a set of fair and simple ex ante rules that specify criteria for determining who will pay for transmission investments identified as needed in regional plans. A beneficiary pays principle can and should be used to develop these criteria. However, it must be applied in a simple and practical way through ex ante cost responsibility criteria, recognizing the difficulties of precisely determining who the beneficiaries are over time. For example, by selling CRRs created by a regulated transmission investment, and crediting the revenues received against the costs of the projects, those who benefit from the value of hedging congestion price volatility will pay for a portion of the transmission investment directly.[55] Requiring generators to pay for interconnection facilities is yet another source of revenue. If the investment is justified primarily by the need to export energy from a region with a surplus to a region with a deficit, the exporting generators and the importing customers should be responsible for paying the bulk of the costs. The NOPR already contemplates applying criteria like this in connection with its proposals to eliminate embedded cost-based rate pancaking

[54] The 33 transmission projects identified as being needed by the New England ISO in its most recent regional transmission expansion plan are almost all identified as "reliability" projects. ISO New England, RTEP02, November 7, 2002. Yet it is fairly clear that many of these projects have "economic" benefits as well. Of course, if barriers to transmission investment are not removed soon, many more transmission projects will be needed to maintain reliability.

[55] As I will discuss presently, an appropriate incentive regulation mechanism must also be applied to ensure that transmission owners do not profit from creating transmission congestion to enhance the value of CRRs.

associated with inter-regional transfers (¶179 - ¶189). This will be a lot easier to accomplish for costs associated with new transmission investments than it is for the embedded costs of the existing infrastructure. There will be many cases where transmission investments will ultimately accrue to the benefit of the retail customers in a TO or ITP's area where the new facilities are built. These costs can be rolled into license plate rates applicable to the LSEs whose customers benefit from them.

4. Attributes of Real Transmission Investments Must Be Included in the Theoretical Models upon which the Commission Relies

The NOPR contains what I consider to be, at best, an incomplete characterization of the attributes of transmission investment opportunities. The conceptualization of transmission investments in the NOPR focuses on major new transmission lines that would expand the footprint of the transmission network and often must confront a very challenging state-controlled facility siting approval process. While, there certainly are transmission investment opportunities that have these attributes, there are many that do not. The failure to consider the full range of transmission investment opportunities leads to a flawed framework for stimulating transmission investment.

There are many potential opportunities to increase the capacity of transmission networks other than by building major new lines involving new rights of way and expansion of the network's footprint. They vary from no- or low-cost upgrades of the reliability of breakers and other components on the network, better monitoring, communication and control capabilities, to more costly investments in static var compensators, capacitors, substation enhancements, FACTS technology, and reconductoring of existing transmission lines. These types of investment opportunities are typically intertwined with and inseparable from the incumbent TOs' transmission networks from a physical, maintenance and operating perspective.

This expanded characterization of the attributes of transmission investment opportunities is consistent with the 33 priority projects identified in ISO New England's recently released 2002 Regional Transmission Expansion Plan.[56] Of these 33 projects, 29 are projected to cost less than $20 million each and have some or all of the attributes I have just listed. Another project has a cost of $40 million. Indeed, all together these 30 projects account for only $163 million of the $888 million estimated total cost of the entire 33 project program. That is, three projects account for the bulk of the costs. All three projects (all of which have been designated as reliability projects) involve significant enhancements to the existing network, while two of them also anticipate building new 345 Kv loops. Few if any of these 33 real transmission projects are well represented by an economic model that assumes that transmission investment involves building "stand-alone" transmission lines on new corridors from point A to point B that require simple interconnections with the existing network at each end. There may very well be some projects with these attributes, but they are not representative of the full range of transmission investment opportunities and are the ones that are most likely to run into siting problems by expanding the network's footprint.

For these reasons, in evaluating alternative transmission investment frameworks it is useful to think conceptually about (at least) two types of transmission investments opportunities that can increase the capacity of the network.

[56] ISO New England, 2002 Regional Transmission Expansion Plan (RTEP02), November 7, 2002.

a. Network deepening investments:

These are investments that involve physical upgrades of the facilities on the incumbent TO's existing network (e.g. adding capacitor banks, phase shifters, reconductoring existing transmission links, upgrading substation facilities, new communications and relay equipment spread around the network to increase the speed with which the SO can respond to sudden equipment outages and relax contingency constraints). These are investments that are physically intertwined with and inseparable from the incumbent TO's facilities. These investments are firm specific investment opportunities that can be undertaken most efficiently by the incumbent network owner and (physical) operator responsible for maintenance and other physical operating tasks.

b. Network expansion investments

These are investments that involve the construction of separate new links that are not physically intertwined with the incumbent network except at the point at either end where they are interconnected. These investments can (in principle) be made either by incumbent transmission owners, by stakeholders (generators, load-serving entities), or by a third-party merchant investor. The two operating DC merchant links in Australia appear to fall into this category.[57] However, as in Australia, these links may have effects on power flows on the rest of the network, including on parallel lines, though they are physically separable projects from a construction and maintenance perspective.

[57]Two merchant lines supported by differences in spot prices in the two market areas they connect have been placed in operation under this arrangement in Australia. Directlink is a 180 Mw, 40 mile merchant DC link connecting Queensland and New South Wales and began operating in 2000. Murraylink is a 220 Mw, 108 mile merchant DC link connecting South Australia and Victoria which began operating in October 2002. On October 18, 2002, Murraylink applied to the regulatory authorities in Australia to change its status from a merchant line to a regulated line that would be compensated based on traditional cost of service principles combined with a performance incentive mechanism. Neither merchant link appears to be profitable. As far as I can tell, these are the only two merchant transmission lines operating anywhere in the world that have been built in anticipation of recovering their costs entirely from congestion rents arising from the difference in nodal prices between the areas they connect.

5. Merchant Transmission Projects Have Not And Will Not Play A Major Role in Providing Transmission Investments Needed to Benefit Consumers

As previously discussed, merchant transmission investments (or "market driven" transmission investments) will ultimately be supported from congestion revenues and the associated market value of CRRs. As I will discuss, the theory upon which the view that we can rely primarily on merchant transmission investment in a market environment with LMP and CRRs is simplistic and flawed. Moreover, there is no empirical evidence to suggest that a merchant transmission model can succeed in providing the bulk of efficient transmission enhancements that are socially beneficial. The limited empirical evidence available is inconsistent with this assumption. Accordingly, building a transmission investment framework on the assumption that merchant or "market driven" transmission investment can and will play a major role in filling the gaps in the capabilities of the transmission network would be a very serious mistake that will cause significant harm to consumers.

On its face, the proposition that merchant transmission investment supported by revenues from CRRs provides a sound framework to govern transmission investment is <u>inconsistent</u> with the Commission's conclusion that there has been inadequate investment in transmission capacity. Recall, that merchant transmission investments must pay their way from revenues earned from congestion payments; at least that's what the theory tells us (more on the theory presently). If the theory were correct, the assertion that there is underinvestment in transmission would imply that we would find that transmission congestion revenues are high enough today to support both the amortization of the capital costs of the existing network plus the associated operation and maintenance expenses <u>and</u> that there would be some surplus congestion revenues left to support the new investments that the Commission has concluded are needed. I recognize that the Commission is not proposing that existing network investments will be paid for from transmission congestion revenues. However, if the merchant investment theory is correct and there is underinvestment it should also be the case that there would be adequate congestion revenue to support the costs of the existing network if we chose to pay for it in this way. Another way of thinking of this is to think of building the current network from scratch based on a merchant model and then asking the question of whether this network could be financed from congestion revenues as the merchant model proposes.

