Cambridge Working Papers in Economics CWPE 0328

UNIVERSITY OF
CAMBRIDGE
**Department of
Applied Economics**

The Difficult Transition to Competitive Electricity Markets in the U.S.

Paul L. Joskow

The
Cambridge-MIT
Institute

*Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research*

CMI Working Paper 28

Cambridge Working Papers in Economics

UNIVERSITY OF
CAMBRIDGE
Department of
Applied Economics

Not to be quoted without permission

The
Cambridge-MIT
Institute

Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research

CMI Working Paper

THE DIFFICULT TRANSITION TO COMPETITIVE ELECTRICITY MARKETS IN THE U.S.[1]

Paul L. Joskow
MIT
May 2003

ABSTRACT

This paper provides a comprehensive discussion of the causes and consequences of state and federal initiatives to introduce wholesale and retail competition into the U.S. electricity sector between 1995 and the present. Information about the development of wholesale market institutions, the expansion of wholesale power trade, the performance of wholesale market institutions, the entry of merchant generating capacity, and the financial collapse of the trading and merchant generating sector is presented and discussed. Issues regarding the ability of evolving spot wholesale energy market institutions and market power mitigation mechanisms to provide adequate incentives for investment in new generating capacity in the absence of some form of peak capacity obligation are discussed theoretically and evaluated empirically. The diffusion of retail competition and the performance of retail competition programs in eight states is examined empirically. Imperfections in transmission governance arrangements and barriers to efficient expansion of the transmission network are identified. The analysis leads to the overall conclusion that the development of efficient competitive wholesale and retail electricity markets continues to be a work in progress and faces many technical, institutional and political challenges in the U.S. Suggestions for successfully confronting, or at least better understanding, these challenges are presented.

[1] Prepared for the conference "Electricity Deregulation: Where From Here?" at the Bush Presidential Conference Center. Texas A&M University, April 4, 2003. This draft reflects helpful comments that I received from the conference organizers and participants on an earlier version. I am grateful for financial support from the MIT Center for Energy and Environmental Policy Research (CEEPR) and the Cambridge-MIT Institute (CMI).

THE DIFFICULT TRANSITION TO COMPETITIVE ELECTRICITY MARKETS IN THE U.S.[1]

Paul L. Joskow
MIT
May 2003

If deregulation is to play a role in helping to improve the efficiency with which electricity is produced and used, it must be introduced as part of a long-term process that also encompasses regulatory and structural reform (Joskow and Schmalensee, 1983, p. 221)

[A]ny deregulation scheme must be carefully structured to conform to the basic technological and economic conditions that characterize the supply and demand for electricity (Joskow and Schmalensee, 1983, p. 212)

[T]here is significant uncertainty as to the likelihood of effective competition … in many areas of the country. The key unknown is the impact of transmission capacity constraints on the effective extent of geographic markets (Joskow and Schmalensee, 1983, p. 190)

Electricity restructuring … is likely to involve both costs and benefits. If the restructuring is done right…the benefits … can significantly outweigh the costs. But the jury is still out on whether policymakers have the will to implement the necessary reforms effectively (Joskow, 1997, p. 136)

INTRODUCTION

Academic discussions about the opportunities and challenges associated with introducing wholesale and retail competition into the electric power sector have gone on for decades. However, serious considerations of comprehensive electricity sector restructuring and deregulation initiatives in the U.S. only began in the mid-1990s, following the first comprehensive privatization, restructuring, wholesale and retail competition program undertaken in England and Wales (E&W) in 1990. The first

---

[1] Prepared for the conference "Electricity Deregulation: Where From Here?" at the Bush Presidential Conference Center. Texas A&M University, April 4, 2003. This draft reflects helpful comments that I received from the conference organizers and participants on an earlier version. I am grateful for financial support from the MIT Center for Energy and Environmental Policy Research (CEEPR) and the Cambridge-MIT Institute (CMI).

comprehensive U.S. programs did not go into operation until early 1998. Of course, wholesale power markets in which proximate vertically integrated utilities traded power on a daily and hourly basis subject to very limited regulation, have existed in the U.S. for many years. In addition, during the 1980s the Public Utility Regulatory Policy Act (PURPA) of 1978 stimulated the development of a non-utility power sector selling electricity produced primarily from cogeneration facilities and renewable energy facilities to local utilities under long-term contracts (Joskow, 1989). The Energy Policy Act of 1992 also removed important barriers to the broader development of unregulated non-utility generating facilities and expanded the Federal Energy Regulatory Commission's (FERC) authority to order utilities to provide transmission service to support wholesale power transactions. However, these developments largely reflected modest expansions of competition at the wholesale level built upon a basic model of regulated vertically integrated franchised monopolies.

The primary impetus for more fundamental restructuring and competition initiatives can be traced to electricity policy debates that began in California and a few states in the Northeast (Massachusetts, Rhode Island, New York, Pennsylvania, Maine, and New Jersey) in the mid-1990s, combined with supporting transmission and wholesale market rules and regulations issued by FERC (e.g. Orders 888 and 889) at about the same time (Joskow, 2000). These debates eventually led to regulatory decision and state legislation in a number of states to embraced competitive electricity market models. The first retail competition programs began operating in Massachusetts, Rhode Island and California in early 1998 and spread to about a dozen states by the end of 2000. By that time about a dozen additional states had announced plans to introduce similar programs in the near future.

The early state restructuring and competition programs included the unbundling of the retail supply of generation services from the supply of distribution and transmission service and giving retail customers with the opportunity to choose their power supplier from among competing retail suppliers. These programs also included various utility restructuring requirements designed to separate (functionally or structurally) competitive services (e.g. generation, retail supply) from monopoly services (e.g. distribution and transmission) that would continue to be subject to (better) regulation, as well as various transition arrangements involving stranded cost recovery, generation assets sales, and regulated retail supply services. These transition arrangements typically involved a mandatory reduction of regulated retail prices charged to all consumers (or at least residential consumers) and some type of "default" service arrangement to supply retail customers with a regulated backstop retail power supply option until they migrated to competitive retail suppliers during what was expected to be a short transition period. As the year 2000 began it appeared that these types of competitive electricity sector reforms would sweep the country within a few years as a growing number of states jumped on the bandwagon. There was also a serious prospect that supporting federal legislation would be enacted by Congress to remove remaining legal and policy barriers to effective wholesale and retail competition and to harmonize diverse state policies.

Since the year 2000, however, no additional states have announced plans to introduce competitive reforms of this type and about nine states that had planned to implement reforms have delayed, cancelled or significantly scaled back their electricity competition programs. Federal pro-competition electricity legislation has also been stalled. The California electricity crisis of 2000-2001 (Joskow 2001), Enron's bankruptcy, the financial collapse of many merchant generating and trading companies, volatile wholesale market prices, rising real retail prices in some states, phantom trading and fraudulent price reporting revelations, accounting abuses, a declining number of competitive retail supply options for residential and small commercial customers in many states, and continuing allegations of market power and market abuses in wholesale markets have all helped to take the glow off of electricity "deregulation" in many parts of the U.S. The average real retail price of electricity in the U.S. increased for the first time in 15 years in 2000 for industrial customers and in 2001 for residential customers (Figure 1), though preliminary data indicate that real prices fell in 2002. FERC has found itself at war with many states in the Southeast and the West as they resist its efforts to expand institutions it believes are necessary to support efficient competitive wholesale markets in all regions of the country. In response to the resulting political pressure, in a White Paper issued on April 28, 2003, FERC indicated that it would provide states and regions with more time and flexibility to implement the wholesale market reforms --- the Standard Market Design (SMD) --- that it proposed in a Notice of Proposed Rulemaking (NOPR) issued in August 2002.

At the very least, the pace of wholesale and retail competition and the supporting restructuring and regulatory reforms has slowed considerably since 2000. Many states have concluded that these types of electricity sector reforms are not in the interest of consumers in their states, or that it is prudent to wait to see if policymakers can figure out how to make competition work well and can demonstrate that these reforms will bring long term benefits for consumers. At the same time, most of the states in the Northeast, a few in the Midwest, Texas, and FERC are committed to moving forward with the development of competitive wholesale and retail markets and to making them work well.

In this paper, I will discuss what has gone right, what has gone wrong and the lessons learned from the experience with wholesale and retail competition in the U.S. during the last five years. My focus is selective rather than comprehensive, reflecting my own interests and assessment of what aspects of the reform program are most important to understand and where improvements are needed. Since California has attracted so much attention by industry analysts, the media, and policymakers, this paper will say little about California and focus on developments in the rest of the country.[2]

WHY WHOLESALE AND RETAIL COMPETITION?

Background

Electricity sectors almost everywhere on earth evolved with (primarily) vertically integrated geographic monopolies that were either publicly owned or subject to public regulation of prices, service obligations, major investments, financing, and expansion into unregulated lines of business. That is, the primary components of electricity supply --- generation, transmission, distribution, and retail supply --- were integrated within individual electric utilities. These firms in turn had de facto exclusive franchises to supply residential, commercial and industrial consumers within a defined geographic area. The performance of these regulated monopolies varied widely across countries and between utilities in the U.S. While there is much to criticize about the institution of

[2] My views on the causes of the California electricity crisis can be found in Joskow (2001) and my analysis of price formation and market power during the summer of 2000 in Joskow and Kahn (2002).

regulated monopoly, it is a mistake to assume that the performance of the U.S. electric power industry during the 20th century was extremely poor. The electric power sector in the U.S. had a very high rate of productivity growth from 1900 to 1970 and it's performance compared very favorably to international norms regarding labor productivity, production costs, penetration rates, reliability and prices. Serious problems began to emerge during the 1970s and 1980s as fossil fuel prices rose, inflation and interest rates rose, nuclear power plant costs exploded and some states embraced PURPA with excessive and costly enthusiasm (Joskow 1974, 1989). Real retail electricity prices rose significantly in many states during the 1970s and early 1980s for the first time in the history of commercial electric power. Moreover, it became clear that there were significant variations in performance across utilities, but an industry structure which provided limited opportunities for more efficient suppliers to expand and to place pressure on less efficient suppliers to improve or contract.

On a national (average) level, real retail electricity prices began to fall once again after the mid-1980s, and continued to fall until 2000-2001, as fossil fuel prices and interest rates declined and inflation moderated significantly. Nevertheless, the legacy of costly nuclear investment and long-term contracting decisions made during the 1970s and 1980s continued to be reflected in regulated retail prices during the 1990s in those states that had made major commitments to these resources (e.g. California, New York, New England). While on average real retail electricity prices fell after 1985, in the Northeast, California and a few other states real retail prices continued to rise into the late 1980s and early 1990s as legacy costs of nuclear plants and QF contracts, combined with excess generating capacity, continued to be reflected in regulated retail prices. Moreover, large disparities emerged between regulated retail electricity prices in different states and regions <u>and</u> between the embedded cost of electric generation services used to set regulated retail prices and the apparent <u>market value</u> of electric generation services in the regional wholesale power markets. In the Northeast, California, Illinois and a few other states, a large gap appeared to exist between the regulated price of generation service and the wholesale market value of these services. Industrial customers began to view regulated utilities as representing a barrier to obtaining lower-priced power available in wholesale markets. From the regulated utilities' perspective, the "gap" between regulated generation prices and competitive generation prices indicated the potentially "stranded costs" that utility investors would incur if generation service prices dropped from their regulated levels, reflecting historical capital costs of generating facilities and long-term contractual commitments to buy power from PURPA Qualifying Facilities, to reflect their current wholesale market values.

Three things are worth noting about this gap between the generation component of regulated retail prices and wholesale market values. First, it largely (though not entirely) reflected sunk costs associated with historical investment and long-term contracting decisions. Second, the size of the gap varied widely from region to region and utility to utility. In several areas of the country the gap was negative --- i.e. regulated generation prices were <u>below</u> their wholesale market value. Third, wholesale market prices observed in the early 1990s were a misleading indicator of what wholesale prices would be in the long run in an industry where generation services were sold largely through market mechanisms rather than through vertical integration. The wholesale markets in the early 1990s were largely short- term "excess energy" markets in which utilities with capacity in excess of their retail customers' needs could sell it to neighboring utilities that had generating capacity with higher short-run marginal cost which could be displaced by lower cost energy purchased in the wholesale market. There was excess generating capacity overall in most areas of the country and the operating costs of a vertically integrated utility's own generation placed a natural cap on wholesale market prices.

Initial interest in electricity sector reform started in the states with the highest retail electricity prices and where the apparent gaps between wholesale and retail prices were the largest. They included California, Massachusetts, Rhode Island, New York, New Jersey, Maine and Pennsylvania. As discussed elsewhere (White, Joskow (2000)), the political pressures for reforms in these states, and in particular for retail competition, came from lobbying activities by industrial customers, independent power producers, and would-be electricity marketers with experience in the natural gas industry. Enron played a major role in stimulating interest in restructuring and competition in almost every one of these "pioneer" states. The primary selling point to state regulators and legislators was that by introducing competition, retail prices would fall significantly to reflect the lower priced power available in the wholesale market. Incumbent utilities in these states initially opposed these retail competition proposals due to the potential for stranding sunk costs, but ultimately negotiated settlements that provided for recovery of a significant fraction of these sunk costs. How retail prices could both fall dramatically to reflect lower wholesale prices <u>and</u> utilities could recover their stranded costs (roughly the difference between regulated generation prices and the expected wholesale price of electricity) was a bit of questionable arithmetic that was largely glossed over.

Public Interest Goals

While, the political debates at the state level focused on retail price reductions, stranded cost recovery, and the creation of opportunities for incumbents and hungry new entrants, the intellectual debates focused on a broader set of public interest goals and implementation strategies. I think that there is fairly wide agreement about the goals that electricity sector reforms should achieve and even on the basic architecture of a model for creating competitive wholesale and retail markets to achieve these goals. It is less clear that there was broad understanding of what would have to be done to achieve these goals and how long it would take to achieve them.

The overriding reform goal is to create new governance arrangements for the electricity sector that will provide long-term benefits to consumers. Theses benefits will accrue by relying on competitive wholesale markets for power to provide better incentives for controlling capital and operating costs of new and existing generating capacity, to encourage innovation in power supply technologies, and to shift the risks of technology choice, construction cost and operating "mistakes" to suppliers and away from consumers. Retail competition, or "customer choice" would allow consumers to choose the supplier offering the price/service quality combination that best met their needs, and competing retail suppliers would provide an enhanced array of retail service products, risk management, demand management, and new opportunities for service quality differentiation based on individual consumer preferences.

It was also widely recognized that significant portions of the total costs of electricity supply --- distribution and transmission --- would continue to be regulated. Accordingly, reform to traditional regulatory arrangements governing the distribution and transmission networks have generally been viewed as an important complement to the introduction of wholesale and retail competition to supply consumer energy needs. This is the case for at least two reasons. First, regulatory mechanisms with good incentive properties would lead to lower distribution and transmission costs and this is turn would help to reduce retail electricity prices. During the first decade of the electricity restructuring and competition program in England and Wales, as much as 35% of the reduction in real electricity prices was associated with cost reductions in distribution and transmission. Second, the efficiency of wholesale markets in particular depends on a well functioning supporting transmission network and its efficient operation by a system operator. Good operating and investments incentives are important for providing an efficient network platform upon which wholesale and retail competition would proceed.

In the long run, the promise was that these reforms would lead to lower costs and lower average retail price levels reflecting these cost savings <u>compared to regulated monopoly alternative,</u> while maintaining or enhancing system reliability and achieving environmental improvement goals. Anyway, this was the dream.

<u>The Basic Model for Competitive Wholesale and Retail Markets</u>

Whatever the political motivations, the basic architecture for transitioning to competitive electricity markets had already been developed in theory and applied in practice in other countries (e.g. England and Wales, Norway, Argentina). It involves several key components that are depicted in Figure 2:

a. Vertical separation of competitive segments (e.g. generation, marketing and retail supply) from regulated segments (distribution, transmission, system operations) either structurally (through divestiture) or functionally (with internal "Chinese" walls separating affiliates within the same corporation).

b. Horizontal integration of transmission and network operations to encompass the geographic expanse of "natural" wholesale markets and the designation of a single independent system operator to manage the operation the network, to schedule generation to meet demand and to maintain the physical parameters of the network (frequency, voltage, stability) so that the lights would stay on except under extremely rare conditions.

c. The creation of wholesale spot energy and operating reserve market institutions to support requirements for real time balancing, to respond quickly6 and effectively to unplanned outages of transmission or generating facilities consistent with the need to maintain network voltage, frequency and stability parameters within narrow limits, and to facilitate economical trading opportunities among suppliers and between buyers and sellers.

d. Creation of institutions to facilitate access to the transmission network by buyers and sellers to facilitate economical production and exchange, including mechanisms efficiently to allocate scarce transmission capacity.

e. Horizontal restructuring, forward supply commitments and/or behavioral rules to mitigate regional and localized market power in wholesale markets.

f. Unbundling retail tariffs to separate retail power supplies and associated support services to be supplied competitively from distribution and transmission services that would continue to be provided by regulated monopolies.

g. Requiring retail consumers to purchase their power supplies from competing retail suppliers which in turn buy their power in wholesale markets, or own generating facilities to support their retail supply commitments.

This is the basic architecture, but if we have learned anything in the last several years of U.S. and international experience, the devil is in the details of actual implementation in practice.

UNDERESTIMATING THE RESTRUCTURING CHALLENGE

In the U.S., electricity sector restructuring and competition initiatives got off on the wrong foot in many parts of the country at least in part because policymakers and many of their advisers underestimated the nature and magnitude of the technical and institutional challenges associated with successfully introducing competitive wholesale and retail markets and the uncertainties associated with how best to respond to these challenges. To some extent the underestimation of the magnitude and extent of the challenge was strategic, reflecting efforts by some participants in the process to feather their own nests. However, it also reflected a combination of ignorance, political barriers, and true uncertainty about how best to restructure to support competition and how to design effective wholesale and retail market, transmission, and system operations institutions. The experts did not, and in many areas still do not, agree on exactly how best to proceed with these structural and institutional reforms. This situation reinforced the natural inclination of policymakers to treat the details of the restructuring program as a political rather than a technical problem. This in turn resulted in numerous political compromises over restructuring and market design issues and the mixing and matching of pieces of alternative restructuring models that did not fit very well together.

Why is the transformation of a regulated monopoly electric power industry into one that relies on competition to supply power at the wholesale and retail levels so challenging? There are several sets of reasons. First, electricity has an unusual set of physical and economic attributes that significantly complicate the task of replacing hierarchies (vertical integration and multilateral agreements) with decentralized market mechanisms. These attributes include:

a. Electricity cannot be stored economically and demand must be cleared with "just-in-time" production from generating capacity available to the network at (almost) exactly the same time that the electricity is consumed.

b. Physical laws governing electricity network operations in real time to maintain frequency, voltage and stability of the network, along with network congestion, interact with non-storability to require that supply and demand be cleared continuously at every location on the network. Creating a set of complete markets that operate this quickly, at so many locations, and without creating market power problems is a significant challenge.

c. The short-run demand elasticity for electricity is very low and supply gets very inelastic at high demand levels as capacity constraints are approached. As a result, spot electricity prices are inherently very volatile and unusually susceptible to the creation of opportunities for suppliers to exercise market power unilaterally.

d. Network congestion, combined with non-storability, may limit significantly the geographic expanse of competition by constraining the ability of remote suppliers to compete, further enhancing market power problems.

e. Loop flow, resulting from the physics of power flows on AC networks, introduces additional complex interactions between generators at different points on the network, creating unusual opportunities for suppliers to take actions unilaterally to affect market prices, complicating the definition of property rights, and creating coordination and free riding problems when, as in the Eastern and Western networks in the U.S., there are multiple system operators responsible for interconnected portions of a single synchronized AC network.

f. Electricity demand varies widely from season to season, between day and night, with extreme temperatures, and between weekdays and weekends (and holidays). The difference between the peak demand and the lowest demand over the course of a year is a factor of about three. Because electricity cannot be stored and varies widely over the year, a significant amount of the generating capacity connected to the system operates for a relatively small number of hours during the year to meet peak demands. Historically, there has also been little reliance on real time prices to ration peak demands. This means that the ability of generators that provide services for a small fraction of the year to recover their investment and fixed operating and maintenance costs is heavily dependant on the price formation process during periods when demand (and prices) are at their highest levels.

g. The combination of non-storability, real time variations in demand, low demand elasticity, random real time failures of generation and transmission equipment, the need to continuously clear supply and demand at every point on the network to meet the physical constraints on reliable network operations, means that some source of real time "inventory" is required to keep the system in balance. This "inventory" is generally provided by "standby" generators that can respond very quickly to changing supply and demand conditions, though demand side responses can also theoretically provide equivalent services as well. Compatible market mechanisms for procuring and effectively operating these "ancillary services" are therefore necessary. Designing well functioning integrated markets for energy to meet demand and the need for multiple ancillary services to maintain network reliability consistent with all of the other constraints and attributes enumerated above is very challenging.

h. The performance of competitive markets for electricity depend critically on the way the regulated transmission network is operated, access to it priced, and scarce transmission capacity is allocated. There are important complementarities between energy markets and transmission operations, especially congestion management and responses to emergencies. Integrating spot energy and ancillary services markets with the allocation of scarce transmission capacity is necessary to wholesale power markets to operate efficiently.

