

Revised December 29, 2005

SUPPLY SECURITY IN COMPETITIVE ELECTRICITY AND NATURAL GAS MARKETS

Paul L. Joskow
Massachusetts Institute of Technology¹

1. INTRODUCTION

Perhaps the most frequently expressed concern about electricity sector liberalization reforms that I hear from government policymakers is that competitive electricity markets are not consistent with achieving acceptable levels of reliability or supply security. They point to rolling blackouts, voltage reductions and public appeals for emergency conservation in California, Ontario, Chile, New Zealand and Brazil, the network collapses in the Eastern and Western U.S., Italy and elsewhere, and what appears to be inadequate investment in new generation and transmission capacity to meet forecasts of “need.” This question is asked much less frequently with regard to liberalized natural gas markets. However, the decline in UK North Sea production and the expected increase in UK reliance on imports through interconnectors and liquefied natural gas (LNG) shipments to meet future demand has led to similar questions being raised in the UK. Growing demand for natural gas, rapidly rising natural gas prices, disappointing supply responses in North America, and recent cold winter natural gas “shortage” alerts in the Northeastern U.S., are starting to raise similar questions in the U.S. as well. The increasing use of natural gas to generate electricity has also led to

¹ Elizabeth and James Killian Professor of Economics and Management, MIT Department of Economics and MIT Sloan School of Management. Director, MIT Center for Energy and Environmental Policy Research. Prepared for the Beesley Lecture in London on October 25, 2005.

increased interest in the implications for supply security of the resulting linkages between liberalized electricity and natural gas markets on supply security.

Are there “supply security” problems that result from the structure, behavior and performance of liberalized electricity and natural gas markets, or the way that the transmission and distribution infrastructures they rely upon are regulated, or a combination of both? If so, what can be done to improve performance? Or is it just the concerns of nervous politicians or special pleadings of interest groups that might benefit from regulatory interventions into these markets?

There is no inherent conflict between the liberalization of electricity and gas sectors that meet reasonable supply security goals as long as the appropriate market, industry structure, market design, and regulatory institutions are developed and implemented. Moreover, there is little evidence that liberalization has, at least yet, reduced supply security in most developed countries and considerable evidence that supply security has improved in some developing countries that have adopted comprehensive liberalization programs. However, the effective liberalization of the electricity and gas sectors does create a number of challenges for institution building and governance that must be recognized and addressed for liberalized systems to perform reasonably well from a supply security perspective.

In the next section I provide what I consider to be a reasonable definition of what “supply security” means in liberalized gas and electricity sectors in the short run and the long run. This leads to a brief discussion of why supply security issues are of more concern in electricity and gas sectors than they are with goods and services bought and sold in other competitive markets. I then turn to a discussion of supply security issues

that may arise in the regulated network segments of the electricity and natural gas sectors. Supply security issues associated with the supply of commodity natural gas and investment in new electric generating capacity are discussed next. This leads to a brief discussion of the role of voluntary demand response for supporting good performance of these markets. The paper concludes with a discussion of growing linkages between liberalized natural gas and electricity markets and their potential implications for supply security. The paper draws primarily on examples from the UK and the US and compares and contrasts their approaches to liberalization of electricity and natural gas markets.

2. WHAT IS “SUPPLY SECURITY?”

Policymakers are not always very clear about what they mean by “supply security” and why they are particularly concerned about it in the case of electricity and natural gas. Accordingly, my first task is to define more precisely what it is that I think policymakers mean when they express concerns about “supply security” in liberalized electricity and gas markets. First, they are concerned about “involuntary rationing” of demand in the form of controlled rolling blackouts, uncontrolled transmission network failures, distribution network failures, and the process of public appeals and government exertions to reduce demand that accompany “supply emergencies.” Involuntary rationing of demand can be very costly to individuals and businesses. These costs grow as outages are more sudden, more frequent, of longer duration and the geographic expanse of involuntary rationing expands. Public appeals to reduce demand in response to supply emergencies reflect badly on policymakers and the liberalization policies they support.

Public reactions to such appeals are also sensitive to how frequently policymakers must resort to them.