We can look at PJM, which has LMP and CRRs similar to what the Commission is proposing in the SMD, to see the inconsistencies between the merchant investment theory and the Commission's conclusion that there is a significant need for additional transmission investment. The original cost of transmission investment by the TOs in PJM is about $5.8 billion.[58] Let's assume, conservatively, that the cost of building this network today would be twice its original cost, or about $12 billion. With a 12% amortization rate (return on equity investment, interest charges on debt, depreciation, property and income taxes) this would yield annual capital costs of $1.4 billion per year. In addition, the TOs incurred about $100 million in maintenance costs in 2001 for a total annual capital and operating cost of a nice round $1.5 billion per year.[59] In 2001 there were $271 million of congestion costs in PJM.[60] Accordingly, congestion revenues would have covered less than 20% of the capital and maintenance costs of the PJM

[58] These numbers have been taken from the 2001 FERC Form 1 for each of the TOs in PJM at that time.

[59] There are other operating costs that I have not included because the Form 1 data for transmission operating costs includes expenses for purchases of transmission service from others. To avoid double counting I simply excluding all operating costs other than maintenance costs.

[60] PJM Interconnection, *State of the Markets Report 2001*, page 11.

network if it had been built on a merchant basis. This 20% recovery figure is consistent with simulation studies that have been performed by others.[61]

There are a number of conclusions one might draw from these numbers. One could conclude that the PJM network is overbuilt and that there has been too much transmission investment in the past. This is, of course, inconsistent with the Commission's conclusion that there has been too little transmission investment and that there is a need for more. Alternatively, one might conclude that the bulk of the transmission investment (i.e. 80%) is for interconnection and reliability purposes and should not be expected to be built on a merchant basis. If this is the case, then it also implies that in the overall scheme of things, merchant investment opportunities represent a small piece of the overall transmission investment challenge and that we have focused too much attention on it in the NOPR. Yet another alternative is that the transmission network is characterized by very significant economies of scale or "lumpiness." As I will discuss presently, lumpiness undermines both efficient merchant investment and a project-by-project approach to participant funding. Finally, one might conclude that the merchant investment model is itself flawed. My analysis leads me to conclude that there are good reasons to believe that it is flawed.

Nor does the experience in Australia where a mixed model that provides for both merchant and regulated transmission investment paths suggest that the merchant investment model is particularly promising. As I have already indicated, there are only two small merchant lines supported by differences in spot prices in the two market areas they connect that have been built and placed in operation in Australia under the type of

[61] Perez-Arriaga, I. J. et. al., "Marginal Pricing of Transmission Service: An Analysis of Cost Recovery," *IEEE Transactions on Power Systems*, 10(1), 1995, pp. 546–553.

merchant model proposed in the NOPR. Directlink is a 180 Mw, 40 mile merchant DC link connecting Queensland and New South Wales which began operating in 2000. Murraylink is a 220 Mw, 108 mile merchant DC link connecting South Australia and Victoria which began operating in October 2002. On October 18, 2002, almost immediately after its completion, Murraylink applied to the regulatory authorities in Australia to change its status from a merchant line to a regulated line that would be compensated based on traditional cost of service principles combined with a performance incentive mechanism.[62] Neither merchant link appears to be profitable. Moreover, in Australia, this mixture of merchant and regulated transmission investment paths has led to extensive litigation between proponents of regulated and merchant transmission links, delaying investments in both. The Australian experience is hardly a poster child for relying on merchant transmission investment to meet transmission investment needs in the U.S. and indicates how difficult it will be to mix regulated and merchant investment options together as well.

6. Flaws in the Merchant Transmission Theory

The NOPR provides little in the way of theoretical or empirical support for relying primarily on merchant transmission investment as the foundation for its transmission investment framework.[63] There is a small academic theoretical literature that envisions a combination of spot LMP and point-to-point transmission revenue rights

[62] Murraylink Transmission Company on Behalf of Murraylink Transmission Partnership, "Application for Conversion to a Prescribed Service and a Maximum Allowable Revenue for 2003-2012," October 18, 2002.

[63] The only reference to the literature in the NOPR that is directly on point is the paper by Hogan cited in footnote 116. The Schweppe et. al. book cited in that footnote contains a brief discussion of the mathematical relationships that define efficient transmission investment, but it is fairly clear that it does not contemplate the merchant investment model the Commission favors; it focuses only on marginal investments, is most relevant to enhancements to existing lines rather than to investments in new lines, and ignores cost recovery considerations (Schweppe, et.al., page 253).

as the foundation for "market driven" merchant investment in transmission infrastructure.[64] An investment that increases network capacity would be rewarded with the associated incremental transmission rights. The value of these transmission rights, which are typically equated to the expected congestion charges either avoided (physical rights) or rebated by the system operator (financial rights) over the life of the transmission investment, then provides the financial incentive to invest in new transmission capacity. I must assume that this theoretical literature is the basis for the merchant transmission model that the NOPR proposes to rely upon.

Research on this model has focused almost entirely on simple cases where the transmission network and the effects of transmission investments on it satisfy a long list of assumptions. These include: The costs of transmission investments, operating and maintenance activities have no increasing returns to scale (or "lumpiness") or other non-convexities;[65] there are no firm or relationship-specific investments;[66] nodal energy prices fully reflect consumers' willingness to pay for energy and reliability at each location; all network externalities are internalised in nodal prices; fixed Mw capacity property rights that are "good all the time" are well defined and match simultaneously feasible transmission capacity "under normal operating conditions;" "normal operating conditions" can be mapped to boundaries of a feasible set of simultaneous physical schedules that have fixed Mw capacities; there is no market power; markets are always cleared by prices rather than involuntary administrative rationing; the boundaries of the

[64] Hogan, W. (1992) "Contract Networks for Electric Power Transmission," *Journal of Regulatory Economics*, 4: 211–242; Bushnell, J. and S. Stoft (1996) "Electric Grid Investment Under a Contract Network Regime," *Journal of Regulatory Economics*, 10: 61–79. Bushnell, J. and S. Stoft (1997) "Improving Private Incentives for Electric Grid Investment," *Resource and Energy Economics*, 19: 85–108.

[65] TO operating and maintenance decisions are ignored completely.

[66] As described in O. Williamson. *The Economic Institutions of Capitalism*, The Free Press, 1985.

feasible set used to define and allocate CRRs are constant under "normal operating conditions;" there is no discretionary behavior by TOs and SOs that may affect the effective transmission capacity and nodal prices over time; transmission investments that increase the boundaries of the feasible set of simultaneous physical point-to-point schedules do so under all "normal operating conditions;" and the allocation of CRRs created by these investments fall on the expanded feasible set and do not infringe on existing rights. Under these very restrictive assumptions it can be demonstrated that (a) efficient transmission investments that create transmission rights satisfying certain simultaneous feasibility constraints will be profitable and (b) that inefficient transmission investments will not be profitable. These two results are the primary economic foundation for relying on a merchant transmission model.

These are potentially powerful results that may appear to transform the transmission investment problem from one that appears to be almost intractable to one that requires only a simple implementation of a property-rights based market system. Accordingly, merchant transmission investment's appeal is that if the theory is a good representation of reality it allows unfettered competition to invest in new transmission capacity, placing the risks of investment inefficiencies and cost overruns on investors rather than consumers, and bypassing many planning[67] and regulatory issues associated with a structure that relies on regulated monopoly transmission companies. In addition, in theory, it allows investment in new generating capacity in the constrained area to

[67] Unlike most textbook market models, the merchant transmission model still requires some amount of regional planning activity. At the very least, the regional planning process would have to determine whether the investment expanded the boundaries of the feasible set, by how much it expanded these boundaries, that the selected allocation of new CRRs are on the frontier of the expanded feasible set and do not infringe on rights held by others (after any voluntary secondary market reallocations). As I will discuss presently, this is the case because defining and allocating appropriate property rights for transmission networks that do all of the good things that have been loaded on to them is a much more difficult (and complicated) challenge than is suggested in the NOPR.

"compete" with new transmission investment that reduces the import constraint. In this way, it is no surprise that market driven transmission investment is an economist's dream, avoiding dealing with issues associated with imperfect regulation of a "natural monopoly" transmission company and aligning competitive transmission investments with the newly developed competition in the generation segment.