While there are many competitive industries that have one or perhaps two of these attributes, it is hard to think of any commodity market that has all of them.[3] Moreover, it is precisely these attributes of electricity that led to vertical integration between generation and transmission and either to extensive horizontal integration or to multilateral cooperative agreements between neighboring vertically integrated system operators. Ignoring these unusual attributes of electricity, and ignoring how and why historical governance arrangements evolved for dealing with them (Joskow, 1997, 2002), is a very bad mistake. Replacing these hierarchical governance arrangements with well functioning decentralized market mechanisms is a very significant technical challenge, about which even the best experts have disagreements. Accordingly, it should not be surprising that electricity restructuring and competition programs have inevitably been a

[3] So, for example, hotel rooms are non-storable. An empty room cannot be stored for another day. However, the demand for hotel rooms is likely to be quite elastic, "stockouts" frequently occur when demand is high, average utilization factors are 80-90%, one hotel cannot affect the supply of rooms at another hotel in the city by closing down several rooms, and we do not expect hotels to rent all of their rooms at a uniform price equal to the hotel's short run marginal cost.

process that involves a lot of learning by doing and ongoing changes to market rules, regulatory arrangements, and governance institutions.

These technical challenges have been further complicated in the U.S. by a number of institutional, legacy investment and political factors that many other countries have been able to avoid. First, the U.S. industry has been characterized by an unusually large number of private vertically integrated utilities of widely varying sizes that own and control generation, transmission, and distribution facilities in or near their distribution franchise areas. Many of these vertically integrated utilities are control area operators (about 140 in 1995) that were, and in many cases still are, responsible for operating portions of one of the three synchronized AC networks in the U.S., subject to rules established by the regional reliability councils and a variety of bilateral and multilateral operating agreements. Only in the Northeast did multi-utility power pools emerge to centrally dispatch generation resources and manage the operation of transmission networks with different owners of individual pieces.

This legacy industry structure is not conducive to creating well functioning competitive wholesale and retail markets (Joskow and Schmalensee, 1983). Ideally, a restructuring program would have separated competitive generation and marketing functions from regulated transmission and distribution activities. Generation ownership would have been further decentralized if ownership concentration created significant additional market power problems. Horizontal integration of transmission assets would have taken place to create regional transmission companies to own and operate transmission networks spanning large geographic areas. This was the basic approach of the restructuring programs in England and Wales, Spain, Norway, Argentina, Australia, and Alberta.

Third, the electric power industry in the U.S. has historically been regulated primarily by the states. The states have divergent views about the desirability of transitioning to competitive wholesale and retail electricity markets and restructuring the utilities in their states to do so effectively. Unlike most other countries that have gone down this path, the U.S. has no clear and coherent national laws that adopt a competitive wholesale and retail market model as national policy and that give federal authorities the tools to do the necessary restructuring and market design work required to make it work. Congress has passed no legislation mandating the implementation of a comprehensive wholesale and retail competition model for the electricity sector. Instead the U.S. has relied heavily on individual state initiatives and efforts by FERC to use its existing but limited Federal Power Act authority to cajole and encourage the states and their utilities to create competitive wholesale markets and supporting transmission institutions. It is hard to force states to adopt policies they don't like, especially when the regulated utilities in these states don't like them either. It is also difficult to force private firms to divest assets and restructure vertically and horizontally without providing them with financial incentives to do so. In most other countries, the restructuring program was implemented in conjunction with the privatization of state-owned assets so that they did not have to confront issues associated with government takings of private property. As a result, to make progress, FERC has had to rely on a variety of alternative regulatory and institutional arrangements, and various regulatory carrots and sticks to provide incentives for cooperation, in order to compensate for its inability to require the kind of restructuring program that can most effectively support wholesale and retail competition.

Fourth, the combination of many relatively small vertically integrated utilities, many operating small control areas, combined with state regulation, has had the effect historically of limiting investments in transmission capacity that provides strong linkages between generating facilities over large geographic areas. So, for example, New England has only 1500 Mw of transmission capacity connecting this six-state region with the rest of the United States. The Pennsylvania, New Jersey, Maryland, Delaware and Washington, D.C. area, where the major utilities have participated in a power pool (PJM) since the 1920s, has a strong internal transmission network, but only about 3500 Mw of firm simultaneous transmission capacity with neighboring states. Moreover, the configuration of the control areas' internal networks typically reflected a century of evolution of the utilities that began supplying electricity early in the 20th century, with generating plants first located in or near load centers and then gradually expanding as more remote generating sites became necessary to accommodate larger generating stations. Interconnections with neighboring utilities were built primarily for reliability purposes rather to gain access to lower cost power supplies located remotely from the local utility's franchise area. In addition, there was no need for these transmission investments to take account of the potential market power problems caused by transmission constraints. The legacy transmission networks therefore represent a potentially serious constraint on effective competition when wholesale power markets are deregulated due to the resulting limitations on the geographic expanse of wholesale power markets (Joskow and Schmalensee, 2003).

These institutional, legacy investment, and political realities have significantly complicated the kind of industry restructuring that is necessary for effective implementation of what is already the very significant technical challenge of creating well functioning competitive wholesale and retail markets for electricity. They continue to be a barrier to effective national reforms today.

GENERATION RESTRUCTURING AND MERCHANT GENERATION INVESTMENT

The approaches to restructuring to support wholesale and retail competition have varied widely across the states. States that have implemented retail competition programs have also typically strongly encouraged or required the affected utilities to separate their regulated T&D businesses from their wholesale generation and marketing activities. The first few states to implement retail competition programs also (effectively)[4] required their utilities to divest substantially all of their generating capacity through an auction process (e.g. California, Massachusetts, New York, Maine, Rhode Island). Other states that have implemented retail competition programs permitted the utilities under their jurisdiction to retain the bulk of their generating assets and to move them into separate unregulated wholesale power affiliates within a holding company structure (e.g. Pennsylvania, Illinois, Maryland, Ohio, Texas, New Jersey). (A few utilities in these states chose voluntarily to divest their generating assets anyway.) Whether the generating assets were divested or transferred to affiliates, the utilities affected typically retained some type of transition or "default service" obligation to continue to supply retail customers who had not chosen a competitive retail supplier

[4] Otherwise the magnitude and speed of stranded cost recovery was threatened.

(ESP) at prices determined through some type of regulatory transition "contract." The terms, conditions and durations of these obligations vary widely from state to state and sometimes vary significantly among utilities in the same state (e.g. New York). I will discuss default service obligations further below when I discuss the evolution of retail competition.

Table 1 displays the patterns of the migration of utility generating assets from regulated utility to unregulated "non-utility" status[5] through either divestiture to third parties or through transfers of generating assets to affiliates of the legacy regulated utility owners. It is evident from the data in Table 1 that generation restructuring activity peaked in 1999-2000 and that the initial focus on divestiture of generating assets was replaced by generating asset transfers to affiliates of the regulated utilities within a holding company structure. The latter holding companies still have common ownership of generation, marketing, retail supply, distribution and transmission and from this perspective continue to be vertically integrated. However, FERC and state regulations place restrictions on communications between utilities owning and operating transmission and distribution assets and those owning and operating unregulated generating plants, wholesale marketing and retail supply businesses. Generation restructuring of either form now appears to have largely been halted as a result of no additional state retail competition and restructuring programs being implemented after the California electricity crisis.

During the last few years there has also been a significant amount of entry of new unregulated generating capacity seeking to supply power to both unintegrated distribution

[5] Typically becoming Electric Wholesale Generating (EWG) companies as provided for in the Energy Policy Act of 1992.

companies and to vertically integrated utilities that have been encouraged or required by state regulators to meet their incremental generation needs through wholesale market purchases. Table 2 displays the patterns of entry of new generating capacity between 1997 and 2002. About 80% of this new generating capacity is unregulated “merchant” or “non-utility” capacity built to make sales in competitive wholesale markets. Very little new generating capacity was added anywhere in the U.S. during the mid-1990s, reflecting the perception that there was excess generating capacity in most regions of the country and uncertainties about the direction of restructuring and competition policies at the state and federal levels. As FERC issued new regulations governing transmission access and related wholesale market rules, as a growing number of states adopted retail competition programs, and as wholesale power prices rose, a large number of merchant plants were announced, began to seek construction permits, went into construction and were ultimately completed. By 2002, the amount of generating capacity completed reached 55,000 Mw per year, an order of magnitude greater than in 1997 and 1998 and a in total of nearly 140,000Mw of new generating capacity was completed between 1999 and 2002.[6] Most of this capacity is gas-fired and relies on clean and thermally efficient combined-cycle generating technology. Up to roughly the middle of 2001, investments in new merchant generating projects and trading power in wholesale markets was perceived to be a booming business with enormous profit opportunities and was pointed to as a notable success of policies aimed at stimulating competition in electricity. However, as I shall discuss in more detail presently, the boom has now turned into a bust with abundant generating capacity in service in almost all regions of the country, a merchant generating and trading sector in difficult financial shape, and many planned

[6] The U.S. has about 760,000 Mw of generating capacity in 1995.

new generating plants (even some under construction) being cancelled or indefinitely postponed.

The combination of divestiture of existing generating plants by utilities, transfers of utility generating plants to unregulated utility affiliates, and entry of new merchant generating plants (a significant amount of which is owned by utility affiliates), has led to a large increase in the fraction of electricity generating by unregulated "non-utility" generating plants. Table 3 displays the fraction of total electricity supplied in the U.S. that has been accounted for by unregulated "non-utility" power suppliers between 1990 and 2002.[7] By 2002 unregulated generators were producing about 33% of the total electricity supplied in the U.S. If we deduct generation supplied by municipal utilities, federal power projects, and cooperatives, unregulated private power generation now accounts for about 40% of the energy supplied by investor-owned companies. However, a significant fraction of this energy comes from generating plants held by affiliates of vertically integrated utilities and located in the same areas as their distribution and transmission facilities.

All things considered, there has been a very significant restructuring of the generating segment of the electric power industry in the last few years. However, there are significant uncertainties about how quickly the pace of generation restructuring will be in the future.

a. Since Texas, no additional states have adopted and implemented retail competition and industry restructuring programs. There does not appear to be a lot of enthusiasm in the remaining states to implement such reforms quickly or at all. These states are likely to continue to require or provide incentives for their vertically integrated utilities to look to the wholesale market for their incremental

[7] In 1990, virtually all of the non-utility generation was accounted for be cogenerators, industrial plants, and renewable energy facilities.

power supply needs, but there is an emerging trend toward utilities again looking to build regulated generating plants to serve their native load customers. While FERC has continued to push reforms to wholesale market and transmission institutions (more on this below), there has been growing resistance from many states to pro-competition policies.

b. The merchant generating sector is in terrible financial shape. What had been an enormous boom has now turned into an enormous bust. Table 4 displays the market values of a share of common stock for each of eleven companies with major financial commitments to energy trading and merchant generating capacity in May 2001 (peak week) and on March 10, 2003. The equity values of these companies has fallen dramatically in less than two years and only two of these companies (both with large regulated utility affiliates) now have investment grade credit ratings. About 125,000 Mw of new power projects that had been announced prior to 2001 have now been cancelled or indefinitely suspended in the last two years. A significant amount of additional capacity under construction or in the permitting process has been delayed.

The fact that many announced projects have now been cancelled is not surprising. Many more generating projects were announced than could possibly have been absorbed profitably by the market. As new generating capacity was completed and demand slowed with the slowdown in the economy, wholesale market prices softened and it became clear that many regions would have capacity surpluses in the immediate future. The market responding to these price signals with reductions and delays in investment in new capacity is what we should expect in a competitive market. However, the slowdown appears to be more than simply an ordinary market response to changes in supply and demand. The merchant generating and trading sector benefited from a "financial bubble" similar to the one supporting telecom and internet stocks, giving them easy access to cheap capital to finance new generating capacity without support from longer term contracts, as well as large energy trading operations that expanded well beyond trading around their own assets into highly speculative trading activities. The end of the stock market bubble, a better understanding by investors of the real economics and market risks associated with building merchant generating plants and trading commodity electricity, trading and accounting improprieties, credit downgrades, refinancing problems, and uncertainties about the future direction of industry restructuring, wholesale market rules, and retail competition, have decimated the merchant generating and trading sector, slashing investment and trading activity and dramatically increasing the cost of capital for new generating plants and very significantly reducing liquidity in forward electricity markets. Changes in capital market conditions and imperfections in wholesale market institutions are likely to create barriers to stimulating efficient investment in new generating capacity and efficient retirement decisions of existing generating capacity in the near future, a subject that I will return to in the section on "resource adequacy" below.

Despite all of the restructuring that has taken place, overall, the U.S. electric power industry continues with a substantial amount of vertical integration between competitive segments and regulated segments in the same geographic area. A majority of the states have decided against pushing forward with competitive market and restructuring initiatives, at least for the time being. In addition, transmission ownership and system operations continue to be very fragmented with ongoing barriers to the development of regional wholesale markets. Transmission network congestion has continued to increase. FERC has recognized and tried to respond to the problems that arise when transmission owners are not independent of the market participants that rely on the network, to problems of excessive fragmentation, and wholesale market inefficiencies. I turn to developments in these areas in the sections below.

WHOLESALE SPOT MARKET INSTITUTIONS

One of the most challenging things to explain to people who are not familiar with the unusual attributes of electricity is that wholesale electricity markets do not design themselves but must be designed as an integral central component of a successful electricity restructuring and competition program. Unlike England and Wales, Norway, Sweden, Spain, Australia, New Zealand, Argentina and most other countries, the U.S. did not proceed with its wholesale and retail competition initiatives with a clear coherent blueprint for wholesale market design, transmission institutions, or vertical and horizontal restructuring. Moreover, despite experience in other countries, the U.S. did not take adequate account of wholesale market imperfections and the incentives and opportunities created by market imperfections for individual suppliers to engage in bidding, scheduling, and trading practices that could increase prices above competitive levels and harm consumers (the distinction that some have drawn between "market power" and "exploiting market inefficiencies" to the same end, is in my view, meaningless).

FERC's efforts to reform wholesale market institutions began with Orders 888 and 889 in 1996. These orders basically established rules under which jurisdictional transmission owners were required to provide access to their transmission systems to third parties, and associated requirements to provide balancing and operating reserve services using formulas or procurement methods specified in FERC regulated transmission tariffs, to make information about the availability of transmission service, purchasing, and scheduling transmission capacity easily available to all market participants. However, Order 888's basic regulatory framework presumed that the prevailing structure of the electric power industry would remain largely unchanged, followed traditional utility transmission practices regarding the provision of transmission service to third parties. These practices were based on utility-specific "contract paths" that were often inconsistent with actual power flows and average "postage stamp" embedded cost transmission price ceilings.[8] Order 888 also gave the incumbents first refusal on available transmission capacity (they had paid for it after all), and relied on administrative rationing, rather than economic rationing, to allocate transmission constraints. Order 888 did not require utilities to operate transparent organized day-ahead or real time markets for energy or operating reserves but rather required transmission owners to provide balancing services and operating reserves as cost-based prices. The transmission owners administering the Order 888 tariffs generally owned generating capacity and used the same network to buy and sell wholesale power as did their would be competitors.

The three Northeastern power pools, California, Texas, and, most recently, several Midwestern states (Midwest ISO), took a more comprehensive approach to developing new wholesale market institutions. They created independent system operators (ISOs) to schedule and dispatch generation and demand on transmission networks with multiple owners, to allocate scarce transmission capacity, to develop and apply fair interconnection procedures for new generators, to operate voluntary public real-time and (sometimes) day-ahead markets for energy and ancillary services, to coordinate planning for new transmission facilities, to monitor market performance in cooperation with independent market monitors, and to implement mitigation measures

[8] Unlike a First Class postage stamp which allows one to send a letter to anywhere in the U.S. for 37 cents, "postage stamp" transmission tariffs only provide for transit to or through an individual utility's transmission network. The number of postage stamps required for transmission service depended upon how many utilities with transmission facilities happen to be on the contract path available between an injection point and a delivery point. This phenomenon is known as "pancaking."

and market reforms when performance problems emerged. FERC's proposed Standard Market Design Rules (SMD) would extend what it views as the best practices drawn from the experience with these ISOs to the rest of the country. However, at the present time, the U.S. has a patchwork of different wholesale market institutions operating in different regions of the country. Moving power between portions of the networks operated by different system operators is sometimes difficult or costly and coordination imperfections between control area operators increase costs of energy, operating reserves, and congestion especially during emergency conditions.

The debates over the design of these wholesale market institutions in California in 1996 and 1997 got the reforms off on the wrong foot. These debates were contentious and highly politicized, reflecting perceptions by various interest groups about how different wholesale market institutions would advance or constrain their interests and, in my view, an inadequate amount of humility regarding uncertainties about the performance attributes of different institutional arrangements. The discussion of alternative institutions was polluted by an unfortunate overtone of ideological rhetoric that attempted to characterize the debate about wholesale market institutions as one between "central planners" and "free market" advocates (Joskow, 1996). The market design process in California in 1997 and 1998 also demonstrates how market design by committee, reflecting interest group compromises and mixing and matching pieces of different market models, can lead to a system with the worst attributes of all of them (Joskow 2001).

There have been a number of studies of the performance of the wholesale market institutions performed by market monitors, consultants, and academics that have been created over the last several years. Let me summarize what I believe we have learned from the experience of the last several years.

a. All wholesale market designs work reasonably well in the short run when demand is low or moderate, supply is very elastic and generating resource ownership not too highly concentrated, there is little congestion on the transmission network, and barriers to scheduling power by competing generators are removed with reasonable transmission access and energy balancing procedures and prices. The performance challenges for wholesale market institutions arise during the relatively small number of hours each year when demand is high, supply is inelastic, transmission congestion is significant and widely dispersed. The performance of wholesale market institutions under these conditions is very important because this is when markets have to work hard to facilitate an efficient allocation of scarce resources. It is during these hours when competitive market prices <u>should</u> be high, when congestion is likely to lead to significantly different prices at different locations on the network, when a significant fraction of competitive market "rents" are produced to pay for the capital costs of investments in new generating capacity, when demand-side decisions are most important for providing investors with accurate price signals reflecting consumer preferences for investment in "reserve capacity" and reliability and for signaling consumer demand for risk hedging instruments to financial intermediaries.