Second, policymakers are also concerned about high prices, or at least sudden increases in prices, for electricity and natural gas that naturally emerge to balance supply and demand when supplies are “tight.”² Of course, high market prices resulting from a tight supply situation provide the economic signals that provide potential suppliers with the incentives to expand supplies. And periods of relatively high and relatively low prices are to be expected in all competitive markets. Rationing by price is also generally far superior from a social efficiency perspective to involuntary administrative rationing. Nevertheless, it should not be surprising that consumers are unhappy when prices for electricity or natural gas increase significantly. There is something about electricity and natural gas (as well as gasoline, diesel fuel, and heating oil) that makes consumers especially unhappy about large sudden increases in prices and about even relatively brief involuntary outages. Energy costs are a significant fraction of consumer budgets and the short-run demand elasticities for these energy sources are very low. As prices rise, consumers cannot easily avoid paying the piper by switching to substitutes. In most developed countries, electricity and natural gas systems have been highly reliable as well for the last few decades and regulation has partially shielded consumers from price volatility.

To further burden elected government officials, involuntary rationing of demand (blackouts) and unusually high prices are highly correlated with one another in both the short run and the long run. In electricity and natural gas markets tight supply

²By “tight” I mean (loosely) that supply capacity constraints are being approached and relatively small shifts in demand lead to large swings in prices.

contingencies are first revealed in higher prices and associated price-driven demand responses. However, in electricity and natural gas markets supply and demand cannot always be balanced with prices and related market mechanisms such as interruptible contracts. At some point conventional price-driven demand response may not be available to reduce demand to match available supply fast enough to satisfy the need to maintain the physical integrity of the network. This leads system operators to turn to involuntary rationing of demand to maintain the integrity of the network. When system operators resort to involuntary rationing they also reduce market efficiency and (typically) create a large gap between market prices and the value of unserved energy or lost load.

Developments in California's electricity markets in 2000-2001 became an "electricity crisis" (Joskow (2001)) both because of involuntary rationing of demand (California Independent System Operator (2001)) and because of a sudden and dramatic increase in wholesale electricity prices (California Independent System Operator (2002)). Recent "electricity crises" in Ontario, Brazil, and Chile were characterized by a similar combination of high prices, voltage reductions and rolling blackouts. The California electricity crisis also sensitized policymakers to potential market power problems in electricity markets (Borenstein, Bushnell and Wolak (2002) and Joskow and Kahn (2002)) and in the U.S. led to the expanded use of wholesale market price caps and other market power mitigation mechanisms.

Involuntary rationing of demand on electricity and gas networks is not always associated with high prices. Failures of the transmission and distribution network can lead to involuntary outages with no visible effect on market prices and no role for price

driven demand response. The cost to consumers of such outages can be quite high, especially if they are unanticipated and are sustained for a long period of time. In a large-scale network collapse such as those that occurred in the Northeastern U.S. and Italy in 2003, tens of millions of people lost their power in a few seconds. When these transmission networks collapsed there was excess demand, a surplus supply of generating capacity, and a zero price. This is not the typical configuration that we see in a textbook model of supply and demand in competitive markets!

Although perhaps an oversimplification, it is useful to group “supply security” concerns into two categories: (a) short run system *operating reliability* and (b) long run *resource adequacy*.

a. *Operating reliability*: This dimension of supply security refers to the short-run performance attributes of the system as it works physically to balance supply and demand in real time given the existing physical capacity of the system. In electricity, the physical capacity of the system encompasses the generation, transmission and distribution network (including metering and control) facilities. In the case of natural gas, it the physical capability to produce commodity gas, the physical capacity of natural gas storage, transmission (including interconnector), LNG import, and distribution facilities. The relevant indicia of performance here include (a) success in maintaining the network’s physical operating constraints (e.g. frequency, voltage, pressure), the number, duration and resulting costs of non-price rationing (involuntary blackouts), the speed of service restoration when non-price rationing occurs, and the overall costs of operating the system given the physical capacity that is in place.

b. *Resource adequacy*: This dimension of supply security in the case of electricity refers to the long-run performance attributes of the system in attracting investment in generation, transmission (including interconnectors), distribution, metering, and control capacity at the right times and the right locations *to minimize the long run costs of power supplies, including the costs of involuntary rationing of various kinds*. In the case of natural gas the relevant supply segments where long term investment needs emerge include natural gas production, storage, LNG import terminal capacity, as well as natural gas pipeline transmission, distribution and metering.