Unfortunately, the large number of simplifying assumptions upon which the optimality of the market driven merchant transmission approach depends are unlikely to be consistent with the actual attributes of transmission investments and the operation of wholesale markets in practice. While there has been some recognition in the literature that relaxing its restrictive assumptions undermines key results regarding the optimality of merchant investment[68], little analysis of more realistic cases has been forthcoming.[69] The best one can say about the state of the relevant theory is that it is "immature" --- certainly too immature to serve as the foundation of the Commission's framework to govern transmission investment in light of the disappointing experience with the kind of merchant transmission investment envisioned in the NOPR. Let me outline how relaxing some of the simplifications that characterize the theories upon which the merchant investment model line seriously undermine its attractive properties.[70]

a. The Market and Institutional Imperfections that Motivate Resource Adequacy Obligations for Generating Capacity and Demand Response Also Undermine Incentives to Invest In Merchant Transmission

---

[68] James Bushnell and Steven Stoft, 1997 op. cit, pp. 102-105 and James Bushnell & Steven Stoft, *Grid Investment: Can a Market Do the Job?,* The Electricity Journal, January/February 1996, Volume 9, Number 1 at p. 77 recognize that economies of scale and free rider problems can lead to underinvestment. See also Oren, S., P. Spiller, P. Varaiya, and F. Wu (1995) "Nodal Prices and Transmission Rights: A Critical Appraisal," *The Electricity Journal,* 8(3): 24–35.

[69] An exception is, Perez-Arriaga, I. J. et. al (1995) "Marginal Pricing of Transmission Service: An Analysis of Cost Recovery," *IEEE Transactions on Power Systems*, 10(1): 546–553.

[70] This discussion draws heavily on my ongoing research with Jean Tirole. However, the presentation and conclusions presented here are my responsibility alone.

The NOPR contains a thoughtful discussion of market and institutional imperfections that lead spot energy and operating reserve markets to fail to provide adequate incentives for investment in generating capacity and demand response. Earlier in these Comments I have provided an expanded list of relevant market and regulatory imperfections that provide further support for the NOPR's conclusions. <u>However, the NOPR fails to recognize that exactly the same considerations that adversely affect incentives for generation and demand response capabilities also adversely affect incentives for merchant transmission investment.</u> The value of merchant transmission (and associated CRRs) reflects the expected difference in nodal spot prices for energy between affected injection and receipt points on the network over the life of the investment. However, the set of market and institutional imperfections that I identified above that affect the level of spot prices for energy and operating reserves <u>also</u> affect their geographic (locational) distribution. Accordingly, the market value of CRRs that provide the revenue source for merchant transmission must be distorted by these same market and institutional imperfections. Indeed, the problems are likely to be worse for merchant transmission than they are for generation investments. All of the economic value they can receive emerges only when there is congestion on the network. The NOPR's proposed local market power mitigation proposals and ITP discretion affecting nodal prices when there is significant congestion[71] are especially likely to distort the locational price signals and associated congestion rents that merchant investment will depend upon to finance transmission investments that will relieve congestion that is

[71] Glachant, J-M, and V. Pignon (2002) "Nordic Elecrtricity Congestion's Arrangement as a Model for Europe: Physical Constraints and Operators' Opportunism," mimeo.

creating load pockets where local market power is already a problem. The problems caused by imperfections in forward markets and the unsettled state of retail competition are exacerbated in the case of merchant transmission investments that will rely on congestion revenues to cover their costs. Robust liquid forward markets for these instruments do not exist, the application of feasibility tests to define and allocate additional CRRs are complex, uncertain and subject to enormous ITP discretion, and the legal status of CRRs as property rights is uncertain.

b. <u>Lumpiness and Economies of Scale Lead to Underinvestment With A Merchant Transmission Model</u>

It is widely recognized, but rarely emphasized, that if transmission investment is characterized by economies of scale ("lumpiness") and related "non-convexities" that there is likely to be underinvestment under a merchant transmission model. Basically, when there are economies of scale the investor in a merchant transmission project will recognize that the completion of a project of optimal size (and timing) will depress congestion prices and the project's profitability. As a result, the <u>private benefits</u> of the project will be less than the <u>social benefits</u> of the project. The effects can be no investment in a beneficial project, excessive delays in building projects of efficient size, the construction of projects that are too small, inefficient utilization of scarce transmission corridors and pre-emption of other more socially beneficial projects. The problems caused by lumpiness, economies of scale and other non-convexities associated with transmission investments have been dismissed by some with the argument that market participants who benefit from the investments will easily get together to internalise the gap between private benefits and social benefits. As I have already discussed, there are good reasons to believe that this kind of negotiated participant funding process will be costly and highly imperfect. One reason why merchant projects sponsored by one buyer (Cross Sound Cable) or one seller (Basslink) are moving forward is that the single buyer or seller is in a position to internalize the social benefits that a merchant investor would not otherwise be able to capture. These arrangements have also addressed financing problems associated with imperfections in forward contract markets by supporting the projects with long term contracts.

Network expansion investments, as defined earlier are most likely to exhibit significant scale economies ("lumpiness"). That is, the average cost of a new link declines as its capacity increases, other things equal.[72] Many *network deepening investments*, as defined earlier, may be less lumpy, but these investments have other attributes that are most conducive to efficient investment by the incumbent network owner rather than a third party. (I will discuss these issues further under "asset specificity" below.) Accordingly, the kinds of investment opportunities that are most conducive to merchant investment from a physical development and operating perspective are also those where inefficiencies resulting from economies of scale and the resulting gap between private and social benefits are likely to be most serious.[73]

Another source of lumpiness associated with *network expansion* investments arises because there may be a *scarcity of rights of way*, for example a unique corridor between a cheap generation and an expensive generation area whose scarcity value is not

[72] R. Baldick and E. Kahn (1992) "Transmission Planning in the Era of Integrated Resource Planning: A Survey of Recent Cases," Lawrence Berkeley Laboratory, LBL–32231. I.J. Perez-Arriaga et. al (1995) "Marginal Pricing of Transmission Service: An Analysis of Cost Recovery," *IEEE Transactions on Power Systems*, 10(1): 546–553.

[73] In addition, network expansion investments are the most likely to face state and local siting challenges. The incumbent TOs have experience working with state and local authorities to get transmission facilities approved. While this may be a difficult challenge for the incumbent TOs as well, they are more likely to be successful in obtaining siting permits than is a new merchant developer who does not have this experience or relationships or equivalent legal authorities under state law to get the project licensed.

reflected in the price of the associated land. The difficulties that new transmission corridors face in obtaining siting authority suggests that the available corridors for new lines through many areas will be limited in the sense that, for example, one additional corridor may be available through the mountains between Arizona and Southern California, and it may accommodate *one* new link that could be of any size between 100 MW and 1000 MW. Merchant investment is then likely to end up in a "preemption and monopoly" situation . A merchant investor will install a transmission line to occupy the scarce corridor and will later expand this capacity, underinvesting in the initial project and its expansion over time. For example, if the optimal investment is 1200 MW, a merchant developer may find it most profitable to invest in an 800 MW enhancement in the scarce corridor instead, pre-empting additional investment. Lumpiness also may make merchant investment occur too early when it takes place in order to pre-empt competition for the use of the scarce corridor.