   Unfortunately, it is also under these "tight supply" conditions when market power problems are most serious, when system operator discretion is most important, and when non-price rationing to balance supply and demand to maintain the network's physical parameters within acceptable levels are most likely to be necessary because incomplete markets cannot respond fast enough to rapidly changing system conditions that threaten network reliability. Price formation during these relatively few hours each year works its way through the system to affect forward prices, incentives for investment in new generating capacity and retirement decisions about existing generating capacity, and overall system reliability. In the end, the performance of wholesale market institutions should be judged primarily by how they function during these tight supply conditions.

b. The debate about whether wholesale markets should be organized around "bilateral contracts" or "centralized dispatch" that took place in California in 1996 and 1997 was an empty and unnecessarily confusing debate. Bilateral financial contracts and self-scheduling and dispatch of generators and load can and should be an important component of any wholesale market design. Nevertheless, there is an important role for the single system operator responsible for maintaining the necessary physical parameters of the network and facilitating economical decisions by buyers and sellers of energy and ancillary services. This role includes operating voluntary, consistent, and transparent real-time balancing and (ideally) day-ahead scheduling bid-based auction markets that are cleared by the system operator using a security-constrained dispatch algorithm that incorporates, as accurately as possible, the physical topology of the network (Hogan 1992, 1993). This process necessarily will yield shadow prices at each supply and demand node reflecting network constraints. The difference between the nodal prices at any point A and at any other point B on the network (ignoring losses) is a measure of the costs of congestion associated with an incremental injection of generation at node A and an incremental increase in demand at node B. At least in the real time balancing market, these are the proper prices to be used for settlement purposes for net sales and purchases of energy (deviations from day-ahead schedules) and pricing of congestion. Aggregation to larger zones on the supply side is, in my view, potentially problematic (and unnecessary) while aggregation on the demand side is probably not a big problem for groups of customers that unlikely to be subject to real time metering and pricing anyway.

c. The allocation of scarce transmission capacity day-ahead and in real-time should be fully integrated with the operation of day-ahead and real time energy markets as discussed above. Where this integration has not been achieved, as in California, New England, and Texas, congestion costs (or at least congestion rents) have increased beyond the efficient level.

d. The definition and allocation of transmission rights is an important aspect of wholesale market design and has implications for the design and performance of wholesale market institutions. In the short run, these rights serve as (imperfect) hedging instruments against (basis) differences between prices at different points on the network. While these hedging properties can in principle be replicated through combinations of forward contracts to buy and sell power at different locations, or by third-parties marketing derivatives on such transactions, thin markets and potential market power problems may limit the development of liquid forward and derivative markets to a much smaller number of hubs. <u>Long-term</u> transmission rights are also likely to be important for securing financing for investment in new transmission facilities that are built on a merchant basis. There has been substantial debate about whether such rights should be point-to-point financial rights in conjunction with nodal pricing or physical flowgate rights in conjunction with a zonal pricing system. In theory these alternatives may look very similar if there is little intra-zonal congestion, though physical rights raise additional operational issues (Joskow and Tirole 2000). In practice, my view is that financial rights are much easier to implement, involve lower transactions costs and are more difficult to use to exercise market power than are physical rights on a meshed network. They are also easier to make compatible with real time balancing actions and the associated real time balancing prices. On radial lines and DC inter-connectors, physical rights may be more attractive, especially if the projects are developed on an unregulated merchant basis. I realize that the relative performance attributes of physical and financial rights, and of nodal vs. zonal pricing systems, is a matter of dispute and I would welcome careful empirical analysis of the relative performance attributes of the different system.

e. Day-ahead and real time markets for energy and ancillary services should be fully integrated and reflect the efficient optimization of energy supply and operating reserve resources that can both supply energy (or reduce demand) and stand in reserve to respond to short run fluctuations in demand and unanticipated outages of generating or transmission facilities. Wholesale market designs that separate energy and individual ancillary services markets have performed poorly and are subject to unilateral behavior that increases prices and reduces efficiency.

f. The unusual attributes of electricity discussed above create unusual opportunities for suppliers to exercise market power unilaterally. Market power problems have been extensively documents in the U.S. and other countries (Wolfram 1999, Borenstein, Bushnell and Wolak 2002, Joskow and Kahn 2002). The incentive and ability to exercise market power is <u>enhanced</u> by poor wholesale market design attributes that expand opportunities for suppliers to take actions that affect market prices. Accordingly, the arguments about whether market performance problems are due to poor market design <u>or</u> market power is an empty debate. The former simply enhances the latter. Because electricity demand is very inelastic in the short run and electricity cannot be stored, individual suppliers may be able to move prices significantly even in markets that are not very highly concentrated by traditional standards. This is most likely to be the case when capacity constraints are approached and a large fraction of demand is being served by spot market purchases in day-ahead and real time markets rather than pursuant to forward contracts that establish the transactions price in advance. The allocation of transmission rights may enhance market power in generation markets as well (Joskow and Tirole 2000). As a result, market design, transmission rights allocation, mergers and acquisitions of generating facilities, and market monitoring institutions must be sensitive to the potential for serious market power problems to emerge in unregulated wholesale power markets.

g. Expanding significantly the requirement that, as the default, larger retail consumers be billed based on their real time consumption and associated real time prices for energy can help to improve wholesale market performance in a number of dimensions. This will allow consumer preferences for reliability and market price volatility to be more accurately represented in the wholesale market, help to mitigate market power, flatten load duration curves and reduce the need for capacity that operates for only a few hours each year, encourage risk-averse consumers to cover their demand with forward contracts, and reduce the need for market power mitigation regulations and resource adequacy rules (see below). At least some of the benefits of real time pricing are public goods that benefit all consumers. Accordingly, it would be worthwhile to provide financing to distribution companies to expand the default utilization of real time meters.

h. While it is important to move price sensitive consumers on to real time pricing for their price sensitive demand, it is also important to provide appropriate incentives for load serving entities and consumers to rely more on forward contracting and less on spot market purchases to reflect their preferences for risk and the costs and benefits of insurance against price volatility. This will help to mitigate market power by reducing supply incentives to withhold output from the spot market and drive up spot market prices and provide better signals and financing support for investments in new generating capacity. The real time pricing goal and the forward contracting goal are not incompatible. If the default is real time pricing, consumers who are risk averse will have a stronger incentive to go out into the market and enter into forward contracts for some or all of their demand. Consumers who are less risk averse or can adjust more easily to real time price variations will leave more of their demand in the spot market to enable them to respond to short-term price signals and will cover less with forward contracts.

i. There are important linkages between wholesale market institutions, retail procurement, and retail competition institutions. The infirmities of retail competition institutions in many of the states that have implemented retail competition programs (see below) and the uncertainties about whether, when and how retail competition will spread to other states, has undermined incentives for distribution companies and other load serving entities to enter into longer term contracts for power and transmission service with potential investors in generating and transmission capacity. This situation will exacerbate the boom-bust cycle of the industry, undermine investment incentives, increases incentives for suppliers to exercise market power when supplies are tight, increases prices volatility, and ultimately increases political pressures for regulatory interventions when prices are high.

j. Fraudulent and misleading accounting practices, false reporting of trading information to private entities publishing price indices upon which suppliers, traders and consumers rely, and the general trend of the late 1990s to make as little information as possible available to the public, has undermined the confidence of consumers, regulators, and their representatives in state legislatures and Congress in competitive electricity markets. It has also led to financial disaster for many merchant generating and energy trading firms and destroyed liquidity in forward electricity markets. Well functioning competitive electricity markets depend upon the existence of efficient liquid forward markets for energy and associated derivative financial instruments to allow consumers and generators to manage their risks efficiently. Restoring

> public and investor confidence in energy trading and creating liquid transparent markets for physical contracts and supporting financial instruments to meet the needs of the ultimate buyers and sellers of electricity in these markets is an important policy priority.

Many features of FERC's Standard Market Design (SMD) NOPR issued on July 31, 2002 reflect these considerations and for this reason provide a useful, though far from perfect, blueprint for wholesale market design.[9]

RETAIL COMPETITION

Most of the academic research and many of the public policy debates about competitive electricity markets have focused on the design and performance of wholesale market and transmission institutions. However, from a political perspective, the primary selling point for competition in electricity among consumers and government officials has been the prospect for <u>retail competition</u> or retail "customer choice" to lead to lower <u>retail</u> electricity prices. Yet, there has been very little work assessing the performance of retail competition programs in the U.S. and there is a growing perception that, at the very least, retail competition programs have had disappointing results, especially from the perspective of residential and small commercial customers. It is too early to provide a comprehensive assessment of retail competition programs, many of which have not been in operation for very long. Moreover, inadequate information, especially on prices and value added services, makes it difficult to perform a good assessment. Nevertheless, there are things to learn from the experience to date and the data that are available.

With a retail competition program, an electricity customer's bill is "unbundled" into regulated components P_R (transmission, distribution, stranded cost recovery, retail

---

[9] The SMD NOPR is much less successful in dealing with transmission investment, organization and incentive issues. See the Comments that I submitted to FERC on these issues at http://econ-www.mit.edu/faculty/pjoskow/papers.htm

service costs to support default services) and a competitive component P_C (generation service, some retail service costs, and perhaps an additional "margin" to induce customers to shop). The customer continues to buy the regulated component from the local distribution company and is free to purchase the competitive component from competing retailers or retail Electricity Service Providers (ESP).[10]

In most jurisdictions that have introduced retail competition programs, the incumbent distribution company is required to continue to provide "default service" of some kind to retail consumers who do not choose an ESP. The terms and conditions of default service vary across the states, but typically have been calculated in the following way. The regulators start with the incumbent's prevailing regulated cost of generation service. A fraction of this regulated generation cost component is determined to be "stranded generation costs" that can be recovered from retail consumers over some time period and is included in P_R. The residual, reflecting an estimate of the competitive market value of generation services, plus some fraction of retail service costs (metering, billing, customer call centers) is then used to define the initial "default service" price P_C or the "price to beat" by ESPs seeking to attract customers from the regulated default service tariffs available from the incumbent utility.[11] The value of P_C is then typically fixed for several years (sometimes with adjustments for fuel prices) but is expected

[10] The ESP may, however, bundle the regulated delivery services and competitive services together, reimbursing the distribution company for P_R.

[11] Some states (e.g. Massachusetts) define the default service price, "price to beat," or "price to compare" with reference to the competitive component P_C only since the P_R component is set by regulation and must be paid to the distribution company, regardless of which ESP supplies a customer with unbundled competitive services. Other states (e.g. Texas) define the "price to beat" with reference to the total of the regulated and competitive components (P_R + P_C). In the latter case, ESPs quote customers a price for the combination of the regulated T&D charge and the competitive generation service charge they are offering. Arithmetically, the two approaches are the same, but they may have different implications from a marketing perspective.

eventually to reflect the competitive market value of providing competitive retail services to consumers.

Where incumbents have significant stranded costs, which is the case in most of the states that have introduced retail competition to date, regulators have been able to capture a small fraction of this sum for customers by mandating an initial retail price reduction that is reflected in P_R, though in many cases this is simply a deferral of recovery of these costs to future years. Regulators in a number of states have consciously left some of the stranded costs in the default service price P_C so that it exceeds the wholesale market value of the associated generation services in order to encourage customers to switch to competitive retailers, though in a number of cases the utility may ultimately be able to recover any associated losses in stranded costs in the future through surcharges included in P_R.

Things become more complicated when the incumbent utility has a prevailing regulated price whose regulated generation cost component is <u>below</u> the competitive wholesale market value of the associated electricity. And there are many utilities around the country that have regulated generation cost components of their retail rates that are <u>below</u> the competitive wholesale market value price of electricity (mostly in states that have decided against implementing retail competition programs and this is not a coincidence) In this case, the bundled price is also below the price that would prevail if a customer purchased regulated services at the regulated price and competitive services at their market values; the incumbent utility effectively has negative stranded costs. Regulators could handle negative stranded costs symmetrically with positive stranded costs by setting the price for competitive service P_C at its competitive market value and then providing a "stranded benefit" credit in the distribution charge P_R to settle up on the historical "regulatory bargain." However, this is not how regulators have handled prices for the few companies with negative stranded costs in states that have implemented retail competition programs. Instead they have just unbundled (more or less) the low prevailing regulated generation cost component of the regulated bundled rate and established that as the default service price. In these cases, ESPs that purchase power in competitive wholesale markets cannot compete with the default service price P_C since it reflects prevailing embedded costs of generation that are below the competitive market value of generation services. Moreover, as wholesale market prices have risen over time, the initial values of P_C in states where utilities did have stranded costs and that were typically frozen for several years, wholesale market prices for electricity subsequently rose (unexpectedly) to levels above P_C. In many states, customers are free to return to the default service tariff if the prices offered by competitive retailers are higher than the distribution company's default service prices and this is exactly what has happened in some states.

Consumers can benefit in at least four ways from the introduction of retail competition. First, even if they do not switch to an ESP they may benefit from reductions in regulated prices that have typically accompanied the restructuring process as an outcome of the bargaining over stranded cost recovery and the terms and conditions under which the incumbents can move their regulated generating plants into unregulated affiliates. Second, consumers can benefit by receiving lower prices than the default service price P_C from an ESP that has competed successfully for their business. Third, ESPs may offer consumers a variety of value added services, including price risk management, demand management, and energy efficiency services. Finally, competing ESPs may be able to provide "retailing" services more efficiently that can the incumbent. However, here we must recognize that retail service costs are a small fraction of a typical customer's bill, amounting to 0.3 to 0.4 cents/kWh or about \$3 - \$7 per month for a typical residential customer (depending on assumptions about fixed vs. variable components of retail service costs) (Joskow 2000b). Since the incumbent monopolies did not have to incur marketing and advertising costs to attract customers, these are additional costs that are not now reflected in regulated retail prices but would have to be incurred by ESPs

Switching Activity

Rational residential consumers will switch to an ESP if the incumbent utility's default service price P_C plus any transactions costs associated with switching suppliers is less than the price offered by competitive ESPs plus the value of any value added services that the ESP provides. There is reasonably good data available (with some effort) to measure the extent to which customers have switched to ESPs and I will focus on that information first. Switching behavior by consumers reflects their revealed preference for default service or ESP service, though absent information on ESP prices and value added services we cannot disentangle switching costs, commodity price differences and the value of any value added services provided to consumers. In Tables 5 to 9 and Figures 3 and 4, I have displayed the most recent data available for the fraction of residential, commercial and industrial customers who have switched to ESPs in several representative states that have implemented retail competition programs. The states that have introduced retail competition provide different amounts of information regarding customer switching activity, so the information is not always comparable across states.

Let us start with Massachusetts (Table 5), which was the first state to make retail competition available to all customers in March 1998, so that retail customer choice has been available for almost five years. During that period of time, less that 3% of the residential customers and less than 9% of the small commercial and industrial customers have switched to ESPs. About 30% of the largest industrial customers have switched to ESPs. Within each customer class, the customers consuming more electricity are more likely to switch to an ESP. Overall, 3.4% of the retail customers and 22% of the retail demand is supplied by ESPs. The remaining customers are served on one of two default service tariffs (called “standard offer service” and “default service” in Massachusetts, with the latter tracking wholesale market prices). During the period of time that retail choice has been available, the standard offer tariff has often provided electricity at a price below the wholesale market price, making ESP supplies unprofitable. The default service price fluctuates with wholesale market conditions, but leaves little if any margin for ESPs to recover their retailing costs. While all consumers have the right to choose their electricity supplier in Massachusetts, there is presently little if any competition from ESPs to serve smaller customers.

Let us turn next to New York (Table 6), where retail competition became available between May 1998 and July 2001 depending on the utility service area and the type of customer. There has been somewhat more switching than in Massachusetts, but the basic patterns and economic incentives are very similar. In both Massachusetts and New York, there has been some switching back and forth between ESPs and default service tariffs as wholesale prices have fluctuated and ESPs have exited the market. In Maine (Table 7) as well, residential and small commercial customers have not switched to ESPs (or those who switched have now switched back to utility default service), though there has been more switching by larger industrial customers. We can compare Massachusetts, New York, and Maine to New Jersey (Table 8), where retail competition began in November 1999. After over three years of customer choice there are almost no customers of any kind who have chosen to be supplied by an ESP in New Jersey, largely because the regulated default service price has been below the comparable wholesale market price for electricity. The situation in New Jersey appears to have evolved into a wholesale competition model in which the incumbent distribution company purchases wholesale generation services through a state-approved competitive auction process and then pass along the associated costs to retail customers. At the present time in New England, New York and New Jersey, a residential or small commercial customer who wanted to be supplied by an ESP would find it hard to find one. Most of those that were active in this region initially have now exited the market or are not actively marketing services to residential and small commercial customers.

Pennsylvania provides an interesting example of the interactions between changing wholesale market prices, default service prices, and the initial regulated utility cost conditions that had an impact on the level of these default service prices. Figure 3A and Figure 3B display the fraction of retail demand that has been served by ESPs at various points in time from April 2000 through January 2003 for each of six major Pennsylvania utilities. Figure 3A displays this information for residential customers. Except for Duquesne, all of these utilities are in PJM and buy and sell power out of the same wholesale market. Yet each of them has a different default service price, reflecting the different levels of their regulated retail prices (and stranded generation costs) prior to restructuring. Note first that there is a very large variation across utilities in the fraction of residential demand served by ESPs. Those with the highest switching rates (PECO and Duquesne) serve Philadelphia and Pittsburgh, had high regulated retail prices and significant stranded costs prior to the implementation of the retail competition and restructuring program.[12] In their restructuring settlements, the value of P_C was set relatively high compared to then-prevailing wholesale market prices, reflecting stranded cost allocations. The "price to beat" faced by ESPs was sufficiently high relative to then-prevailing wholesale market prices to make it profitable, at least initially, for ESPs to make offers to customers that were sufficiently far below the default service price to get a significant number of customers quickly to switch to an ESP. The other utilities in Pennsylvania had relatively low regulated rates prior to restructuring and relatively low values for P_C. Since ESPs must buy power in the wholesale market at higher prices than the utility default price available to retail customers they have not been able profitable to offer attractive competitive alternatives to retail customers in these areas. Note as well that over time, the fraction of residential customers in Pennsylvania receiving service from an ESP has <u>declined</u> very significantly. This reflects the fact that as wholesale prices rose after the default prices were set, and retailers realized how costly it actually is to serve residential customers, they raised their prices, or withdrew from the market, and residential customers they were serving found it advantageous or necessary to return to their utility's default service tariff. Based on the results of PECO's recent effort to

[12] Duquesne's stranded costs have now been recovered and the associated charges eliminated from the prices that it charges for distribution service.

auction off some of its small commercial default service customers to ESPs, there seems to be little interest among ESPs in provided service at a price significantly below PECO's default service price.[13]

Table 3B displays the same data for industrial customers. A larger fraction of industrial customers initially switched to ESPs. Then many of them returned to default service when wholesale prices rose. Some now appear to be returning again to ESPs in a few utility service areas. This back and forth between utility default service and competitive ESP service is a consequence of the "safety net" provided by the regulated default service prices available from the incumbent utility supplier.

Ohio also provides an interesting example of how the interaction between regulation, default service terms, and wholesale market prices affects consumer switching behavior. Retail competition began in Ohio in January 2001. Table 9 displays the fraction of customers served by ESPs in different utility service areas. In the areas served by subsidiaries of AEP and Dayton Power and Light, essentially no customers are served by ESPs. These are utilities that had very low regulated retail prices and the value of P_C establish at the beginning of retail prices reflected these low regulated rates. The regulated default prices in these areas are lower than the wholesale market price. However, in the areas served by First Energy's subsidiaries there has been very significant movement of retail customers to ESPs. Indeed, residential customers appear to have switched at the same rate or even at a greater rate than industrial customers. First Energy's subsidiaries had high regulated retail prices, the value of P_C was consciously set

[13] The Pennsylvania Public Utilities Commission required PECO to auction off 64,000 commercial customers because too few customers had chosen voluntarily to be served by ESPs. The average discount offered by the winning bidders was 1.25% off the default service price. Customers can opt out of the "draft" if they fill out a postcard or call PECO. Dow Jones Business Service, February 27, 2003. http://biz.yahoo.com/djus/030227/1853001167_1.html .

at a relatively high level to put recovery of some of the utilities' stranded costs at risk, and there were some other incentives related to stranded cost recovery that made it in First Energy's interest to migrate customers to ESPs. Moreover, the data for residential customers is misleading. Ohio has a municipal aggregation program that allows municipalities to purchase power on behalf of the customers within their municipal boundaries (customers have an opt-out option). About 90% of the residential customers listed as being served by ESPs are actually being served through municipal aggregation programs in which an ESP can effectively serve all of the customers located in that municipality. This is really a wholesale competition program rather than a retail competition program.

Finally, let's turn to Texas (Figures 4A, 4B and 4C). Retail competition began officially in Texas in January 2002, though there was a pilot program implemented before that and customers who had switched before the official program began could stay with the ESPs they had chosen. By the end of 2002 between 4% and 10% of the residential customers had switched to ESPs, and the fraction continues to grow (rather than to decline as in Pennsylvania). For commercial customers, 10% to 15% had switched to ESPs by the end of 2002, while 20% of the largest customers, accounting for 50% of the revenues from large industrial customers, had switched to an ESP, and virtually all of the largest customers have negotiated competitive contracts either with the retailing affiliate of their incumbent utility or an unaffiliated ESP. So, in only a year, and despite some initial technical problems experienced with switching customers, Texas appears to be on a trend in which customers are migrating relatively quickly to ESPs. It is also the state that has the largest number of active ESPs competing to sell service to retail consumers.

We can compare the switching experience with that in England and Wales since 1990. In England and Wales retail competition started with the largest industrial and commercial retail customers (> 1 Mw initially) and was gradually expanded to industrial and commercial customer groups with smaller consumption levels (peak demand greater that 100 Kw in 1994). By 2000, about 80% of industrial and commercial customers with peak demands above 1 Mw and about 70% with peak demand below 1 Mw (and greater than 100 Kw) had switched to a competitive retail supplier other than the incumbent distribution company's ESP affiliate. These switching shares have increased monotonically over time (DTI, 2000). In May 1999, retail competition was opened up to residential (domestic) customers. Initially, the prices that could be charged to residential customers by the incumbent electricity distributor-affiliated ESP were regulated with a price cap and the incumbent retail supplier had a continuing obligation to provide service to residential customers in their service areas at a price no higher than the regulated supply price. By the end of 2001, about 30% of the residential customers had switched to competing ESPs and by mid-2002 this fraction had reached 34% (DTI 2003, p.97). In April 2002, the default service price caps applicable to residential customers were lifted and retail supply prices were fully deregulated for all types of customers (OFGEM 2001, DTI 2003). The retail competition program in England and Wales has been much more successful in facilitating migration of retail customers from default service to non-incumbent ESPs than have the programs in the U.S.

Retail Price and Value Added Service Effects

Unfortunately, there is very little information available to evaluate systematically the price effects of retail competition or even the level of retail prices charged by ESPs in the U.S. Nor is there information about the diffusion of the value added services that many retailers argued would be made available to retail customers. The Energy Information Administration (EIA) has begun to collect retail price data separately for utility and ESP supplied retail service, but sufficient data are not yet available to do a proper comprehensive analysis. Moreover, the relevant comparison is not between what retail prices are today and what they were when retail competition began, but rather the difference between what they are with retail competition and what they would have been without it. This will be a difficult counterfactual analysis to undertake, requiring an incorporation of fuel price changes, other changes in O&M costs, changes in the utilities' rate bases, cost of capital, etc., to measure what regulated prices would have been absent restructuring settlements.

Since so few customers have switched to competitive retailers in most states, it must be the case that the bulk of any savings that they have achieved are attributable to reductions in regulated prices implemented as a consequence of restructuring compared to what these regulated prices would have been absent restructuring. This is especially true for residential customers. The Texas PUC estimates that "price-to-beat customers" (residential, commercial, and industrial customers with peak demands below 1 Mw) saved \$902 million in 2002 if they stayed on the "price to beat" rates. This sum is composed of \$262 million attributed to a mandated 6% retail price reduction included as part of the restructuring program and \$677 million attributed to reduced fuel costs and the expiration of fuel surcharges.[14] Roughly 10% of these customers have switched to ESPs and, assuming (contrary to fact) that they all switched to the ESP offering the lowest price, this would imply another $64 million in savings directly attributable to retail competition for residential customers in 2002.[15] Figure 5 displays information on the price to beat and the competitive ESP prices being offered to residential customers with different average monthly consumption levels for one utility service area in Texas during 2002.[16] The discounts from the incumbent's price to beat are in the 4% to 10% range. This is a similar magnitude to the discounts offered in England and Wales by non-incumbent ESPs competing for residential customers in England and Wales. Discounts of this magnitude are apparently sufficient to overcome transactions costs associated with switching by roughly 10% of the residential and smaller commercial customers during the first year that the retail competition program is in effect. However, so far it appears that the bulk of any retail price savings resulting directly from retail competition have accrued primarily to industrial customers in Texas and most other states that have implemented retail competition programs.

We can gain some further insights into the challenges associated with measuring the impact of retail competition on retail prices by comparing the patterns of retail prices charged to residential customers by utilities between states that have introduced retail competition and those that have not. 90% or more of residential customers continue to

[14] It is difficult to believe that it is appropriate to attribute fuel price reductions and the end fuel surcharges as a benefit of restructuring. These costs were traditionally treated as pass-throughs under regulation. Moreover, the relevant comparison is not with the 6% mandatory reduction from 1999 prices, but what prices would have been if standard regulatory procedures had been followed.

[15] Public Utility Commission of Texas, February 2003 Report Card on Competition.

[16] Many ESPs offer flat rates annual rates while the incumbent supplier's price to beat may vary from month to month. Accordingly, annual comparisons provide a better picture than do monthly comparisons.

take service from their local utility in most states that have introduced retail competition. If there were significantly more attractive prices offered by ESPs we would have seen much more switching by now, as in Texas. Figure 6 displays the patterns of average residential prices for eight states with the most active retail competition programs, along with the average residential price for the U.S. as a whole for the period 1995 through 2002.[17] Figure 7 displays the percentage change in nominal residential prices for each of these eight retail competition states and for the U.S. as a whole. Note first that except for Texas, all of these pioneer retail competition states had prices in the mid-1990s that exceeded the national average, typically by a substantial amount.[18] Retail prices for residential customers in Texas in 1995 were just slightly below the national average. On average residential prices fell by 3.72% in these eight retail competition states between 1995 and 2002, compared to an average increase of 0.5% for the U.S. as a whole. The two states with the largest reduction in residential prices are New Jersey and Illinois. Both states have essentially no residential customers who have switched to an ESP, so any savings from restructuring are reflected entirely in changes in the prices of regulated default service. Since seven of the eight retail competition states had residential prices that were above the national average it is also likely that the prices in these states would have fallen more than the average if regulation had continued, reflecting the declining

[17] These price data come from the March Issue of the EIA publication *Electric Power Monthly*. They are far from ideal since they include municipal and cooperative as well as investor-owned utility sales and revenues. The 2002 data are preliminary and the numbers for California look wrong to me. Further analysis as additional data become available would be desirable.

[18] Connecticut, Rhode Island, Illinois, Michigan, Maryland, and Delaware also had above average retail prices. The only states that have introduced retail competition for residential customers that had below average prices were Texas and Virginia. Oregon, which has very low retail rates, technically introduced retail competition for large industrial customers only in 2002. However, no customers have chosen to be served by an ESP because the rate options offered by the local utilities (including a floating market-based rate option) are cheaper.

rate bases associated with costly nuclear power plants, increased capacity utilization as demand increased, and the gradual end of high-priced QF contracts signed during the 1980s. Texas, the only state in the group which had residential prices below the national average in 1995, experienced the largest increase in prices by 2002. If retail competition per se has had any effect on residential prices in the U.S. it must be small. However, this is an area where a more comprehensive analysis of counterfactual regulated prices and ESP prices would be valuable when and if the necessary data become available.

I think that it is fair to say that retail competition is still a work in progress in the U.S. and that retail competition has been a disappointment in many of the states that have implemented it. It should be no surprise that the remaining states, which typically have much lower regulated prices than did the retail competition pioneer states, would not find the performance to date to provide a particularly compelling case to introduce retail competition. Factors that need to be recognized and issues that need to be addressed in reforming existing retail competition programs or designing new programs include the following:

a. If the primary selling point for introducing retail competition continues to be that it will result in significantly lower retail prices from those that prevail today, retail competition will be a very tough sell politically in many of the states that have not already embraced it. In the states that have introduced retail competition, few residential or small commercial customers have chosen to be served by ESPs and the direct benefits of retail competition per se, in terms of lower prices, appear to be very small. Large industrial and commercial customers appear to have done much better in states that had relatively high regulated rates, at least based on revealed preference in the absence of publicly available retail price data for these customers. Many of the states that have not introduced retail competition have relatively low regulated retail prices; many with regulated prices for generation services that are lower than comparable wholesale market prices. Unless state regulators are willing to allow retail prices to rise to reflect the competitive market value of generation services, ESPs will not be able to compete against these low regulated generation prices if they are unbundled at their current or lower levels. Accordingly, it is not surprising that there is little interest in retail competition in those states that have low regulated retail prices.

b. If default service prices are set below the comparable wholesale market price of power, ESPs will not be able to compete for retail customers. Moreover, allowing customers that choose to take service from an ESP to return to a regulated tariff when wholesale prices are high without being charged an appropriate price for this option, seriously undermines the development of retail competition because it effectively provides a subsidized option for retail customers who switch back and forth and a very unstable customer base for ESPs.[19] This subsidized default service option also has adverse effects on wholesale markets by discouraging both utilities with default service obligations and ESPs from entering into long term forward contracts since under the rules in place in most states, consumers are free to come and go from the regulated default tariff as they choose.

c. Even if the regulated default service price P_C is higher than the comparable competitive wholesale market value of the power supplied, ESPs need an additional margin both to induce sticky retail customers to switch suppliers and to cover their retail supply costs. The retail supply costs for the mass market (residential and small commercial) are much higher than many retailers had anticipated. Billing, customer service, bad debt, advertising and promotion costs add up quickly. Moreover, there are significant economies of scale associated with several components of these costs. Accordingly, P_C may have to be much higher than the comparable wholesale market price to induce much customer switching. The evidence from England and Wales suggests that price reductions of 5% to 10% of the <u>total bill</u> are necessary to get significant customer switching for mass market (residential and commercial) customers. This is consistent with the limited experience in Texas as well. If the generation component of the retail price is 50% of the total bill, then price reductions of 10% to 20% on the generation component are necessary to get significant switching. To this must be added about another 5% to 10% for retail service costs. So, a margin of 15% to 30% between P_C and the comparable wholesale market value of power may be necessary to induce significant switching by residential and small commercial customers. In many areas of the U.S. this kind of margin is incompatible with reducing retail prices from their prevailing regulated levels.

d. It is far from obvious to me that residential and small commercial customers have or will benefit much, if at all, from retail competition compared to a regime where their local distribution company purchased power for their

[19] This is not the case in Texas where customers that cease being served by an ESP are put on a Provider of Last Resort (POLR) rate that is market-based (through an auction) and reflects the attributes of customers who are likely to be dropped by their ESP. However, it also is not the case in Massachusetts and in New York City where default service customers are served on a separate market-based default service tariff.

needs by putting together a portfolio of short term forward contracts (from days to several years) acquired in wholesale markets (Joskow (2000, 2000b).[20] This is effectively what is happening in most of the states that have introduced retail competition anyway, except the wholesale contracts are largely short-term and the threat of draconian policy changes to “make retail competition work” lead both utilities and ESPs to be very reluctant to make longer term commitments for power supplies of any kind. There is no evidence that residential and small commercial customers are getting any significant value added services that they find sufficiently attractive to shop and choose an ESP. For those states that have not introduced retail competition it would be worth considering introducing it first for larger customers and then expanding eligibility over time as the retail market institutions develop.

If policymakers are committed to fostering retail competition for residential and small commercial customers, despite the possibility that retail prices will rise in the short run due to increased transactions costs, switching costs and market power, the framework adopted by Texas, a framework with many similarities to what was adopted in England and Wales, is likely to be the most successful in stimulating retail shopping and the development of a viable retail supply sector. It has several components:

a. All retail supply functions of the incumbent regulated utility are shifted to a retail supply affiliate, and ideally, this retail supply affiliate should be separated through sale or flotation from the regulated delivery business. All of the associated retail supply costs are unbundled from the distribution and transmission charges and included in the default service price. ESPs then have a shot at a P_C that includes all of the incumbent’s retail supply costs. The retail supply affiliate has the obligation to provide regulated default service to retail customers for a defined period of time or until market conditions are sufficiently competitive to deregulated retail prices completely. The decision to fully deregulate should be sensitive to evidence that residential customers have high switching costs and that the incumbent may have significant market power for some period of time (Giulietti, Price, and Waterson 2003).

b. The default service price should reflect the competitive wholesale market value of electricity over the period during which the default service prices are to be in effect plus retail service costs. The incumbent supplier should be free to cover these commitments with a portfolio of contracts and ownership of generating assets and to try to reduce retail supply costs under a price cap and subject to quality of service criteria and penalties for failing to meet them.

[20] See Littlechild (2003) for a different view.

For utilities that have stranded costs, they can be recovered through a non-bypassable surcharge attached to distribution or transmission charges. Utilities whose regulated prices for generation service are below competitive market values should be treated symmetrically. A measure of the associated "stranded benefits" should be developed and accounted for as a rebate attached to regulated distribution and transmission prices. Thus, even in states with very low regulated prices, the unbundled generation component will reflect prevailing market values while consumers still capture their entitlement to the "cheap" power from generating plants they have paid for through regulated prices through lower delivery charges.

c. Once customers turn to an ESP they do not get to come back to the default service price. Instead they are served by a Provider of Last Resort (POLR) at prices determined through an auction of the POLR service responsibility.

d. When wholesale and retail markets are deemed to be sufficiently mature and competitive, retail generation prices are deregulated and the incumbent's retail supply affiliate has the opportunity to charge any remaining customers an unregulated price. This provides the incumbent retail supplier with a potentially valuable opportunity to supply the stickiest customers at unregulated prices if the competition in the retail market evolves sufficiently to trigger the deregulation brass ring. This potential business opportunity also makes it fair for the incumbent to be at risk for some of its retail service costs, many components of which are fixed in the short run and needed to serve the initially large default service customer base. Rather than being indifferent to whether retail customers switch to an ESP or not, which is the case in most states where retail competition has been implemented, the incumbent has an interest in encouraging enough of its customers to turn to competitive retail suppliers so that the competitive retail supply market will flourish.

LONG TERM RESOURCE ADEQUACY

Despite the substantial amount of new generating capacity that has been completed in the last few years, and what appears to be more than adequate capacity to meet peak demands in most regions of the country, there are growing concerns that wholesale energy and operating reserve markets will not provide adequate incentives to bring forth sufficient new generating capacity in the future to meet traditional reliability criteria. These concerns reflect a number of phenomena, including the financial collapse of the merchant generation and trading sector, the cancellation of many planned generating projects, tougher financial requirements and increased cost of capital to finance new merchant projects, very thin markets for medium and long term forward contracts for energy and operating reserves to help to support financing of new projects, proposals to terminate installed capacity (ICAP) obligations that now exist in the three Eastern ISOs, the effects of price caps and various bid mitigation rules that are applied to spot markets, and increasing pressures to close older existing plants as they face low wholesale prices and additional costs to meet new environmental requirements. FERC's SMD 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. It cites a number of market and institutional imperfections that could lead to under-investment in the future.

I think that there are good reasons to believe that spot market prices for energy and operating reserves alone, as they are now constituted and in the present state of partial and somewhat chaotic transition from regulated monopoly to competitive markets, are unlikely to provide adequate incentives to achieve generating capacity levels that match consumers' preferences for reliability. A variety of market and institutional imperfections contribute to this problem. I will present some empirical evidence below that reinforces this conclusion.

Questions about whether competitive wholesale and retail markets would produce adequate generation investment incentives to balance supply and demand so as to match consumer valuations of reliability have been raised since the transition to competitive electricity markets began. Until 2001, the system in England and Wales provided for additional capacity payments to be made to all generators scheduled to supply during hours when supply was unusually tight (high loss-of-load probability).[21] The markets administered by the Eastern ISOs continued their traditional policies of requiring distribution companies (or more generally "load serving entities" (LSE) to encompass ESPs) to enter into contracts for capacity to meet their projected peak demand plus an administratively determined reserve margin. Similar requirements continue to be applied by utilities in those states that have not introduced retail competition and continue to rely primarily on vertically integrated utilities which may purchase some of their forecast capacity needs in the wholesale market. Argentina's competitive electricity market system also included capacity payments to stimulate investment in reserve capacity. In Chile, distribution utilities are required to enter into forward contracts to meet forecast demand plus a reserve margin. However, California's electricity market design did not impose capacity, reserve or forward contract obligations, nor does Texas.

At it's core, the questions about whether wholesale markets will bring forth adequate investments in generating capacity arises from the unusual characteristics of electricity supply and demand discussed earlier: (a) large variations in demand over the course of a year; (b) non-storability; (c) the need to physically balance supply and demand at every point on the network continuously to meet physical constraints on voltage, frequency, and stability; and (d) that even under the best of circumstances (i.e. with effective real time pricing of energy and operating reserves) non-price mechanisms will have to be relied upon from time to time to ration imbalances between supply and demand to meet physical operating reliability criteria because markets cannot clear fast

[21] This payment mechanism was dropped when the New Electricity Trading Arrangements (NETA) system was introduced in 2001.

enough to do so without unplanned outages. These attributes have a number of implications. First, a large amount of generating capacity that is available to meet peak demand plus the associated operating reserve requirements supplies relatively small amounts of energy during the year. For example, in New England in 2001, 93% of the energy was supplied by 55% of the installed generating capacity while the remaining 45% of the capacity supplied only about 7% of the energy.[22] Potential investors in new generating capacity must expect to cover their variable operating costs, their fixed operating and maintenance costs, and their capital costs from sales of energy and operating reserves over the life of generating capacity under consideration. The return of and on the associated capital investment in new generating capacity is the difference between the prices they receive for generation services (including capacity payment, if any) and their operating (primarily fuel) costs. The profitability of generating units that are likely to operate only for a relatively small number of hours in each year ("peaking capacity") are especially sensitive to the level of prices that are realized during the small number of high demand hours in which they provide energy or operating reserves.

Second, the generating capacity available to supply energy at any point in time must always be greater than the demand for energy at that point in time as a result of the need to carry "inventory" in the form of generators providing frequency regulation and operating reserve services. That is, generating capacity (or in principle demand response) must be available that is either "spinning" or available to start up quickly to provide energy to balance supply and demand at each location on the network in response to real time variations in demand and unplanned equipment outages. When these operating reserves fall below a certain level (e.g. 7% of demand), system operators begin to take

[22] Sithe Energy presentation, IAEE, Boston Chapter, February 19, 2003.

actions to reduce demand administratively according to a pre-specified hierarchy of "operating reserve conservation" actions. The final actions in this hierarchy are voltage reductions and non-price rationing of demand (rolling blackouts).