These considerations lead to three conclusions. First, when there are economies of scale associated with transmission investment, underinvestment is likely to result from relying on a merchant model. Accordingly, a merchant model cannot be relied upon to produce the efficient level of transmission investment. Second, merchant transmission proposals should be scrutinized through a regional planning process to ensure that they are not pre-empting more efficient projects and are not misusing scarce transmission corridors whose social value is not fully priced in the market. Third, these results suggest that the NOPR's proposal that merchant projects be given a de facto first-in-line preference if they satisfy feasibility constraints needs further thought to respond to these potential investment distortions.

c. Assets Specificity Considerations Often Make the Incumbent TO the Most Efficient Owner and Operator of Enhanced Transmission Facilities

Network deepening investments, defined earlier can, as a practical matter, only be implemented efficiently by the owner and operator (maintenance and physical operation) of the existing network. Adding facilities owned and operated by third-parties that are fully integrated with and inseparable from the equipment that makes up the existing network from a physical and maintenance perspective, creates significant incentive problems with decentralized ownership and the associated potential for inefficiencies. The problems of defining a good set of rules for investing in and maintaining facilities of this type with decentralized ownership is further exacerbated by the heterogeneous nature of transmission facilities. While it is theoretically possible to devise contractual arrangements that will solve the incentive problems associated with decentralized ownership of physically inseparable assets, including opportunistic behavior of one or more parties, investments with these attributes are most likely to be governed efficiently through ownership by a single firm.[74] If the problems created by incomplete contracts cannot be resolved, the result will be inefficient investment. In addition, one would need to carefully allocate the new capacity of the line between the initial design and maintenance choice of the original owner and the actions of the renters who make deepening investments. This "moral hazard in teams" problem is a substantial obstacle to

[74] Jean Tirole, *The Theory of Industrial Organization*, MIT Press, 1988, pp. 21-29; B. Klein, R. Crawford and A. Alchian, “Vertical Integration, Appropriable Quasi Rents, and the Competitive Contracting Process,” *Journal of Law and Economics*, 21, 1978, pp. 297-326; O. Williamson. *The Economic Institutions of Capitalism*, The Free Press, 1985.

the design of an effective third party access policy for this type of transmission investment.[75]

The importance of *network deepening investments* raises the question of how incumbent transmission owners are to participate in a "market driven" transmission investment framework. On the one hand, precluding them from participating would mean that potentially low-cost network deepening investments will be lost. On the other hand, allowing them to make unregulated merchant investments for network deepening enhancements to which they have unique access would allow them to exercise market power, restrict supplies and capture rents that might otherwise go to consumers under a regulated investment regime.

d. Property Rights (CRR) Definition and Allocation Issues [76]

The efficiency of market driven investments in any sector of the economy depends upon the existence of a clear and well defined set of credible property rights that define ex ante rights and obligations that are consistent with the physical attributes of the assets created by the investments. The attributes of the CRRs contemplated in the NOPR turn out to be poorly adapted to the attributes of real transmission networks. Accordingly, they cannot provide perfect hedges against variations in congestion costs without being subsidized by consumers or TOs and cannot provide efficient incentives for merchant investment in transmission.

To understand why the proposed attributes of the transmission property rights contemplated in the NOPR are problematic it is useful to consider an example from another sector of the economy --- developing and renting space in buildings. I consider

---

[75] B. Holmström, "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 1982, 324–340.

[76] The discussion in this section is based on ongoing research with Jean Tirole.

two cases: A "normal case" and an "unusual case." The proposed CRRs fit the normal case for developing and renting buildings, but the attributes of real transmission networks are more like the unusual case I will develop below. As a result, we have a mismatch between the attributes of the property rights proposed in the NOPR and the attributes of the assets they are supposed to support. This mismatch undermines the NOPR's assumptions about full funding of CRR obligations and efficient investment incentives.

In a the typical or "normal" case, if a developer builds a building with 100,000 square feet of (net) rentable space, she knows that she will have the right to use or rent those 100,000 square feet of space, no more and no less, for whatever price prevails in the market.[77] The value of the space may vary widely with changing market conditions, but the rights to rent 100,000 square feet of physical space in a specific building are clear and cannot be expropriated. If the developer decides to add an extension to the building that creates another 25,000 square feet of rentable space, for example, by adding another floor, then she has another 25,000 square feet to rent in this particular building at the market price. The property rights are well defined and match the physical attributes of the building.

Now, consider an imaginary world where the physical space in this building varies randomly from month to month from a low 90,000 square feet to a high of 115,000 square feet based on factors <u>that are not under the direct control of the building's owner</u>.[78] Moreover, the amount of available physical space in this building can be affected by actions taken by the owners of other buildings in the city. This "unusual" building asset has very different attributes than the normal building asset we are accustomed to as

[77] Subject of course to meeting building code, safety, zoning and other regulatory requirements.
[78] That is, the variation in available space is not a result of maintenance activities undertaken by the building owner.

described in the previous paragraph. The physical capacity that the developer has invested in, and the associated property rights, are now contingent on exogenous random factors and third-party behaviour over which the owner has no control. The owner would have to take these contingencies into account when she voluntarily enters into rental agreements.

With this type of assets and the associated contingent property rights, one type of rental agreement would be to rent the minimum 90,000 square feet of “available all the time” space through a multi-year lease to one tenant and then rent the rest of the space to other tenants (maybe graduate students) on a month to month basis when it becomes (randomly) available. Another type of lease arrangement would be to rent, at a lower rent per square foot of space, the maximum 115,000 square feet under a long term lease agreement and leave it to the lessee to deal with the variations in the actual space available in any given month. Or, for a higher rent per square foot of space, the developer might offer a lease agreement for the maximum 115,000 square feet of space and herself take on the obligation to acquire additional space for the tenant to make up for the shortfalls when they occur. There are numerous other potential leasing arrangements that might be chosen, all of which would have to recognize the stochastic properties of the physical space available in the building.

Assume that this imaginary world has another unusual feature that affects investments in building expansion. When an investment is made that expands the building, the changes in rentable space vary depending on whether it is a “low-space month” or a “high space month.” Indeed, an additional floor increases the rentable space a lot during “low space” months, but actually reduces it during “high space” months. To make the decision regarding whether to invest in a building expansion or not, the developer must weigh the value of the increased rental space available under some contingencies against the reduced value of the rental space available under others.

The differences between the standard building described in the first paragraph of this section and the building in the unusual imaginary world described in the three paragraphs that follow it helps to illuminate an additional set of problems with the merchant investment model and with the nascent theoretical literature upon which it appears to be based. Specifically, the merchant model pretends that transmission assets and associated property rights have the attributes of the normal building described in the first paragraph of this section, while in reality their attributes are more like those of the unusual building described subsequently. This is just another dimension in which electric power networks are different from ordinary goods and services and any well-designed property rights based system for electricity needs to reflect these differences fully.

As proposed in the NOPR, a CRR gives the holders an entitlement (obligation) to receive (pay) the difference in nodal prices at the nodes covered by the rights times a fixed pre-specified Mw quantity of CRRs defined for each point-to-point pair.[79] The fixed Mw capacity of CRRs allocated for each point-to-point pair must be simultaneously feasible under "normal operating conditions." The proposed rights are *non-contingent* rights that do not vary with the actual physical capacity of the network under different

[79] I will focus here on point-to-point financial obligation rights since these are the rights that the NOPR proposes to introduce first. Let me note, however, that existing transmission agreements and contracts are more like point-to-point option rights. This will create complications for transforming current explicit and implicit physical rights into the financial rights proposed in a way that leaves existing rights holders "whole." It would be prudent to better understand the implications of implementing point-to-point obligation rights before implementing other types of rights that will further complicate what will already be a challenging implementation process.

"normal operating conditions." That is, they are rights that are "good all the time" regardless of exogenous changes in the physical feasibility of the network simultaneously to accommodate the associated bilateral schedules. When a transmission investment expands the feasible set, the investors in the transmission capacity are then entitled to an allocation of any additional *non-contingent* CRRs that fall within (or on the frontier of) the expanded feasible set created by the investment and do not infringe on existing rights.

The reliance on non-contingent rights as proposed in the NOPR (fixed MW entitlements to congestion revenues for each point-to-point pair) appears to be designed to achieve two goals simultaneously. One is to make available a set of financial instruments that provide a perfect hedge (or full insurance) against variations in congestion prices for each point-to-point pair. The second is to provide efficient incentives for investment in new transmission capacity. As I will discuss presently, however, except under a set of very restrictive assumptions, it is unlikely to be possible to achieve both of these goals simultaneously with the type of non-contingent CRRs proposed in the NOPR even in theory, let alone in practice. This is the case because the definition and allocation of CRRs proposed in the NOPR are based on a model of transmission network that is akin to the "normal building" described above while real transmission networks have attributes more like the "unusual" building described subsequently.