Finally, there is an important but under-appreciated linkage between the retail procurement environment and the performance of the wholesale market with regard to investments in generating capacity that runs infrequently to meet peak demands for energy and operating reserves. If we are in a world where regulated monopoly distribution companies (i.e. no retail competition) are required to purchase electricity for their retail customers to meet their forecast demand subject to a clearly defined reliability criterion (e.g. specified loss of load probability and associated capacity margin above expected peak demand) then wholesale market prices (spot and forward) would rise, in one way or another, to clear the market at this reliability level. If, however, we are in a retail competition world, or in a world where distribution companies choose to have no predetermined reliability criterion but simply buy enough energy and operating reserves to meet demand at each point in time as it is realized, the implicit reliability and generating capacity level at which the wholesale market will clear will depend on exactly how consumer preferences for electricity with regard to price levels and price volatility are represented in the wholesale market, the associated ability of consumers to respond to real time prices and to match their preferences for market price risk with the costs of forward contracts.

A well functioning perfectly competitive wholesale electricity market will (to oversimplify for this discussion) operate in one of two states of nature. Under typical operating conditions (State 1), market clearing prices for energy <u>and</u> operating reserves should equal the marginal (opportunity) cost of the last increment of generating capacity that just clears supply and demand at each point in time.[23] In the case of wholesale electric energy supply, this price is the marginal cost of producing a little more or a little less energy from the generating unit on the margin in the merit order. See Figure 8 which depicts the sport market demand for electricity and the competitive supply curve for electricity under typical operating conditions (State 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 (state 1) the price of operating reserves will be very small --- close to zero, and far below the price of energy.

The second wholesale market state (state 2) is associated with a relatively small number of hours each year when there would be excess demand at a wholesale price that equals to the marginal supply cost of the last increment of generating capacity that can physically be made available on the network to supply energy or operating reserves. In this case, the market must be cleared "on the demand side." That is, consumers bidding to obtain energy would bid prices up to a (much) higher level reflecting the value (or value of lost energy or load) that consumers place on consuming less electricity as

---

[23] This will, of course, also be the value consumers place on this energy at the margin where supply and demand are equal.

demand is reduced to match the limited supplies available to the market. This second state is depicted in Figure 9. In what follows, I will refer to the conditions depicted in Figure 9 as <u>competitive</u> "scarcity" conditions.[24]

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 from generating capacity available to the network, 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."

These scarcity rents in turn help to cover their fixed capital and operating costs. 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 such 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.). The number of hours in which "scarcity" conditions emerge depends upon the amount of generating capacity that has been installed and is

[24] To distinguish it from contrived scarcity resulting from suppliers withholding supplies from the market to drive up prices.

physically available to operate relative to the tail of the distribution of aggregate demand realization during the year. This amount of generating capacity that is physically available to the network will then depend on investors in generating capacity balancing the costs of additional investments against the net revenues they expect to receive, including the "scarcity" rents produced when this state of nature emerges, from spot market sales and through forward contracts. The prices for such forward contracts are necessarily linked to future expected spot market prices and consumer and supplier preferences for market price risk.

If wholesale markets worked perfectly, prices during state 1 conditions and state 2 conditions would provide the appropriate price signals to link consumer preferences for reliability with the costs of supplying reliability with compatible investments in generating capacity. Moreover, with a set of liquid competitive forward markets for energy and for associated derivative instruments, consumer risk preferences and investor costs of bearing risk (and their cost of capital) could be matched through forward contracting and price hedging actions. Unfortunately, there are a number of market imperfections that undermine the operation and performance of the idealized (and necessarily over-simplified) competitive wholesale market that I have just described. The most important market and institutional imperfections are:

a. Consumer demand for energy and reliability are not well represented in wholesale spot markets today. 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.

c. Scarce generating capacity is not price-rationed during true scarcity conditions. Reliance by system operators 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 consumers over many non-reserve deficiency hours through some form of 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 supra-competitive prices and to incentives for over-investment in generating capacity.[25] The other market imperfections generally lead to under-investment in generating capacity and demand response programs. Whether, on balance, the incentives are for too much or to little generating capacity is an empirical question.

Whether it is too much or too little investment also depends 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: FERC, along with the market monitors in the existing ISOs have imposed 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.[26] 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 <u>competitive</u> "scarcity") are also the conditions when *market power* problems are likely to be most severe (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. They may also fail to mitigate fully supra-competitive prices during other hours. FERC's SMD NOPR reflects the judgment that, on balance, these 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

[25] See, for example, Green and David (1992), examine the inefficiencies associated with excessive entry stimulate by supra-competitive prices resulting from the exercise of market power.

[26] The FERC SMD 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, <u>id</u>. 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, <u>id</u>. at ¶ 433.

response capabilities available during very high demand contingencies to match consumer preferences for reliability, though it offers no empirical support for this conclusion.

b. Discretionary behavior by ISOs/RTOs/ITPs during true scarcity conditions: The level of prices for energy and operating reserves realized during 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, as I have already mentioned, 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 (Tirole 1988, Joskow 1987).

Empirical Evidence on Net Revenues During Scarcity Conditions

Exactly how these market and institutional imperfections balance out and affect investment incentives is ultimately an empirical question. 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." I then apply 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 the ISO's spot markets has been exhausted.

I focus on the marginal generating capacity that supply energy and operating reserves only during "scarcity conditions" and, as a result, have very low capacity factors. Generating units that are expected to operate at very low capacity factors typically have relatively low fixed costs and relatively high marginal operating 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 of the last (highest operating cost) generating unit in the merit order physically capable of providing energy or operating reserves. Let P_s be the average 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 that runs for H_s hours to just break even given any particular probability of scarcity ($H_s/8760$) is then given by:[27]

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

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 average hourly opportunity cost that the marginal consumer

[27] This is obviously an oversimplification. P_s will vary during scarcity hours.

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.[28] 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:

$$C_K = H_s{*}V \tag{2}$$

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).

I now 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.[29] I compare these scarcity rents to alternative measures of C_K. The analysis shows that R

[28] 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 Stoft 2002, Chapter 2-5.

[29] 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.

has been much lower than C_K (defined in a number of different ways) in New England over the last four years.

Many 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.[30] 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.[31] This in turn typically triggers the System Operator'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.

First, I identified all hours when the New England ISO declared an operating reserve deficiency. Operating reserve deficiencies trigger NEPOOL Operating Procedure 4.[32] NEPOOL Operating Procedure 4 (Op-4) has 16 action steps of increasing severity.

---

[30] This discussion assumes that 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 <u>overestimate</u> 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.

[31] 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. If system operators had the right financial incentives if would make sense to expand their contracting options in this way.

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

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 physically capable of supplying energy and operating reserves.

During these scarcity conditions, marginal generators selling in the real time energy and ancillary services markets can earn revenues in one of two ways.[33] 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 equal to the price of energy <u>minus</u> 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

[33] Price formation in the real time markets will work its way back into day-ahead and forward contract price formation through intertemporal arbitrage.

energy exceeded $1000, I set it to the $1000 price cap that was implemented after May 2000.[34] 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 higher 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.[35] To calculate the "scarcity rents" associated with supplying energy during Op-4 conditions, I assumed that the short-run marginal cost 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 10. 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

[34] 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.

[35] The revenue effect is relatively small.

theory suggests they should be).[36] There is significant volatility from year to year in the rents earned, however. On average there are 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. <u>On average there were only six hours per year when the price cap was binding, again with considerable year-to-year variation</u>.[37] This suggests that the $1000 price caps are unlikely to be the primary source of the revenue deficiencies (more on this below).[38] 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 - $80,000/Mw-Yr in New England. Clearly, the scarcity rents are far below what would be necessary to attract CT to the market to be available to supply operating reserves and energy only during scarcity conditions.

---

[36] 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.

[37] One must wonder if 1999 is just an unusual year, with the ISO and market participants learning how to operate within the new wholesale market institutions in New England. 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.

[38] A similar analysis has been done for the New York ISO and comes to similar conclusions. See Patton (2002), pp. 25, 42-64. Patton's analysis of the New York ISO 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.

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 case, 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 that the annual fixed O&Ms of an older fossil steam units is in the range of \$20,000 to \$35,000/Mw/Yr.[39] The scarcity rents that I have measured for New England are not high

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

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.[40] 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.[41] We would expect the implied value of V as defined here to be below the 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

---

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

[41] 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.

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.

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 conclusions in the FERC SMD 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. [42, 43] While there appears to be abundant generating capacity in most regions of the country at the present time, this is a potential problem that

[42] 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.

[43] Of course, New England, New York and PJM have had capacity obligations and owners of generating capacity can receive an additional stream of revenues from sales of capacity. And there is no shortage of generating capacity in New England, PJM or New York (except in New York City where investment in new generating capacity faces additional challenges) and in most other regions of the country. 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.

may undermine future investment and lead to premature retirement of some existing generating capacity.

TRANSMISSION GOVERNANCE AND INVESTMENT

I do not want to conclude this paper without at least a few comments about transmission governance and investment issues. Transmission networks provide the essential supporting platform upon which competitive wholesale markets depend. Transmission congestion effectively reduces the geographic expanse of competition, increases the incidence of locational market power, and can limit entry of competing generators. A well functioning transmission network is a critical component of a program to create robust competitive wholesale and retail markets for electricity. Yet the legacy transmission network that we inherit from the era of large numbers of vertically integrated regulated firms was not designed to promote competition among generators over large geographic areas, focused on interconnecting generators and loads within individual utility control areas and did not take local market power and other market performance problems into account when investments were made. It should come as no surprise that the legacy network is not well suited for supporting competitive wholesale markets and that significant investments will be required to adapt the legacy network to its new role.

As I have already discussed, many countries that have implemented competitive electricity market programs have created independent regional or national transmission companies with responsibilities for system operations (broadly defined) and transmission investment. In the U.S., we have taken a different approach. Most transmission assets are owned, operated and maintained by vertically integrated utilities. There is very significant geographic balkanization of ownership and operation with many companies owning, operating, maintaining and potentially responsible for investing in new transmission facilities at particular locations on the same AC network.

Rather than promoting vertical and horizontal restructuring of asset ownership and operation to deal with these independence and balkanization issues, FERC has taken the existing ownership structure as a constraint and promoted the creation of new not-for-profit independent system operators (ISO) or Regional Transmission Operator (RTO) to deal with these issues. RTOs are to be responsible for scheduling and dispatching generators on regional networks, implementing market-based mechanisms for allocating scarce transmission capacity, monitoring generator, marketer, transmission owner behavior and market performance, coordinating maintenance performed by transmission owners, coordinating regional planning processes for new transmission facilities, and operating voluntary public spot markets (real time and day-ahead) for energy and ancillary services. However, these independent entities own no transmission assets, have no linemen or helicopters to maintain transmission lines and respond to outages, and are not directly responsible for the costs of operating, investing in, or the ultimate performance of the transmission networks they "manage."

These organizational arrangements are further complicated by the distribution of regulatory authority and responsibilities between state and federal regulators. The states are responsible for reviewing applications for major new transmission facilities and granting any necessary permits. FERC has been responsible for regulating the prices for unbundled (wholesale) transmission service, but has no authority over transmission planning or siting approvals, while the states are responsible for regulating charges for bundled transmission service charged by vertically integrated companies to their retail customers, though exactly the same facilities and people are involved in both. The FERC SMD endeavors to extend FERC's jurisdiction over charges for all transmission service, but whether states can be forced to pass along these FERC approved charges in retail prices is a matter of dispute.[44] Until recently, FERC's policies on transmission investment responsibilities and cost recovery rules have been confusing and, in my view much too heavily focused on flawed models that rely primarily on merchant transmission investment (Joskow and Tirole 2003). Moreover, FERC has not embraced any program to adopt performance-based regulatory mechanisms focused on improving transmission network performance and reducing costs. Very recently, FERC has begun to use its ratemaking authority to reward transmission owners that divest their transmission assets, that form independent transmission companies, that operate under the supervision of RTOs, and to improve network performance.[45] This is a step in the right direction, and since these policies were adopted, there has been growing interest in transmission divestiture and the formation in independent transmission companies (ITC).[46] However, there is much work to be done.

Current institutional arrangements governing transmission operations and investment are simply not well matched to creating the transmission network platform

---

[44] It seems fairly clear that FERC does have this authority if it chooses to assert it. However, in a White Paper issued on April 28, 2003 FERC has announced that it will not assert this jurisdiction and allow the states to determine prices for "bundled" retail transmission service.

[45] *Proposed Pricing Policy for Efficient Operation and Expansion of the Transmission Grid,* Federal Energy Regulatory Commission Docket Number PL-03-1-000, Issued January 13, 2003.

[46] My views on FERC's transmission regulatory policies can be found in my Comments submitted in response to FERC's SMD NOPR. http://econ-www.mit.edu/faculty/pjoskow/papers.htm.

necessary to support well functioning competitive markets that can operate with a minimal amount of continuing regulatory “mitigation” measures to respond to market power and related market performance concerns. Stimulating transmission investment has been especially problematic. While generation investment in generating capacity grew enormously during the last five years, transmission investment has been declining for many years (U.S. Department of Energy 2002). Transmission congestion has grown steadily over the last several years. Figure 10 displays the moving average of the number of monthly administrative Transmission Line Reliefs (TLRs) ordered by regional security coordinators to respond to transmission constraints, primarily in the Midwest. Between June 1998 and December 2002, the number of monthly TLRs has increased by a factor of 6. In New England, new generating capacity that has been built in Maine and Rhode Island cannot be run to meet regional energy needs economically because there is inadequate transmission to get it out of these areas. At the same time, Southern Connecticut faces reliability problems because of transmission constraints that significantly limit imports into that part of New England. These trends in transmission congestion are not limited to areas that have not implemented Locational Marginal Pricing (LMP). Table 11 displays the number of hours of transmission congestion on the PJM system from 1998 to 2002, the first ISO to implement an LMP system.[47] Table 12 displays congestion charges in PJM between 1999 and 2002. By both measures congestion has continued to grow rapidly in PJM.

It is often argued that the primary problem causing the decline in transmission investment is local “Nimby” opposition to new transmission lines. This is certainly a

[47] The 2002 data include the incorporation of PJM-West into PJM and are not directly comparable to the data for previous years.

problem. However, many transmission investment opportunities do not involve opening up major new transmission corridors or significantly expanding the footprint of the transmission network. 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, 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 characterization of the diverse attributes of the full range of transmission investment opportunities is consistent with the 33 priority projects identified in ISO New England's recently released 2002 Regional Transmission Expansion Plan.[48] 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

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

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.

FERC has also adopted an excessively narrow conceptualization of "congestion management" that has led to its failure properly to consider important organizational, management and incentive issues affecting both transmission operation and transmission investment. 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 transmission owners, and investments, small and large, in the transmission network. An ISO or RTO like PJM running a security-constrained dispatch, resulting in a set of ex post LMPs is not in a position to undertake this broader set of transmission network 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 ISO 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 recent FERC rulemakings, LMPs themselves do not "manage congestion" in any meaningful way. The ISO'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 need to be taken into account by policy makers. The experience in England and Wales over the last decade indicates that substantial transmission investment, operating performance improvements, reductions in congestion and ancillary services costs, and lower overall transmission costs can be achieved with the right organizational and regulatory incentive institutions in place. This is a goal that electricity policymakers in the U.S. need to place higher on the restructuring and regulatory reform agenda.

CONCLUSION

The development of well-functioning competitive wholesale and retail markets for electricity in the U.S. is a work in progress. It has encountered more problems and proceeded less quickly than some had anticipated when the first restructuring and competition programs were first being implemented in the late 1990s. The most visible success of these initiatives to date has been the substantial investment in new generating capacity completed by merchant generating companies in the last few years, and the shifting of the associated construction cost, operating performance and market risks to suppliers rather than to consumers as under regulation. There has also been substantial growth in the fraction of electricity supplied through competitive wholesale market transactions (physical and financial) as a result of restructuring of incumbent vertically integrated utilities, the entry of new generating capacity built by merchant generating companies, and the development of public and private wholesale power trading institutions. Load-serving entities in all parts of the country rely more on competitive wholesale market purchases to supplement or replace owned-generation subject to cost-based regulation than they did in the mid-1990s. Despite their imperfections, Order 888 and FERC initiatives that have built on it have substantially increased access to transmission networks and associated support services, facilitating these wholesale market developments. We have learned a lot about the performance attributes of alternative wholesale market institutions from recent experience. Reforms aimed at adopting best practices are being implemented in many regions of the country with strong support and encouragement by FERC. Retail customers in a number of states have benefited from lower regulated retail prices negotiated as a component of state restructuring programs, though the direct benefits attributable to retail competition per se have been limited and have flowed primarily to the largest electricity customers.

The transition to competitive wholesale and retail electricity markets has also been plagued by problems and disappointments. The boom in merchant generating investment and the growth in wholesale power trade has now turned into a bust. Many merchant generating and trading companies are in serious financial trouble and cannot raise capital to build or acquire projects. Many generating projects under construction and development have been cancelled. While a significant reduction in investment in new generating capacity should have been expected due to basic supply and demand fundamentals, the financial problems faced by this sector reflect a number of other market and regulatory phenomena. It is fairly clear that the next generation of merchant investment activity will take place in an environment where capital is much more costly than it was during the late 1990s bubble and where lenders will be looking for supporting portfolios of supply contracts of longer durations. Investors will also be looking for much more stability in wholesale market rules, reforms in wholesale market institutions to support investment or continued operation of peaking capacity, more efficient and transparent congestion management arrangements, and stable federal regulatory policies. Many wholesale trading companies have either completely or substantially withdrawn from trading activity beyond trading around their own generating assets and market liquidity has fallen dramatically, reducing the efficiency of short-term markets and making it very difficult for buyers' and sellers' risk preferences to be well-matched in liquid markets for longer-term forward markets. Market power problems and other market imperfections have reduced the efficiency of wholesale power markets and increased costs to consumers.

In California, the combination of market design imperfections, market power problems, and poor federal and state policy responses has managed both to increase retail prices enormously (30% to 40%) and to leave the utility and merchant generating sectors littered with financially crippled and bankrupt suppliers. The performance of retail competition programs has been disappointing almost everywhere, especially for residential and small commercial customers. Imperfections in retail competition programs investments in new generating capacity and undermine the performance of wholesale spot markets as well.

Investment in transmission capacity has stagnated while network congestion has increased. This in turn has increased local market power problems and complicated the smooth operation of wholesale power markets. Many parts of the country continue to rely on inefficient non-price rationing mechanisms to manage congestion and property rights to scarce transmission capacity continue to be poorly defined. The transmission system remains fragmented with too many system operators relying on incompatible scheduling, transmission pricing, and emergency management mechanisms. The bulk of the transmission capacity is owned by companies that also own and operate generating facilities connected to these transmission facilities and trade power in the same regional markets, creating opportunities to increase rivals' costs and reduce competition. FERC has responded to the failure to restructure the industry to match the needs of competitive electricity markets with new institutional arrangements (ISO, RTO) that may be well suited for operating public markets for energy and ancillary services and real time physical system dispatch, but whose long run performance attributes from the broader perspective of transmission network operating costs, transmission line availability, and transmission investment are not particularly promising.

The positions that the various states have taken regarding electricity sector reforms in general, and FERC's SMD in particular, reflect their assessment of the costs and benefits of the electricity competition and restructuring initiatives to date. The states in the Northeast, Texas, and a few states in the Midwest that have gone the farthest down the restructuring path are committed to making these initiatives work better and to implement the wholesale and retail market reforms necessary to do so. They have gone too far to easily reverse course and return to the old system of regulated vertically integrated firms. Their strategies and the basic framework established in the FERC SMD seem to be reasonably well aligned. California remains in shock from the experiences of 2000 and 2001 and its long run electricity strategy remains murky at best. The majority of the states, clustered in the Southeast, the South, and the West have either taken a cautious wait and see attitude or have simply rejected restructuring and competition initiatives. These states tend to have relatively low regulated retail prices, do not face looming supply shortages or reliability problems and face little consumer pressure for change. Why take the risk that a California-like crisis will come home to roost? Since restructuring to rely on wholesale and retail electricity markets involves turning much more of the electricity value chain over to federal from state jurisdiction, these states are also concerned that FERC will not act promptly or responsibly to protect consumers in their states when problems arise. (Texas (or at least ERCOT) is the only state that has been in a position to implement fundamental reforms while retaining state jurisdiction of the reform process.)