What makes it possible to define a set of CRRs that have fixed Mw capacities for each point to point pair, are "good all the time," are also always simultaneously feasible under "normal operating conditions," and where the payment obligations are no less than the congestion rents produced under "normal operating conditions? It depends on a long list of assumptions. Most importantly, the theoretical literature upon which the NOPR's discussion of feasibility and allocation of non-contingent CRRs appears to rely generally assumes (a) that the initial feasible set of simultaneous physical schedules is well defined in the sense that the boundaries of the feasible set do not vary with exogenous random variables that may be realized under "normal operating conditions" and (b) that the shifts in the frontier of the feasible set resulting from efficient investments in transmission capacity do not make any rights/capacity combinations that were previously in the feasible set infeasible, post investment. These assumptions match the attributes of the "normal building" discussed above.

However, this model and the associated assumptions abstract from some important issues that arise even on a transmission network without loop flow, but are especially problematic in more complex networks with loop flow. Consider first a simple 2-node network (no loop flow) with transmission capacity connecting a generation area in the North with a demand and generation area in the South. In practice, even in the two-node model, the actual capacity of the North/South link depends on exogenous environmental parameters. Furthermore, system operators have substantial discretion in defining and implementing security constraints, affecting the actual power flows, constraints, and nodal prices on the network in real time that need not match the boundaries of the feasible set determined ex ante. For example, the physical capability of transmission lines depends on temperature and other exogenous contingencies.[80]

[80] For example, the "rated" (read "feasible") capacity of Path 15, connecting Northern and Southern California falls by about 600 MW as the ambient temperature rises, other things equal. The "rated" (read "feasible") capacity of Path 15 varies by about 1300 MW depending on the availability of various remedial action schemes to respond to transmission and certain generation outages. California ISO, Operating Procedure T-122A, November 6, 2002. It is also important to recognize that in the U.S. there is not a single SO controlling the network, but multiple SOs controlling independent segments of the network. To

Network simulation exercises that incorporate contingency and stability constraints as well as thermal constraints and loop flow typically embody assumption about whether or not specific generating plants are operating or not, with the actual feasible set of transmission capacity combinations varying with the actual operation of these generating plants in real time.[81]

Consider a simple 2-node network with the following attributes: suppose that K is the capacity of the transmission line linking the North with the South, but this capacity is stochastic, varying with some exogenous variable (like temperature) θ: $K = K(\theta)$, $K'(\theta) > 0$ and θ is distributed between θ^- and θ^+. Let's say that the line is congested for all values of θ, but the value of the congestion η will vary with $K(\theta)$. For which value of θ should one compute the number of financial rights? All values of θ reflect "normal operating conditions" except to the extent that the ITP chooses one value arbitrarily to use for defining the feasible set and associated CRRs. The ITP could be conservative and set the number of financial rights equal to $K\left(\theta^-\right)$. The ITP would then issue $K\left(\theta^-\right)$ financial rights and owe the holders $\eta K\left(\theta^-\right)$ in congestion payments. When the realized θ is θ^-, the feasibility and revenue adequacy conditions are satisfied. But what happens when $\theta > \theta^-$? The congestion payments will exceed what is owed to the rights holders. What does one do with the excess and how does the distribution affect investment incentives? At the other extreme, one could set the number of financial rights to reflect

maintain reliability and avoid free riding, less flexible contingency criteria must be defined than might be the case if there were a single SO operating the network in real time.

[81] For example, the simultaneous import transmission capacity into Southern California varies by 700MW depending on the operating status of the three units of a nuclear generating plant in Arizona. California ISO, Operating Procedure T-103, November 6, 2002.

the maximum capacity $K\left(\theta^{+}\right)$. There would be revenue adequacy when $\theta=\theta^{+}$ but not when $\theta<\theta^{+}$, which would be most of the time since the system operator would owe $\eta K(\theta^{+})$regardless of the actual realization of θ. Where does the shortfall come from and how does this distribution it affect investment incentives? The answers to these questions necessarily affect the incentives merchant generators will have to make investments.

Moreover, transmission investments may <u>increase</u> capacity under some system conditions and <u>decrease</u> it under others. (Again like the "unusual" building above, rather than the "normal" building.) Indeed, this is a natural attribute of the standard 3-node model used to illustrate loop flow and a variety of congestion and pricing issues associated with it.[82] The primary rationale for investing in the transmission link connecting the two generation nodes, and then bearing the costs of congestion on this link resulting from loop flow, is because the third link has value when one of the radial links connecting generation directly with load experiences a full or partial outage.[83] Otherwise, it would typically be more efficient to eliminate the third link when it is congested because overall congestion and energy prices faced by consumers would then be reduced. If the potential outages of the radial links are ignored, the third link <u>reduces</u> the size of the feasible set when an optimal dispatch can be achieved without causing congestion on the radial links between the generation areas and the load area.[84] Thus, under "normal operating conditions" the third link would appear to be both infeasible and uneconomical (inefficient). But this analysis, and the "good all the time" CRR framework which is based on it, completely ignores the value of the link when there are

[82] Joskow and Tirole, 2000, op. cit., pp. 475-481; Bushnell and Stoft, op. cit., 1997, pp.94-99.
[83] Joskow and Tirole, 2000, op. cit., page 477.
[84] This should be clear from Figure 4 in Bushnell and Stoft, 1997, page 99.

outages on the other existing radial links. Accordingly, a transmission investment with potentially significant reliability benefits could be deemed to be infeasible and would receive no CRRs from the investment if this framework is adopted. And, of course, even a well-maintained system will have some random outages outside of the TOs control that cause the available capacity of the network to be reduced below what is simulated ex ante for the feasible set. Non-contingent CRRs fail to capture these attributes of real transmission networks, reflecting only one set of contingencies relevant to defining feasibility rather than the full distribution of contingencies that can occur under "normal operating conditions."

In short, transmission assets have attributes that are more like those of the second building described at the beginning of this section than the first. Pretending otherwise, and creating a property rights system that ignores these attributes, obscures the issues that must be addressed to evaluate the likely performance of a merchant investment model. None of these real world considerations appear to have been incorporated in the simple theoretical models upon which the NOPR's proposals regarding merchant transmission investment supported by CRRs are based. Indeed, the NOPR itself sweeps these issues under the rug be referring to feasibility "under normal operating conditions," a phrase that is not defined anywhere in the NOPR, as if there were a single well defined unambiguous set of "normal operating conditions" (e.g. ¶250). These considerations all raise a number of issues that must be confronted under more realistic conditions. Some ISOs have simply "punted" on these difficult issues, effectively allocating more CRRs than the capacity that exists under many "normal operating conditions" and collecting the CRR revenue shortfalls from TOs and their customers through an uplift charge. This is effectively a subsidy to cover the costs of CRRs that are perfect hedges. Such subsidies distort investment incentives as well as overcharge consumers for benefits received by others.

It is unlikely that the reliance on non-contingent CRRs that are "good all the time" can provide proper investment incentives on a real network with the kinds of attributes outlined above. Rather, in theory, *state-contingent CRRs* are likely to be more efficient from a risk hedging and an investment incentive perspective than are "good all the time" rights, as long as the contingencies can be readily described (temperature, output of specific generators that affect contingency limits, conditions in interconnected control areas, etc.) and contracted upon.[85] A state-contingent CRR would give different Mw values to the capacity (K) of CRRs defined for each point-to-point pair (properly) reflecting the varying capacity of the network when different "exogenous" conditions affecting network capacity are realized.[86] A consequence of using the appropriate contingent rights structure is that it necessarily requires an even larger number of contingent rights than the already large number of non-contingent point-to-point rights that go along with LMP, exacerbating a number of transactional problems associated with point-to-point CRRs: large transaction costs, thinness and market power in the secondary markets for these rights.