If the states that have not embraced competitive market reforms move forward voluntarily in the future it will be because the other states that have committed to restructuring, wholesale and retail competition can demonstrate to them that it has in practice, rather than just in theory, brought long-term benefits to consumers. This will require solid empirical analysis of the performance of the electric power sectors in those states that have restructured and implemented comprehensive wholesale and retail competition programs.

REFERENCES

Borenstein, S., J. Bushnell and F. Wolak, “Measuring Market Inefficiencies in California’s Restructured Wholesale Electricity Market,” *American Economic Review*, December 2002, 92(5): 1376-1405.

Department of Trade and Industry (DTI), *UK Energy Sector Indicators*, London, 2000.

Department of Trade and Industry (DTI), *Energy White Paper: Our Future – Creating a Low Carbon Economy*, Her Majesty’s Stationary Office, February 2003.

Federal Energy Regulatory Commission, *Notice of Proposed Rulemaking, Docket No. RM01-12-000,* July 31, 2003. [FERC SMD NOPR]

Giulietti, M., C. Waddams Price, and M. Waterson, “Consumer Choice and Industrial Policy: A Study of UK Energy Markets,” mimeo March 2003.

Green, R. “Can Competition Replace Regulation for Small Utility Customers?”, Centre for Economic Policy Research, Discussion Paper 2405, 2000.

Green R. and D. Newbery, "Competition in the British Electricity Spot Market, *Journal of Political Economy*, 100(5), 929-53, 1992

Hogan, W. (1992) “Contract Networks for Electric Power Transmission,” *Journal of Regulatory Economics*, 4: 211–242.

Hogan, W., “Markets in Real Networks Require Reactive Prices,” *Energy Jou*rnal*,* 14(3), 1993, 171-200.

Holmström B. and J. Tirole (2000) "Liquidity and Risk Management,” *Journal of Money, Credit and Banking*, 32 (3): 295–319.

Joskow, P. L., "Inflation and Environmental Concern: Change in the Process of Public Utility Price Regulation", *Journal of Law and Economics*, October, 1974**.**

Joskow, P.L., “Contract Duration and Relationship Specific Investment, *American Economic Review,* pp. 168-185, 1987

Joskow, P. L., "Regulatory Failure, Regulatory Reform and Structural Change In The Electric Power Industry," *Brookings Papers on Economic Activity: Microeconomics,* 1989.

Joskow, P. L. "Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector", *Journal of Economic Perspectives*, 11(3), Summer 1997, pp. 119-138.

Joskow, P. L. "Deregulation and Regulatory Reform in the U.S. Electric Power Sector," in *Deregulation of Network Industries: The Next Steps* (S. Peltzman and Clifford Winston, eds.), Brookings Press, 2000.

Joskow, P.L. , "Why do We Need Electricity Retailers? You Can Get it Cheaper Wholesale," February 13, 2000. (2000b) http://econ-www.mit.edu/faculty/pjoskow/files/RETAILV2.pdf.

Joskow, P. L., "California's Electricity Crisis," *Oxford Review of Economic Policy*, Vol. 17, No. 3, pp. 365-388, 2001.

Joskow, P. L., "Electricity Sector Restructuring and Competition: A Transaction Cost Perspective," *The Economics of Contracts: Theories and Applications*, Eric Brousseau and Jean-Michel Glachant (eds)., Cambridge University Press, 2002.

Joskow, P. L. and E. 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.

Joskow, P. L. and R. Schmalensee, *Markets for Power*, MIT Press, 1983

Joskow, P.L., "Restructuring to Promote Competition in Electricity: In General and Regarding the POOLCO vs. Bilateral Contracts Debate," presented at the American Economic Association annual meeting in San Francisco, January 1996.

Joskow, P. L. and J. Tirole, "Transmission Rights and Market Power on Electric Power Networks," *Rand Journal of Economics*, 31(3) Autumn, 2000.

Joskow, P. L. and J. Tirole, "Merchant Transmission Investment," MIT Center for Energy and Environmental Policy Research Working Paper, February 2003.

Littlechild, S. C., "Wholesale Spot Market Passthrough," *Journal of Regulatory Economics*, 23(1), January 2003, pp. 61-91.

Office of Gas and Electricity Markets (OFGEM), "Review of Domestic Gas and Electricity Competition and Supply Regulation," November 2001.

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.

Patton, D.S., *Summer 2002 Review of New York Electricity Markets*, New York Independent System Operator, October 15, 2002

Public Utility Commission of Texas, *February 2003 Report Card on Competition*.

Stoft, S., *Power System Economics*, Wiley Interscience, 2002.

Tirole, Jean, *The Theory of Industrial Organization*, MIT Press, 1988.

U.S. Department of Energy, *National Transmission Grid Study*, Washington, D.C., May 2002.

White, M., "Power Struggles: Explaining Deregulatory Reforms in Electric PowerMarkets," *Brookings Papers on Economic Activity: Microeconomics, 1996, 201-250.*

Wilson, R. (2002) "Architecture of Power Markets," *Econometrica*, 70(4): 1299-1340.

Wolfram, C., "Measuring Duopoly Power in the British Electricity Spot Market," *American Economic Review*, 89(4), pp. 805-826, 1999.

TABLE 1

GENERATING PLANT SALES AND TRANSFERS
MW

YEAR	DIVESTED	TO UNREGULATED AFFILIATES	TOTAL
1998	23,413	-0-	23,413
1999	50,962	4,108	55,070
2000	15,334	32,657	47,991
2001	8,135	20,051	28,186
2002	2,154	27,206	29,360

Source: *Electric Power Monthly,* various issues 1998-2003, U.S. Energy Information Administration

TABLE 2

GENERATING CAPACITY ADDITIONS
MW

YEAR	**GENERATING CAPACITY ADDED**
1997[1]	4,000
1998[2]	6,500
1999[3]	10,500
2000[4]	23,500
2001[5]	48,000
2002[6]	55,000

[1] *Inventory of Power Plants*:*1997*, Energy Information Administration
[2] *Inventory of Power Plants:1998*, Energy Information Administration
[3] *Electric Power Annual:1999*, Energy Information Administration
[4] *Electric Power Annual:2000*, Energy Information Administration
[5] Energy Information Administration website, read February 15, 2003
[6] Energy Argus New Power Plant Spreadsheet, February, 2003

TABLE 3

NON-UTILITY POWER PRODUCTION
% Of Total Electricity Supplied

YEAR	% Non-Utility
1990	7.2%
1991	8.0%
1992	9.3%
1993	9.8%
1994	10.5%
1995	10.8%
1996	10.7%
1997	10.6%
1998	11.2%
1999	14.3%
2000	20.7%
2001	30.0%
2002	33.6%

Source: *Monthly Energy Review*, U.S. Energy Information Administration, March 2003, page 97

TABLE 4

SHARE PRICES OF SELECTED MERCHANT GENERATING AND TRADING FIRMS
($/share)

Company	May 2001 Peak Week	March 10, 2003	
		Share Price	S&P Credit Rating
AES	48.5	3.2	B+
AEP	50.4	21.2	BBB
Allegheny	53.8	5.1	BB-
Calpine	54.7	6.1	BB
Dynegy	57.0	2.2	B
Duke	46.1	12.2	BBB[7]
El Paso	64.9	4.4	B+
Mirant	45.4	1.4	BB
NRG	30.4	6.2	D
Reliant	33.8	3.7	B-
Williams	41.0	4.0	B+

Source: Yahoo Finance and S&P web site, accessed March 10, 2003

[7] Duke Energy Trading and Marketing

TABLE 5

RETAIL CHOICE IN MASSACHUSETTS
December 2002

Retail Choice Began: March 1998

Fraction of Customers Served By ESP:

Residential:	2.6%
Small C/I:	8.8%
Medium C/I:	11.1%
Large C/I:	29.2%
Total:	3.4%

Fraction of Load Served By ESPs:

Residential:	2.4%
Small C/I:	11.5%
Medium C/I:	17.3%
Large C/I:	46.2%
Total:	22.2%

Source: http://www.state.ma.us/doer/pub_info/migrate.htm

TABLE 6

RETAIL CHOICE IN NEW YORK
As of December 2002

Retail Choice Began: May 1998-July 2001

Fraction of Customers Served by ESP:

Residential:	**5.0%**
Non-Residential:	**7.1%**

Fraction of Load Served by ESP:

Residential:	**5.6%**
Non-residential:	**32.3%**

Source: http://www.dps.state.ny.us/Electric_RA_Migration.htm

TABLE 7

RETAIL CHOICE IN MAINE
(Central Maine Power)

January 2003

Retail Choice Began: March, 2000

Fraction of Load Served by ESP:

Residential/Small Commercial:	**<1%**
Medium Industrial:	**28%**
Large Industrial:	**73%**
Total:	**33%**

Source:
http://www.state.me.us/mpuc/electric%20restructuring/migrationrates.htm

TABLE 8

RETAIL CHOICE IN NEW JERSEY
As of December 15, 2002

Retail Choice Began: November 1999

Fraction of Customers Served By ESPs:

Residential:	**0.06%**
Non-Residential:	**0.14%**

Fraction of Retail Load Served by ESPs: 1.9%

Source: http://www.bpu.state.nj.us/wwwroot/energy/elecswitchdata.htm

TABLE 9

RETAIL CHOICE IN OHIO
September 30, 2002

Retail Choice Began: January 1, 2001

Fraction of Customers Served by ESP:

<u>AEP Subsidiaries:</u>

Columbus Southern:

Residential:	**0.00%**
Commercial:	**0.25%**
Industrial:	**0.00%**

Ohio Power:

Residential:	**0.00%**
Commercial:	**0.00%**
Industrial:	**0.00%**

<u>First Energy Subsidiaries:</u>

Cleveland Electric:

Residential:	**55.05%**
Commercial:	**54.84%**
Industrial:	**27.69%**

Ohio Edison:

Residential:	**23.85%**
Commercial:	**23.51**
Industrial:	**27.70%**

<u>Dayton Power and Light:</u>

Residential:	**0.00%**
Commercial:	**0.01%**
Industrial:	**0.16%**

Source:
**http://www.puco.ohio.gov/ohioutil/MarketMonitoring/ECC_Switch_Rates_Summar
y**

TABLE 10

SCARCITY RENTS IN ISO-NEW ENGLAND[8]

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[9]	$ 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

[8] Computation procedures are discussed in the text.

[9] 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.

TABLE 11

PJM CONGESTION EVENT HOURS

YEAR	TOTAL	500KV	345KV	230KV
1998	1,244	203	71	588
1999	2,134	189	148	818
2000	6,941	562	14	869
2001	8,435	759	38	744
2002	11,657	1,926	1,107	2,056

Source: *PJM State of the Market Report 2002.*

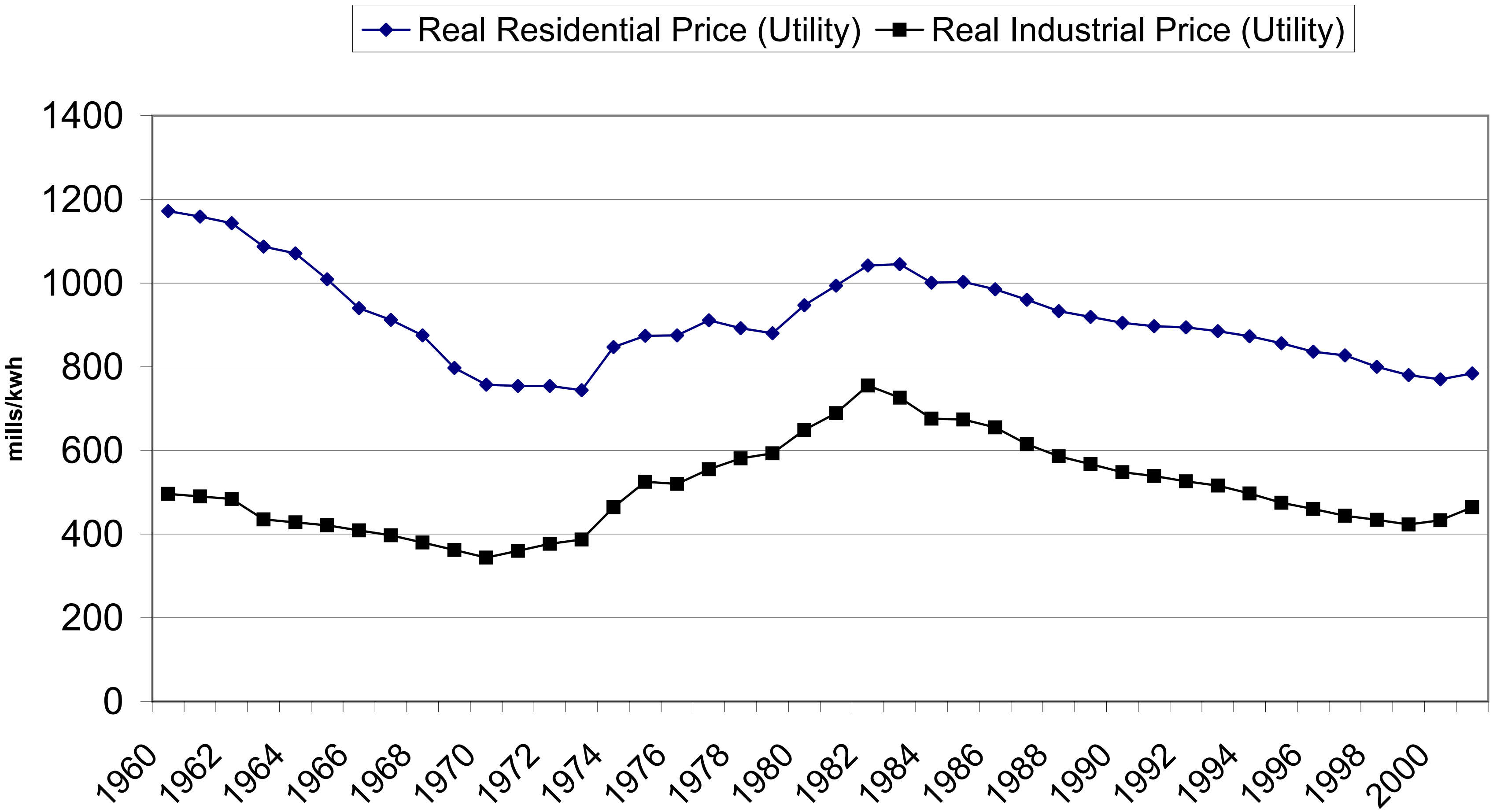
TABLE 12

PJM CONGESTION CHARGES
($million)