An alternative would be to define financial rights that entitle the CRR holder to a <u>proportional allocation</u> of congestion revenues associated with each point-to-point pair rather than create a fixed Mw obligation. Such rights provide better investment

---

[85]The desirability of contingent rights is hinted at in footnote 183 of the NOPR.

[86] Contingent rights are an old idea and are embedded in existing physical rights allocation mechanisms used in some parts of the country. They just have not yet made it into the theoretical models upon which the NOPR appears to rely.

incentives than an arbitrary fixed-MW allocation matched to a specific realization of uncertainty (θ) because they naturally take exogenous variations in feasible transmission capacity into account. This approach would not expand the number of rights to be traded or increase the transactional problems noted above. However, proportional congestion revenue allocation rights no longer provide full insurance against variations in congestion prices. Holders who valued full hedges and were willing to pay for the associated costs would then have to turn to private market financial intermediaries to supplement the insurance provided by the ITP-created CRRs and pay them for the hedges. The NOPR does not recognize the possibility that private market financial intermediaries will provide hedging and insurance instruments (at a price) similar to CRRs or discuss the reasons why these instruments will not be provided by voluntarily by the private sector. [87] Instead, the current proposal forces TOs to subsidize the full insurance that the NOPR requires these rights to provide.[88]

More generally, it may very well be <u>impossible</u> to define a set of CRRs that simultaneously completely insure holders from variations in congestion prices, provide efficient investment incentives, and leave TOs whole in the sense that they are not left to subsidize holders of CRRs. This issues, as well as several other potential problems associated with the kinds of CRRs proposed in the NOPR which I will not elaborate further here, will only be resolved when the research on these issues advances beyond its present immature state.

---

[87] e.g. if there is seller market power in constrained import areas these instruments may not be offered by financial intermediaries P. Joskow and J. Tirole "Transmission Rights and Market Power on Electric Power Networks," *Rand Journal of Economics*, 31(3): pp. 2000, page 460.

[88]It addition to leading to the problems already noted, subsidizing the provision of full insurance embodied in CRRs reduces opportunities for financial intermediaries to offer such insurance and eliminates opportunities for market participants to weigh the true costs and benefits of different levels of insurance.

The bottom line is that there are many issues associated with the appropriate definition of property rights for transmission capacity that still need to be resolved. The existing theory is immature, incomplete and ignores important attributes of real transmission systems. The limited available literature on the subject has established a good foundation for additional academic research. But this research is not yet at a state maturity, let along empirical verification, where it can be released for "prime time" application to form the basis for a new transmission investment framework of critical importance to the nation.

7. There is No Basis for Placing the Full Burden of CRR Revenue Shortfalls on TOs.

The discussion in the previous section also leads to the conclusions that it would be inappropriate for the Commission to adopt a rule that would simply require that TOs bear all of the responsibility for CRR revenue shortfalls. The way CRRs are defined as fixed Mw capacity non-contingent rights it is inevitable that there will be shortfalls and surpluses as "normal operating conditions" vary. Requiring TOs to bear the shortfall would be like forcing the owner of the second "unusual" building described above to rent 115,000 square feet of space in the building (defined by some building "ITP" as "normal operating conditions") at the prevailing market price for equivalent space without any financial provisions for contingencies beyond her control and then requiring her to pay the lessee when the available space falls below 115,000 square feet. This would be a confiscatory policy since, on average, the building owner must lose money. The non-contingent rights imposed on the building owner are also incompatible with the physical attributes of the asset. Obviously, if the developer had known that such a rule would be put in place she would not have invested in the building in the first place.

Returning to CRR revenue shortfalls, if CRR holders have K Mw of point to point CRRs for schedules between node 1 and node 2, they are entitled to $K(p_1 - p_2)$ congestion payments, where p_1 and p_2 are the prices at the 2 nodes and assuming that $p_1 > p_2$. If the actual capacity of the network turns out to be K_a rather than K, ($K_a < K$), say because of very high temperatures, then the system operator will have a congestion revenue deficit equal to $(K - K_a)(p_1 - p_2)$. The NOPR proposes that TOs bear the full burden of congestion revenue shortfalls, subject to a *force majeure* exception (¶250). Once we recognize the true stochastic attributes of transmission network capacity even under "normal operating conditions" and the difficulties of defining a set of "good all the time" CRRs, it is clear that revenue shortfalls will occur even when TOs have implemented optimal maintenance and investment programs because there is no single unambiguous value of θ that reflects the "normal operating conditions" assumed in the NOPR. Whether or not sufficient congestion revenues are produced to cover CRR obligations will depend on how contingencies are treated in the definition of the relevant feasible sets and associated CRR allocations, all decisions made by and subject to some discretion by the ITP.[89] Moreover, given the way that contingency constraints are applied in practice in network simulations supporting the determination of feasible simultaneous transmission capacity combinations (nomograms), actual transmission capacity may be reduced from the "feasible capacity" determined ex ante because of <u>generator</u> outages and behavior of TOs and ITPs in neighboring control areas.[90]

---

[89] Similarly, the "revenue sufficiency" theorems upon which important aspects of the foundations of the merchant transmission model rest ignore the stochastic attributes of transmission networks.

[90] For example, there are simultaneous import limitations into California that depend on the availability of links from the Southwest to Southern California, the Northwest to California, and the operating generating capacity inside California. These limits are presently managed administratively with "nomograms" that define the curtailments that are triggered when the constraints are binding.

Accordingly, there is no way that a TO can protect itself from CRR revenue deficiencies that arise from factors beyond its control. Giving TOs the "upside," a possibility raised in the NOPR (¶251) does not fix the problem since there is no particular reason to believe that the shortfalls and excesses are necessarily symmetrical, even assuming good maintenance practices by the TO.

I agree with the NOPR that it would be desirable to have a performance based regulatory mechanism to provide TOs with incentives to make economic maintenance and operating decisions that reflects the costs of congestion at different times so that transmission lines achieve optimal availability (¶ 251). However, the proposal simply to place all of the burdens of CRR revenue shortfalls on TOs is simply not an appropriate incentive mechanism because it violates at least two of the key criteria that such mechanisms must meet --- they should not be confiscatory[91] and should be targeted at performance variations that can be affected by TO operating and investment decisions not random performance effects or performance resulting from the actions of third parties over whom the TO has no control.

In summary, the theoretical underpinning of the merchant transmission model are immature and fail to take account of important attributes of transmission networks and their components. The limited international empirical evidence is not encouraging either. Creating a transmission investment framework based on the assumption that the bulk of the nation's transmission needs can be facilitated by merchant investors is more likely

[91] Expected revenues must be greater than or equal to expected costs if the firm's performance meets expectations.

than not to lead to serious underinvestment in transmission capacity. This will make the current transmission investment problems worse rather than better.

8. An Alternative Transmission Investment Framework

The sound framework for stimulating efficient transmission investment would recognize that the bulk of needed transmission investment will be made by incumbent regulated transmission owners in response to an open regional transmission plan process and subject to good incentive regulation mechanisms. CRRs would be defined, allocated and sold, but the precise details of how this is accomplished would not be the knife edge upon which investment in new transmission capacity depends. Merchant transmission developers would be permitted to propose and construct projects whose costs are not included in the ITP's tariff in return for CRRs. Such projects would be screened through the regional planning process for feasibility and non-infringing CRR definition and allocation purposes, as well as to guard against the distortions discussed above. Merchant projects could go forward only after the potential problems noted above are fully vetted through the planning process to ensure that more socially beneficial projects are not being preempted. The NOPR's proposal that they can proceed "first in line" merely if they satisfy feasibility and non-infringement constraints is not sufficient. The costs of regulated transmission projects would be apportioned based on a set of ex ante criteria for implementing a simple but fair beneficiaries pay principle as I discussed in my Comments above.