1999	**$53**
2000	**$132**
2001	**$271**
2002	**$430**

Source: *PJM State of the Market Report 2002*

FIGURE 1
Real Retail Price of Electricity ($1996):1960-2001



Source: *Annual Review of Energy* , U.S. Energy Information Administration

Figure 2
COMPREHENSIVE REFORM VISION
COMPETITIVE WHOLESALE + RETAIL MARKETS

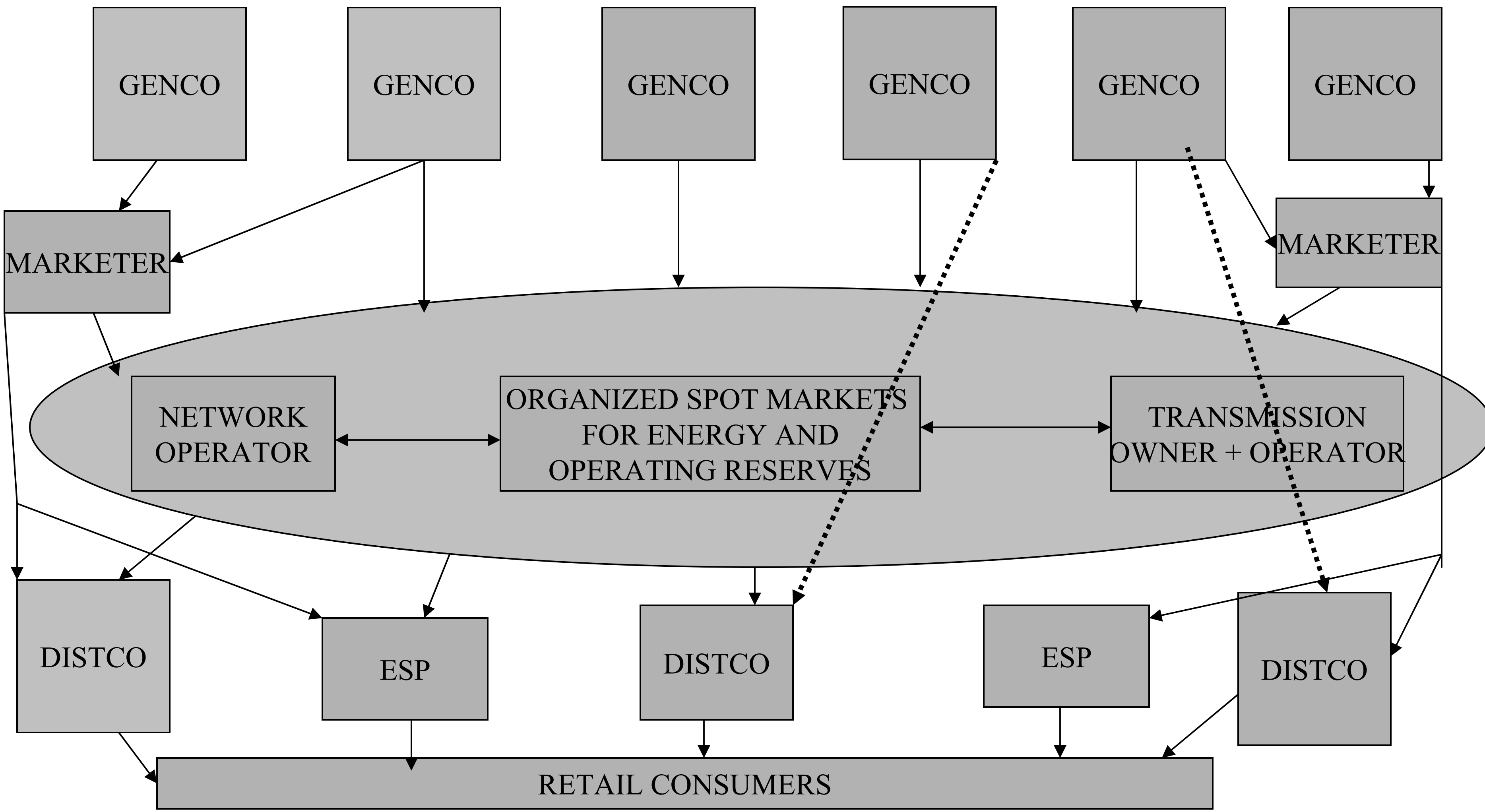
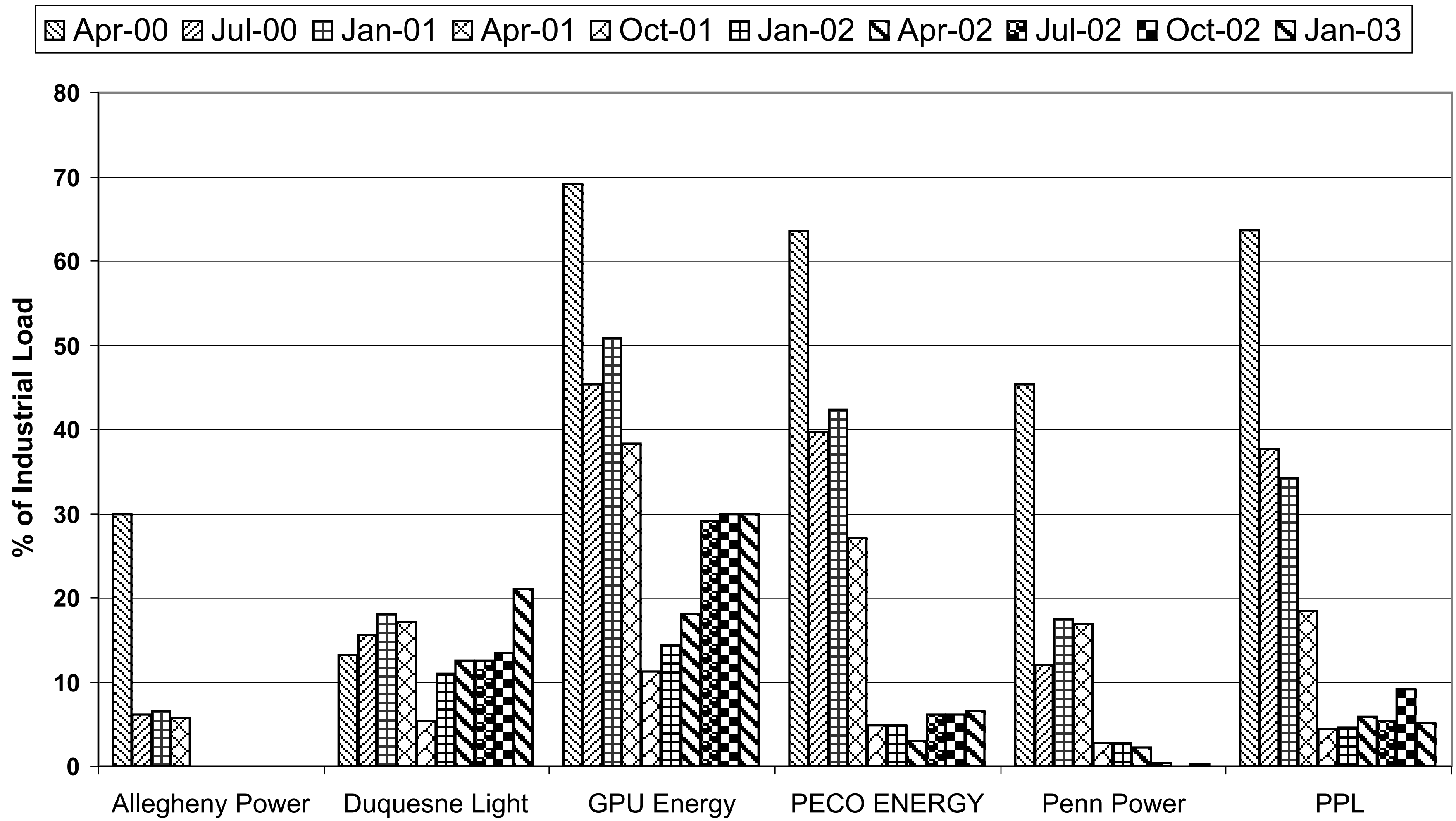
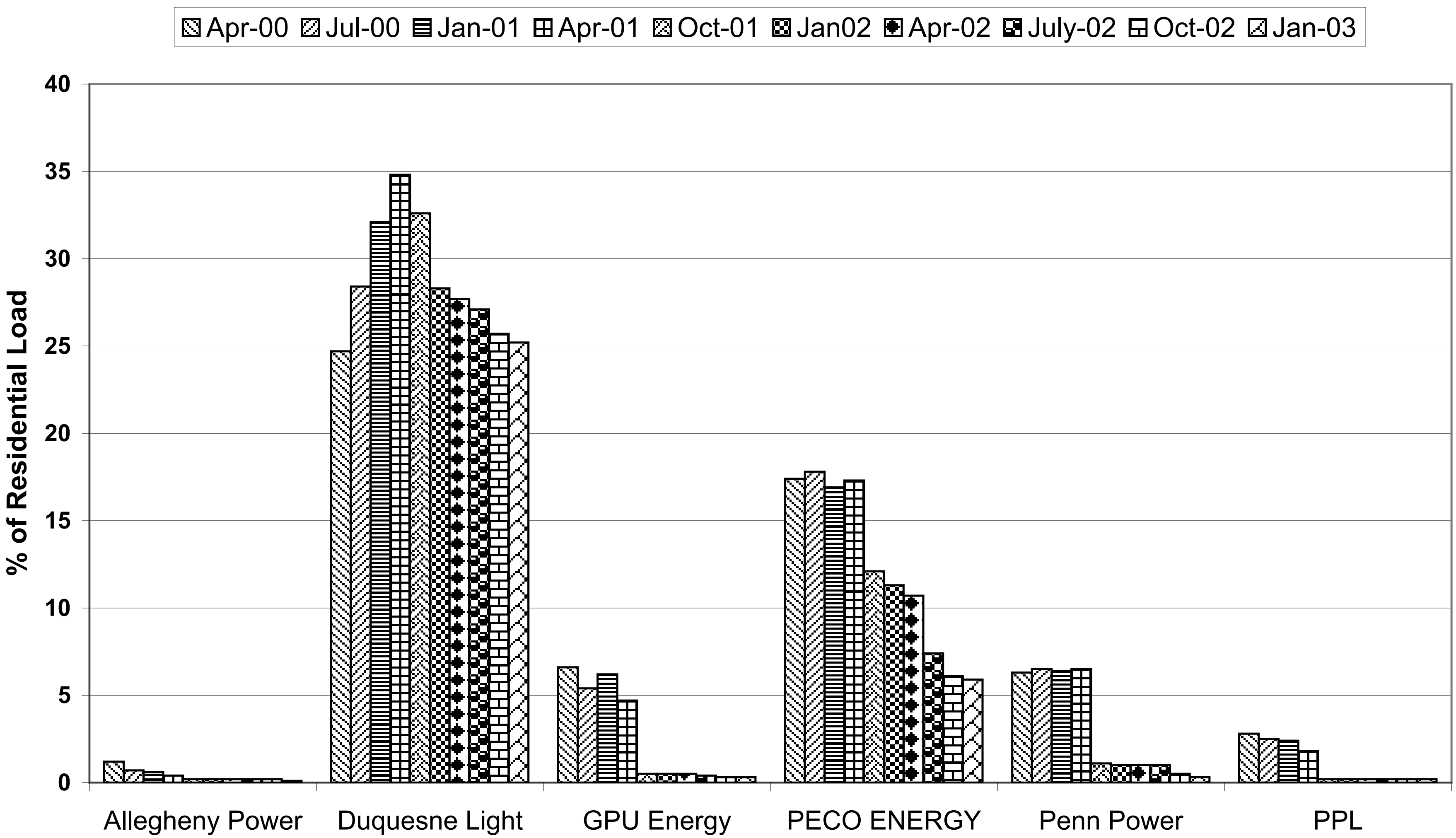


Figure 3A
PENNSYLVANIA DIRECT ACCESS LOAD: INDUSTRIAL (%)



Source: Pennsylvania Office of Consumer Advocate

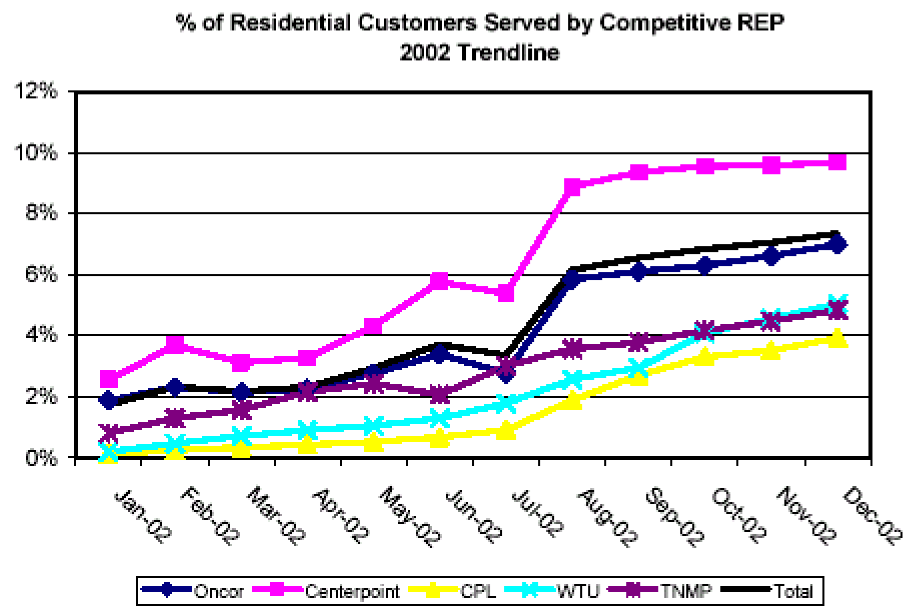
Figure 3B
PENNSYLVANIA DIRECT ACCESS LOAD: RESIDENTIAL (%)

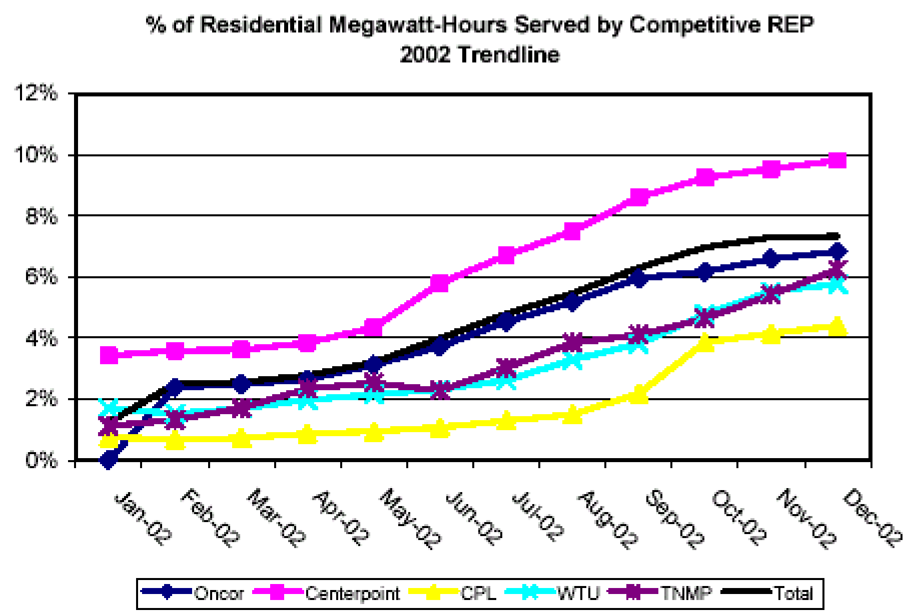


Source: Pennsylvania Office of Consumer Advocate

FIGURE 4A
RETAIL CHOICE IN TEXAS
December 31, 2002

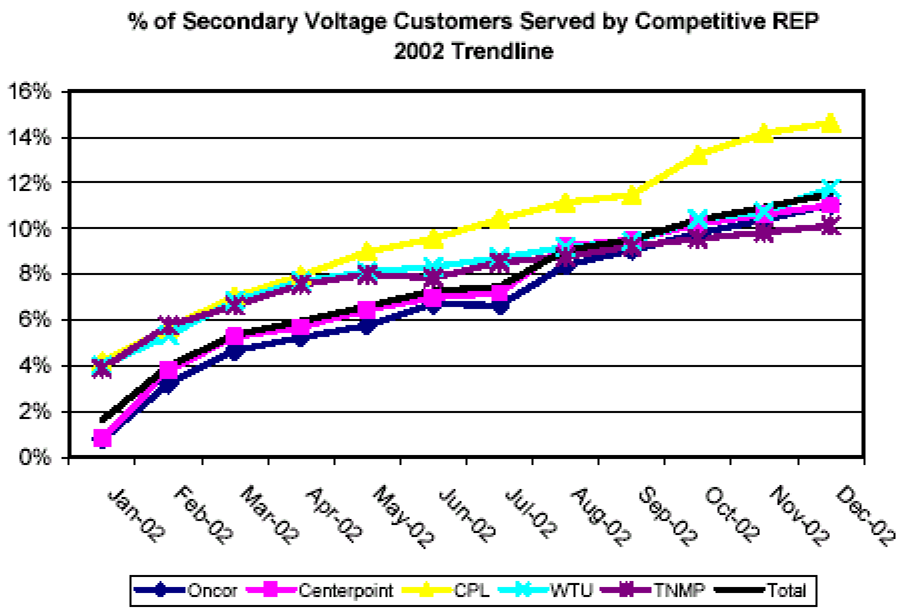
Retail Choice Began: January 1, 2002
(Data include customers in 2001 Pilots)

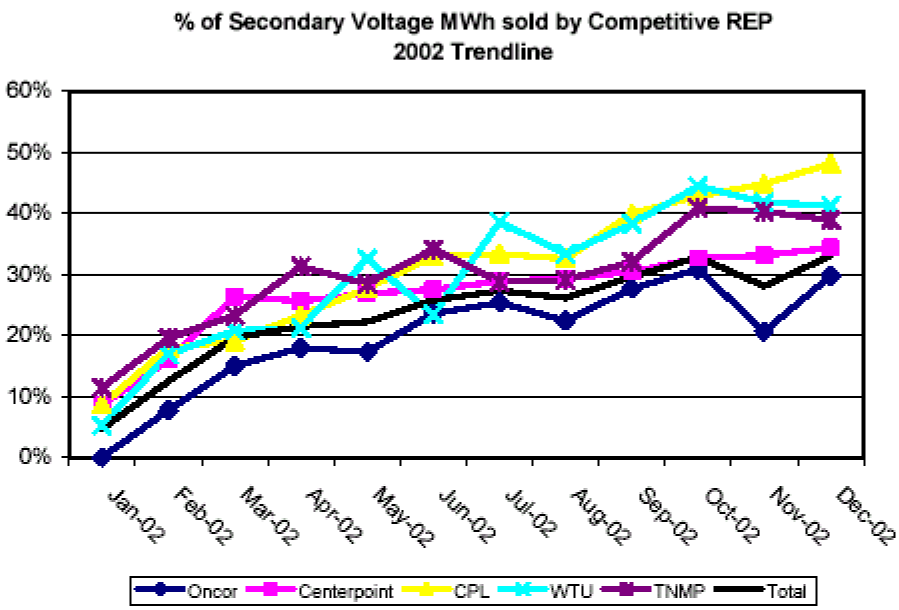




Source: Public Utility Commission of Texas, *February 2003 Report Card on Competition*

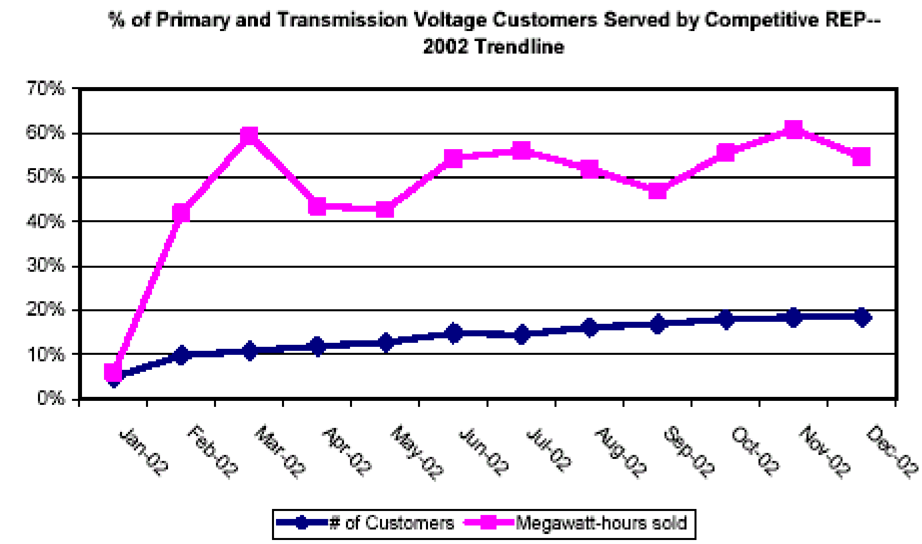
FIGURE 4B
RETAIL CHOICE IN TEXAS
December 31, 2002

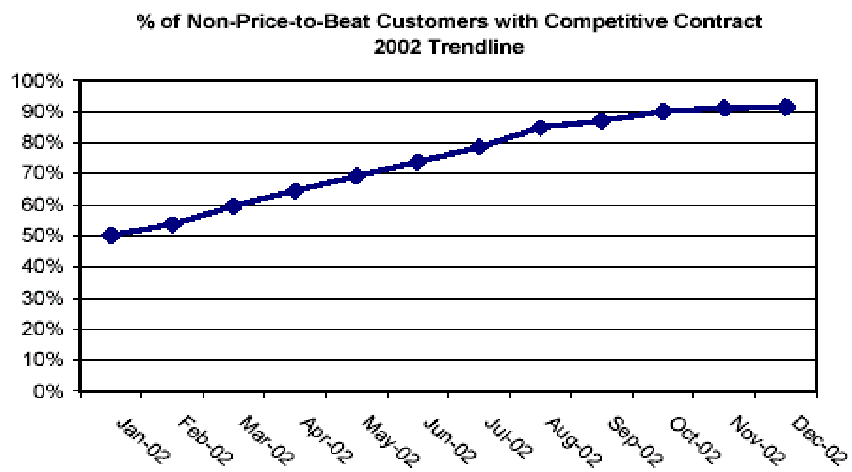




Source: Public Utility Commission of Texas, *February 2003 Report Card on Competition*

FIGURE 4C
RETAIL CHOICE IN TEXAS
December 31, 2002





Source: Public Utility Commission of Texas, *February 2003 Report Card on Competition*

FIGURE 5

PUBLIC UTILITY COMMISSION OF TEXAS
RETAIL MARKET OVERSIGHT SECTION - ELECTRIC DIVISION
RETAIL ELECTRIC SERVICE RATE COMPARISONS

DECEMBER 2002 RATE COMPARISON
AVERAGE ANNUAL RATE (CENTS PER KWH)

TDU Service Area	Retail Electric Provider	500 kWh	% savings	1000 kWh	% savings	1500 kWh	% savings	2000 kWh	% savings
		Average monthly usage and percentage savings off of PTB							
TXU	TXU ENERGY SERVICES - Price to Beat	9.41		8.66		8.30		8.12	
	ACN ENERGY	8.57	9%	8.08	7%	7.91	5%	7.83	4%
	CIRRO CORP	9.26	2%	8.26	5%	7.61	8%	7.29	10%
	ENERGY AMERICA *36 Month Term*	8.98	5%	8.48	2%	8.31	0%	8.23	-1%
	ENTERGY SOLUTIONS	8.64	8%	8.15	6%	7.98	4%	7.90	3%
	FIRST CHOICE POWER	8.69	8%	8.20	5%	8.03	3%	7.95	2%
	GEXA ENERGY	8.30	12%	8.00	8%	7.90	5%	7.85	3%
	GREEN MOUNTAIN ENERGY (100% renewable power)								
	Month to Month	9.49	-1%	9.00	-4%	8.83	-6%	8.75	-8%
	Reliable	9.79	-4%	9.30	-7%	9.13	-10%	9.05	-11%
	RELIANT ENERGY	8.94	5%	8.23	5%	7.88	5%	7.71	5%
	UTILITY CHOICE ELECTRIC	9.19	2%	8.20	5%	7.86	5%	7.70	5%

This information is compiled and provided by the Public Utility Commission of Texas from publicly available information from the Retail Electric Providers and PUC approved price to beat rates. The average price shown on this sheet was compiled from the Electricity Facts Labels of the REPs and is inclusive of all fixed and variable charges. The actual cost per kWh to a customer may vary based on the actual usage of the customer. The PUC makes no recommendations with respect to any REP. Although we believe that these prices are accurate, the PUC makes no warranty that the prices in this table are currently being offered. Please contact the relevant REP for their current pricing offers and terms of service. Information on how to select a REP and contact information for REPs is located at **www.powertochoose.org**.