This approach can be adapted to current transmission planning and regulatory procedures in a straightforward way, can accommodate regional variations, and relies on an approach that has worked well in other countries. It is not dependant on speculation about the validity of immature theories that are poorly matched to the attributes of real transmission network. It permits merchant investment in transmission but does not count on it to provide the bulk of needed transmission capacity. It will help to end the regulatory barriers and gridlock surrounding transmission investment today.

a. Transmission and Generation Do Not "compete" in the same way as do typical goods and services that are close substitutes

Some will no doubt argue that this approach will not allow generation and transmission investments to "compete" with one another. I do not think that this is a real problem. Indeed, the emphasis that the Commission has placed on this kind of "competition" has created barriers to efficient investment in both generation and transmission, reducing real competition between generators. Transmission and generation do not "compete" in the same way as do, say, Dell, Gateway, HP and IBM in the manufacturing and retail distribution of personal computers and workstations. Transmission facilities do not produce any electricity.[92] Rather, they make it possible to move electricity from location to location on a network. When transmission investments expand capacity and eliminate congestion they expand the geographic expanse of competition between generators at different locations on the network. Thus, while a generator located in a constrained import area may not like to see transmission investments take place that increase import capacity, it is because the transmission investment increases competition from other generators, not because transmission investment is a direct competitor itself. The perspective of generators in import constrained areas is no different from that of early 19th century farmers in New England who feared investments in canals and railroads which would open up competition from

[92] Except perhaps in the limited sense that transmission investments may reduce losses.

farmers in the Midwest. Yet, we don't typically think of railroads from Boston to Chicago as competing with corn (once) grown in the Connecticut Valley.

The mantra of "let transmission and generation compete" has become an excuse for suppressing rather than for promoting real competition among generators. There has been a very substantial expansion in the new generating capacity that has been completed and placed in operation in the last three years. There is no evidence that these investments have contributed to a reduction in congestion or even that in areas (e.g. New York and PJM) that now have LMP, that a disproportionate amount of new generating capacity has even been attracted to areas with relatively high LMPs.[93] At the same time, congestion continues to increase around the country as new generating capacity is completed.[94] In short, there is nothing to suggest that generation investment is playing a significant role in helping to solve transmission congestion problems, even where LMP is in place. This is not very surprising. Generator location decisions depend on many variables including the availability and price of land, the availability of cooling water, the costs of transporting fuel, the costs of connecting to the network, and the costs of congestion on the network at different locations. Generators' decisions to continue operating once investments have been sunk, especially decisions of generators that are economical to operate for a relatively small number of hours each year are likely to be

[93] CERA, "Locational Marginal Pricing: Not A Transmission Panacea," December 2002, pp. 2-3. On the other hand the 2001 PJM State of the Markets Report argues that 60% of the generation in the various PJM queues is in the Eastern "region" of PJM where congestion, and local market power problems, are most severe and that these generation investments may help to reduce congestion and local market power problems in this area. 2001 Report at page 28. However, nearly 60% of the PJM load is also in the Eastern Region (2001 report at page 121), so that generation investment in the Eastern region is roughly proportional to the fraction of PJM load in that region, despite the high levels of congestion and the increasing need for local market power mitigation there.

[94] ftp://www.nerc.com/pub/sys/all_updl/oc/scs/logs/trends.htm .

more sensitive to locational prices for energy and operating reserves than are investment decisions by new generators.

Of course, transmission investment and generation investment can be alternative responses to congestion. If prices are high in a constrained import area generators may choose to locate there rather than somewhere else. If a transmission investment reduces the congestion, the generator may decide to invest at another location or competition from existing regional generating capacity that can now serve the area may make additional generation investment uneconomical. However, the nature of the impacts of the generation and transmission investments in this situation is different. The transmission investment permanently expands the geographic expanse of competition and, if it's a regulated investment subject to appropriate incentive regulation mechanisms that rewards reducing congestion and penalizes increasing it, there is no danger that once in place the transmission owners will have an incentive to withhold transmission facility (e.g. by declaring it partially unavailable) to drive up congestion prices. (The situation would be different if it were a merchant transmission project that could benefit from withholding capacity to increase congestion. In this regard merchant transmission and merchant generation are similar.) A generation investment inside the constrained area, in contrast, may not fully mitigate local market power problems (e.g. moving from a local monopoly to a local duopoly) and would always benefit from increased congestion.

Given the record to date, generation investments are simply not playing a significant role as alternatives to transmission investments that relieve congestion. To the extent that such generation projects are proposed they should be taken into account in the regional planning process. That is, transmission needs assessments should reflect reasonably anticipated and credible plans for investments in new generation and retirements of old generation. In this sense, generation would have a slight first-mover advantage in the regional transmission planning process.

b. CRRs and Market Power

The NOPR fails to recognize and deal with another important way in which the proposed transmission framework will affect competition. It is now well established that the distribution of CRRs to sellers or buyers of electricity with market power can significantly enhance their market power.[95] The NOPR is silent on this problem. I suggest that the Final Rule make it clear that market monitors must develop criteria for sale procedures, purchases and accumulations of CRRs that guard against using them to enhance market power.

[95] Paul L. Joskow and Jean Tirole, "Transmission Rights and Market Power," *Rand Journal of Economics*, Autumn 2000. See also Richard Gilbert, Karsten Neuhoff, and David Newbery, "Allocating Transmision to Mitigate Market Power in Electricity Networks," Cambridge University-MIT Project on Liberalized Energy Markets, October 2002.

IV. MANAGING CONGESTION VS. ALLOCATING AVAILABLE TRANSMISSION CAPACITY EFFICIENTLY

The NOPR reflects an excessively narrow definition of “congestion management” and improperly equates LMP and ITP “system operations” with a broader and more appropriate notion of congestion management and network operation and maintenance activities that involve actions by TOs as well. This narrow view of “congestion management” may be one reason that the NOPR fails to consider and propose new regulatory mechanisms that can increase the availability of transmission capacity at particular times when it is scarce. I support the basic components of the proposed SMD for day-ahead and real time markets, the associated security constrained bid-based dispatch, the computations of the resulting LMPs, and what the NOPR repeatedly refers to as “congestion management.” However, it would be much more appropriate to refer to this system as providing a mechanism to “allocate a given amount of available transmission capacity” efficiently rather than a comprehensive system of “congestion management” more broadly.

I think of “managing congestion” as encompassing all actions that can be taken by system operators and transmission owners that can affect congestion and associated congestion costs. Congestion management actions properly encompass maintenance decisions and expenditures, physical operating decisions that are still made (or should be made) by TOs, and investments, small and large, in the transmission network. An ITP running a security-constrained dispatch, resulting in a set of ex post LMPs is not in a position to undertake this broader set of congestion management actions. It does not have the people, the trucks, the materials, the money, or at the present time, the financial incentives to do it. Instead, what an ITP does on a day-to-day basis is to take into account the transmission capacity that is available and, using the bids made by generators and demand response, then calculates the most efficient way to allocate the transmission capacity that is available. Contrary to the frequent references made in the NOPR, LMPs themselves do not “manage congestion” in any meaningful way. The ITP’s security constrained dispatch allocates scarce transmission capacity based on the bids submitted by competing users of the network, supply and demand conditions, and available transmission capacity. The LMPs themselves merely provide ex post measures of congestion on the system given the transmission capacity available, the price and quantity bids made at different locations, and the security criteria specified by the ITP and included in its central economic dispatch program. Few, if any consumers actually see LMPs, so there is little if any response to them on the demand side either.

While an efficient security constrained dispatch that results in an efficient utilization of a given quantity of scarce transmission capacity is an important part of a comprehensive congestion management system, there are other important aspects of congestion management that are and are likely to continue to be in the hands of regulated TOs. The Commission should adopt a broader conceptualization of what “congestion management” actually means and reflect this broader conceptualization in its regulatory policies.