Source: Public Utility Commission of Texas

Figure 6
AVERAGE RESIDENTIAL PRICE (cents/Kwh Nominal)

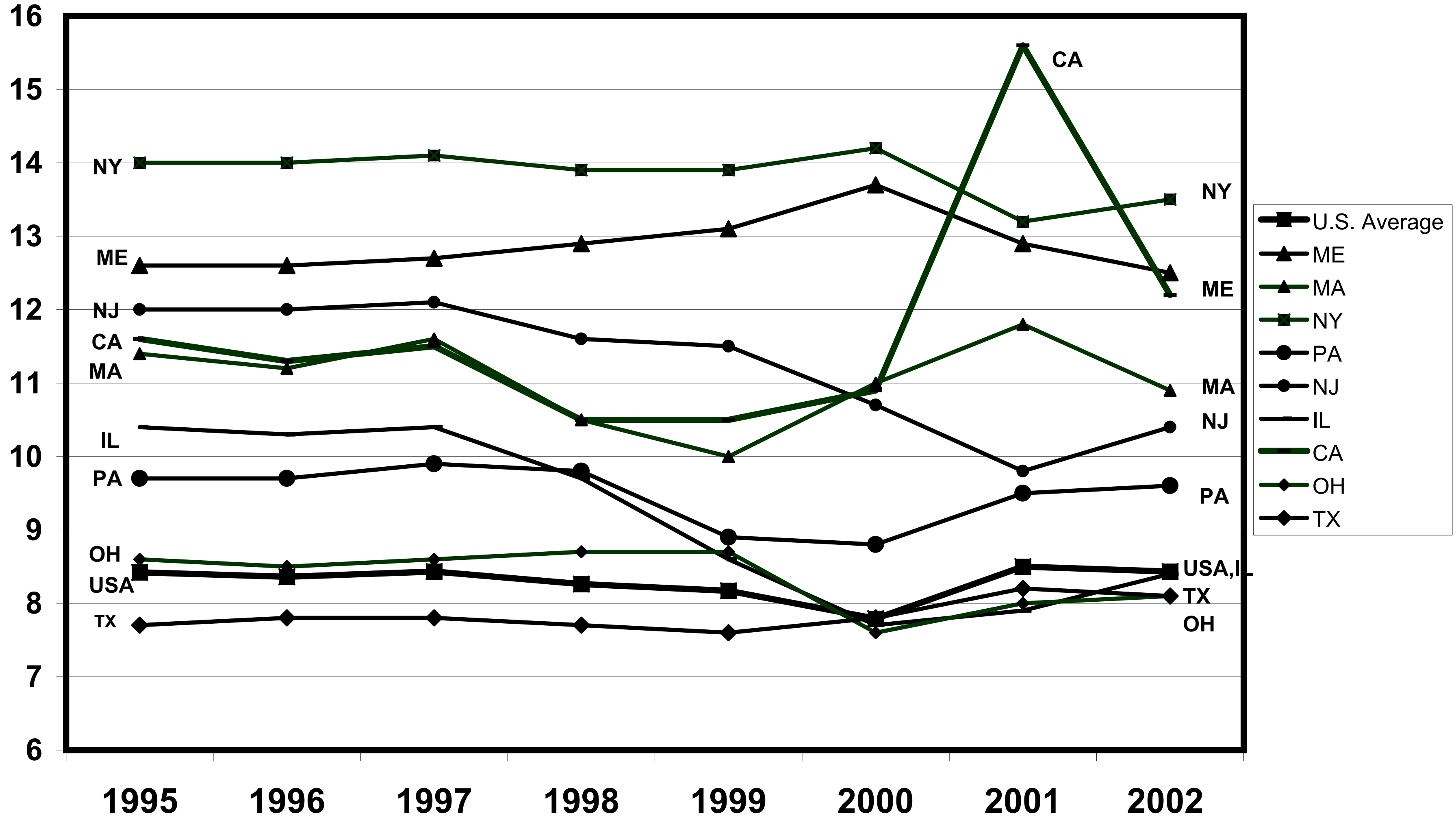


Figure 7
% Change in Residential Retail Price (1995-2002)

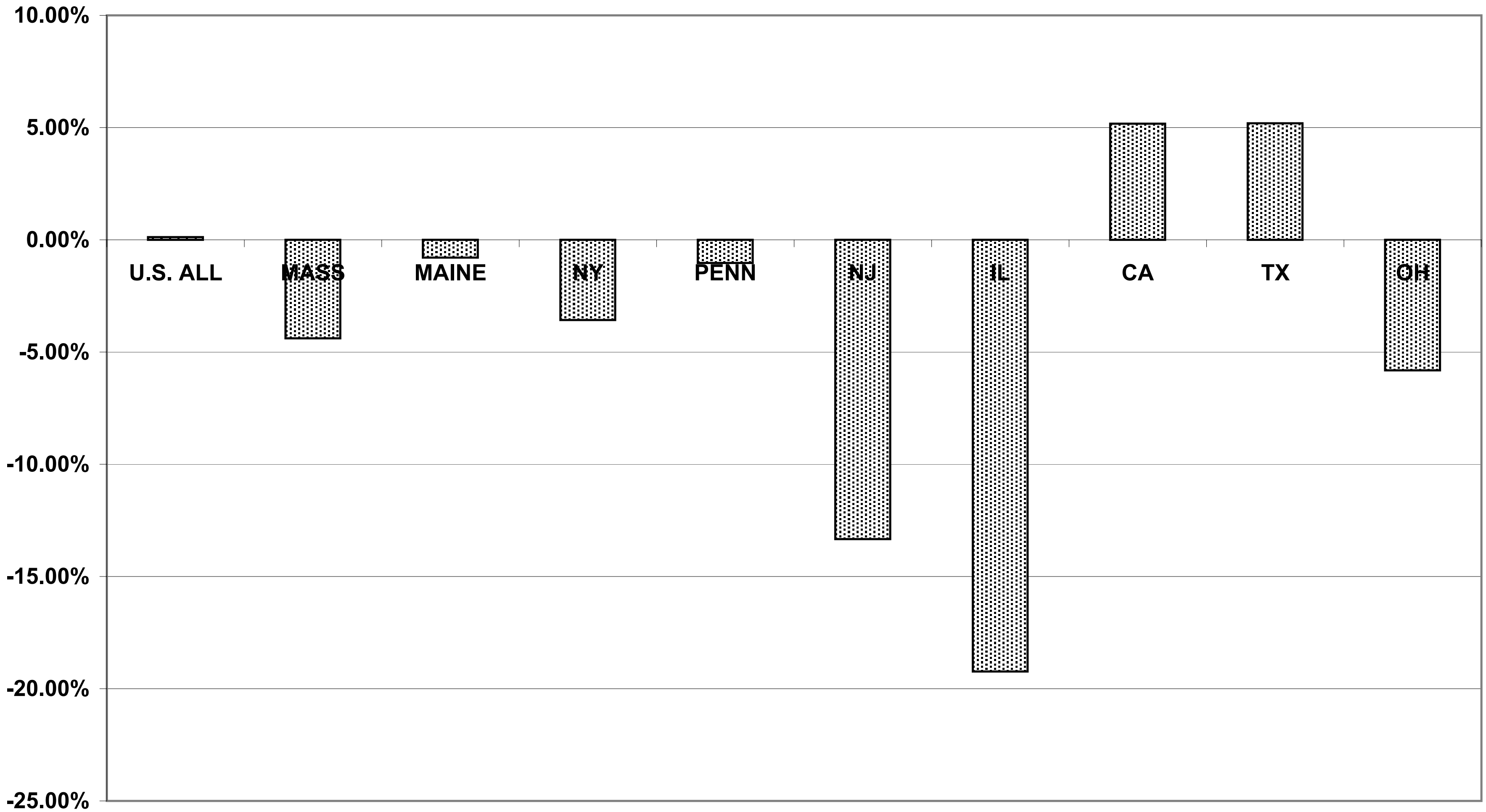


Figure 8
PERFECTLY COMPETITIVBE WHOLESALE SPOT ELECTRICITY MARKET

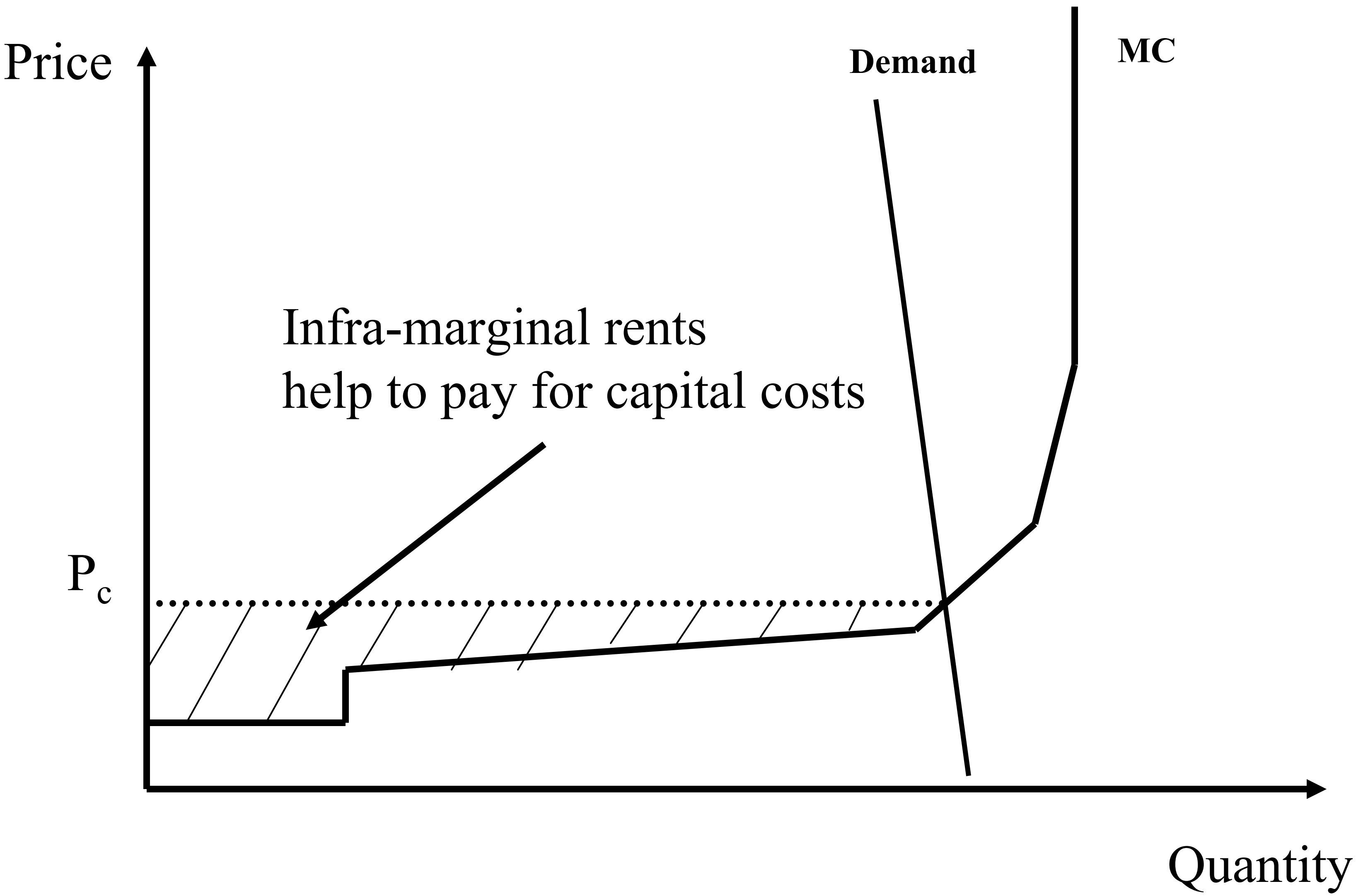


Figure 9
RATIONING SCARCE CAPACITY

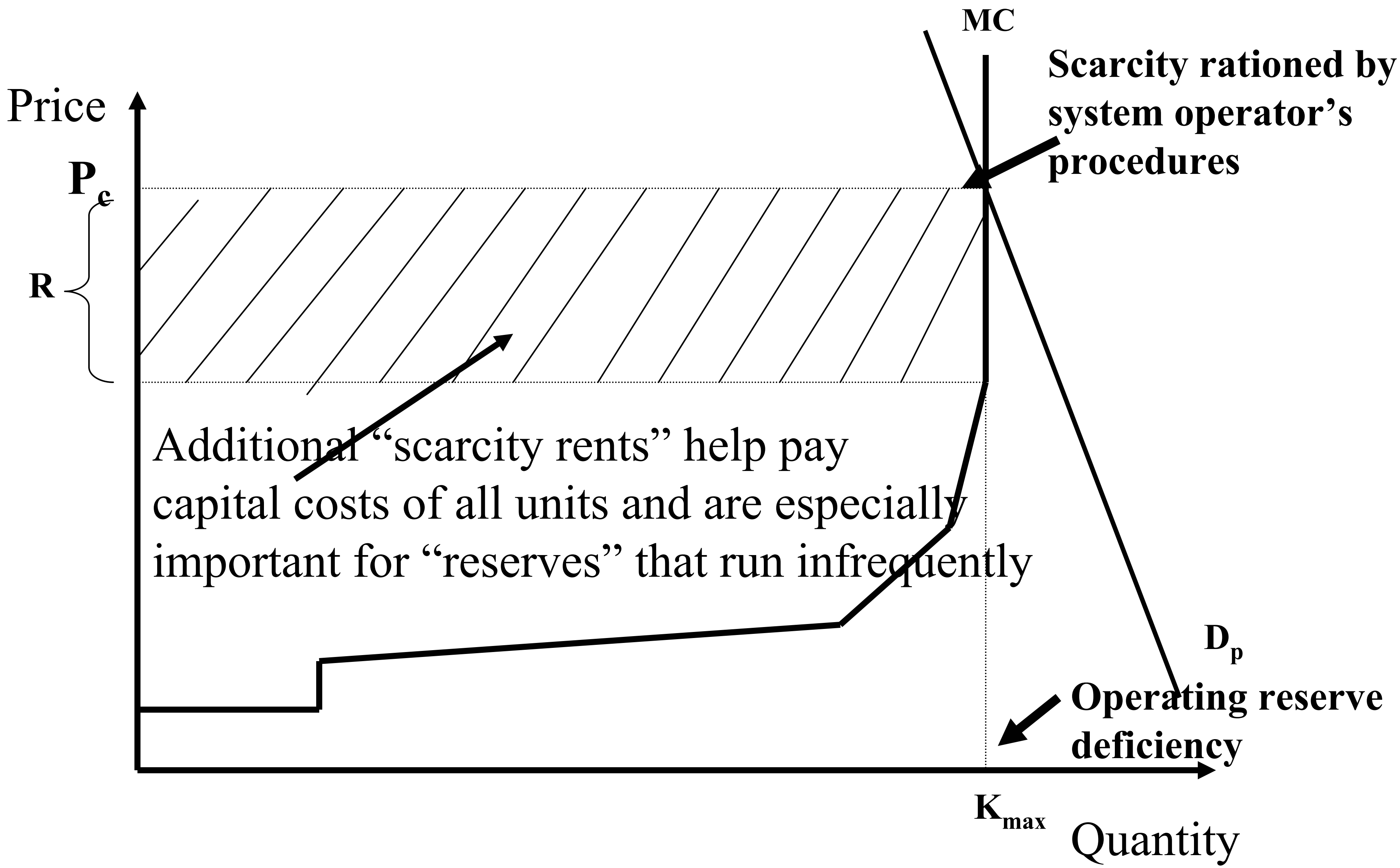
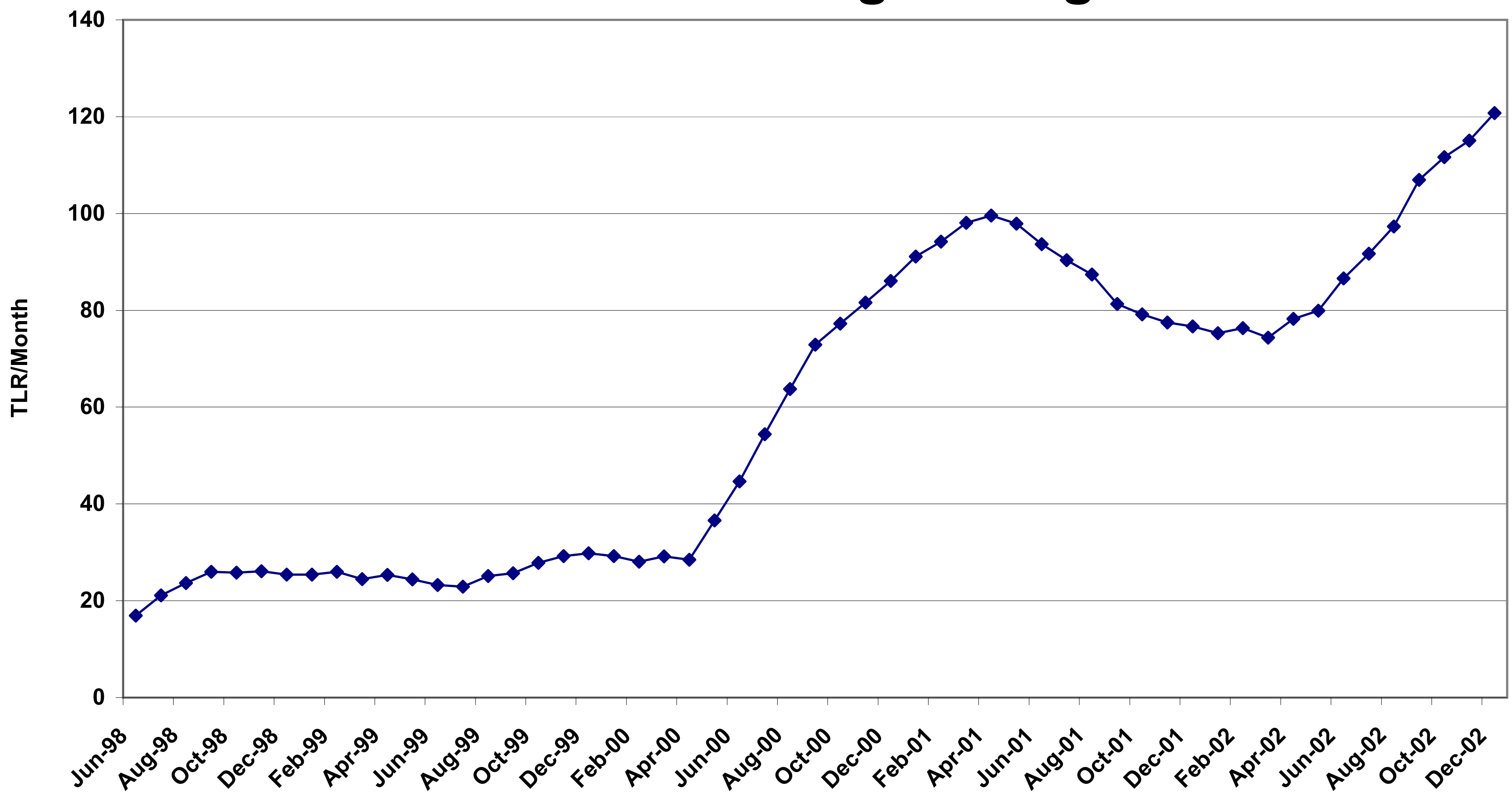


Figure 10
Transmission Line Relief Loadings (TLR)
12-Month Moving Average



Source: North American Electric Reliability Counsil (NERC)