V. PERFORMANCE-BASED INCENTIVE REGULATION ISSUES SHOULD BE ADDRESSED IN THE FINAL RULE

These considerations also lead to the conclusion that the Commission should give a much higher priority to working with ITPs, TOs and state regulators to develop and apply good performance based regulation mechanisms that will stimulate a much broader range of beneficial congestion management efforts. Aside from the proposal to rely on a requirement that TOs be responsible for CRR revenue shortfalls, a proposal that I have already indicated has serious deficiencies, the NOPR largely ignores these regulatory issues. I strongly encourage the Commission to pay more attention to the development of performance-based regulatory mechanisms to be applied both to transmission owners and system operators. I submitted Comments to the Commission on performance based regulation, ownership and related issues over three years ago in response to the RTO NOPR.[96] Little progress has been made by the Commission on the PBR front since I submitted these Comments in 1999.[97]

We can see that these issues are important by examining the experience with transmission congestion, operating costs and transmission investment in PJM over the last several years. PJM appears to be the model for the SMD's proposed spot market design and associated available transmission capacity allocation system. And PJM appears to have done a good job operating a set of consistent spot energy markets, allocating available transmission capacity efficiently using market mechanisms, maintaining reliability given the available transmission capacity, and calculating the

---

[96] Comments of Professor Paul L. Joskow, Notice of Proposed Rulemaking, Regional Transmission Groups, Docket No. RM99-2-000, August 16, 1999.
[97] See Comments that I filed with the Commission, available at: http://web.mit.edu/ceepr/www/99010.pdf

associated spot LMPs. However, in other dimensions the performance record is much less impressive.

We can look first at PJM's record regarding congestion. Table 2 displays the congestion costs that PJM has experienced in the last few years. Congestion costs have grown by a factor of ten in four years.[98] Table 3 displays the number of transmission constraint hours experienced in total and at various voltage levels.[99] These too have increased by a similar order of magnitude.[100] Indeed, while PJM may have allocated the available transmission capacity more efficiently than in other areas of the country, the pattern of congestion over time in PJM is quite similar to the trends in TLR incidence reported by NERC.[101] Finally, Table 4 displays PJM's annual operating expenses. These too have grown by a factor of 10 in four years, heavily influenced by a large increase in the costs of interconnection studies and services performed by TOs and billed through PJM.[102] While perhaps not up to Boston "Big Dig" escalation levels, the rate of growth of PJM's expenses raises cost-control incentive issues that cannot be avoided.

These data should not be interpreted as implying that PJM has necessarily done a bad job. It appears to have done a very good job at running well functioning security

---

[98] I have been unable to obtain data for 2002 from PJM.

[99] The data for the high voltage lines are more comparable over time because PJM has gradually taken over more responsibility from TOs for dispatching the lower voltage facilities.

[100] I have been able to reproduce the congestion incident numbers reported in the annual PJM Interconnection *State of the Market Report: 2001* from other data on PJM's web site for all years except 2001. The numbers available on the PJM web site indicate over 9,600 transmission constraint incident hours in 2001, significantly higher than the 8,200 hours reported. I have not received a response to my inquiries to resolve the difference. During 2002 there were over 10,000 constraint incident hours reported for the first 11 months of the year (through November 25, 2002), though these data appear to include congestion on the expanded PJM West footprint which become operational in 2002, so that comparisons of with earlier data are not comparable without additional adjustments.

[101] ftp://www.nerc.com/pub/sys/all_updl/oc/scs/logs/trends.htm.

[102] If I exclude the costs of interconnection studies, PJM's operating expenses increased by just over 40% between 2000 and 2001, comparable to the rates of increase in operating expenses for the New York ISO (45%) and ISO-New England (38%) for this period.

constrained dispatch and companion LMP-based spot marketas in an internally consistent manner simultaneously with the efficient allocation of available transmission capacity. However, it has not been successful in managing congestion in the broader sense that I have described since congestion has increased significantly rather than decreased in PJM. Indeed, the pattern of increasing transmission congestion displayed in Table 2 and Table 3 is similar to what was experienced in England and Wales during the early 1990s prior to the introduction of a system operator performance-based incentive regulation mechanism. After a performance based incentive regulation mechanism was introduced in England and Wales in April 1994, congestion cost fell rapidly and were only about 10% of the 1993/94 level by 2000. Moreover, transmission investment has lagged in PJM as it has in most of the rest of the country, falling from an average of $166 million per year during the 1994-1997 period (with a peak of $240 million in 1994) to an average of $94 million per year during the period 1998-2001 (with a low of less than $50 million in 1999)[103] and average wholesale spot energy prices have increased since 1998 as well.[104]

It should be clear that the combination of LMP and ITPs is not in and of itself a magic elixir for the transmission investment and broader congestion management challenges and efforts to build on it to support a merchant transmission model are unlikely to be particularly successful.[105]

It is clear to me that the bulk of future transmission investment will be made by regulated TOs, rather than unregulated merchant developers. It is also likely that with a

[103] Based on data from FERC Form 1 filed by the TOs in PJM.
[104] PJM Interconnection, *State of the Market Report: 2001*, page 33, Table 3, page 36, Table 5, and page 37, Table 6.
[105] CERA, "Locational Marginal Pricing: Not A Transmission Panacea," December 2002.

supporting regulatory framework, divestiture of transmission assets by vertically integrated utilities will continue and that ITCs will become a growing presence in the U.S., as they are in most other countries with liberalized electricity markets.[106] Their geographic scope is likely to expand as well. Moreover, dividing some system operating functions between ITPs and TOs is likely to create coordination inefficiencies that will increase transmission costs. For-profit ITCs, unburdened by conflicts created by vertical integration, are likely to be in a good position to improve transmission network performance if they are permitted to take responsibility for additional system operating functions and operate under a well-designed PBR mechanism.

These observations lead to two conclusions. First, as I have already discussed, the Commission should devote much more attention to working with TOs and ITPs to develop state-of-the-art performance based incentive regulation mechanisms. Second, ITPs/ISO/RTOs should not be allowed to use this rulemaking to obtain a permanent monopoly position over a specific set of system operating functions, especially if they are organized in a way that makes it impossible to make them financially responsible for their performance. The worst kind of monopoly is one that has absolutely no financial responsibility for its actions. In the future, ITCs are likely to be in the position and should be permitted, indeed encouraged, to propose system operating functions that they can perform more efficiently than a not-for-profit ITP, supported by an associated performance-based regulatory program.

---

[106] Alex Henney, "What the U.S. Could Learn From Western Europe and Elswehere," *Electricity Journal*, December 2002, pp. 53-64.

TABLE 2

PJM TRANSMISSION CONGESTION CHARGES
$ Millions

	Per Audited Financial Statements	Per Annual Report on State Of the Markets
1998	$ 20.0	N/A
1999	$ 67.4	$ 53.0
2000	$108.1	$132.0
2001	$208.9	$271.0
% Change 1998-2001	+1044%	
1999-2001		+511%

Sources:
1. The numbers in the first column from the audited financial statements for PJM, LLC contained in PJM's Annual Reports. They are for the amounts billed for transmission congestion in each financial year.

2. The numbers in the second column come from the PJM Interconnection *State of the Market Report* for 2001, page 119.

TABLE 3

PJM TRANSMISSION CONSTRAINT EVENT HOURS
Hours

	Total	500Kv	345Kv	230Kv
1998	1,244	203	71	588
1999	2,134	189	148	818
2000	7,040	684	491	1,461
2001	8,227	1,326	725	2,317
% Change 1998-2001	+661%	+653%	+1021%	+394%

Source: PJM Interconnection *State of the Market Report* for 2001, page 124.

TABLE 4

PJM TOTAL OPERATING EXPENSES
$ Millions

1997	$ 27,297	
1998	$ 42,445	
1999	$ 57,532	
2000	$ 62,908	
2001	$123,141	
2002	$251,169	[$188,377 reported 9-month 2002 expenses annualized]

Source: Audited financial statements for PJM, LLC contained in PJM's Annual Reports, except for 2002 which relies on the 9-month (ending September 30, 2002) (unaudited) financial statements and annualizes the 9-month figure by multiplying it by 4/3. "Other expenses" and "interest expenses" reported on the financial statements have not been included in these figures. These additional expenses were $15.3 million in 2001, $6.5 million in 2000, and $4.0 million in 1999 and $12.2 million for the first 9 months of 2002.