Obviously, operating reliability and resource adequacy considerations are interdependent. Operating an electric power or natural gas system reliably is a lot more challenging and costly when efficient investments in supply resources have not been forthcoming. Similarly, protocols to meet short run reliability criteria may affect incentives for investment in new facilities in the long run. I will focus here primarily on resource adequacy issues, but we should not forget this interdependence.

3. WHY WORRY ABOUT SUPPLY SECURITY?

Why should policymakers be worried about supply security in electricity and natural gas markets any more than they worry about supply security in any other competitive market? There are a few reasons: (a) important segments of these industries continue to be regulated; (b) physical and economic attributes of these products makes the design of well-functioning competitive markets a significant technical and political challenge; (c) competitive market institutions are still evolving through “reforms of the reforms” and sometimes subject to residual regulation; (d) liberalization is incomplete in

some areas where electricity and gas are supplied and sold competitively; (e) these markets cannot always be cleared by prices, rely on involuntary rationing under extreme conditions, and the costs of involuntary rationing can be very high.

The title of my paper commits a common sin in referring only to “competitive” electricity and natural gas markets. In fact, the phrase “liberalized markets” used extensively in Europe provides a more productive context for evaluating supply security issues. The liberalization of the electricity and natural gas sectors involves a complex institutional transformation from industries composed of vertically integrated regulated monopolies (typically state-owned) to industries with unregulated competitive segments (e.g. generation of electricity, retail supply) and regulated (primarily) monopoly transmission and distribution network segments. For liberalized systems to work well it is necessary to implement sound market institutions and market designs for the competitive segments, vertical and horizontal restructuring, unbundling of competitive and regulated network services, and a compatible regulatory framework to govern the regulated network segments. Poorly performing network segments can undermine the performance of the competitive segments and adversely affect supply security directly and indirectly through their effects on the performance of competitive power markets. Liberalization initiatives have tended to focus a lot on the competitive segments (unbundling, market design, vertical separation, ring-fencing) and much less on the remaining regulated network segments. Outside of the UK, the importance of developing and implementing a good incentive or performance-based regulation framework for the transmission and distribution networks has not been given adequate recognition (Joskow (2005d)). The failure to build liberalized electricity and gas sectors with both good

market designs for the competitive segments and good performance-based regulatory mechanisms for the regulated segments can be a major source of supply security problems.

It is also important to recognize that electricity in particular has a set of unusual physical and economic characteristics that create significant challenges for the development of good market institutions, for developing compatible regulatory institutions, and for integrating supply security considerations of various kinds into market and regulatory institutions. Natural gas networks share some of these attributes, though they are quantitatively less important. The differences between some of the attributes of gas and electricity are reflected in both institutional design challenges and in market performance.

Most discussions of electricity sector liberalization recognize that electricity has some unusual characteristics that create challenges for creating well functioning competitive power markets. These attributes include (a) electricity cannot be stored economically; (b) electricity demand varies widely within days, between days, and between months of the year --- a factor of three from peak to trough and peak demands are sustained for only a few hours each year; (c) the short run elasticity of demand for electricity is very low; (d) electric power networks are physically delicate in the sense that they must meet stringent physical criteria for network frequency, voltage, and stability to be able to supply electricity from dispersed generators to dispersed consumers at all; (e) supply and demand must be balanced continuously in real time to meet these physical criteria and the associated balancing mechanisms must react very quickly to changes in system conditions, including equipment outages, to meet these physical

constraints; (f) most consumers cannot see or respond to short-run price movements that signal supply scarcity at different times and at different locations and which can come and go very quickly; (g) there can be very significant intra-day and day-to-day price volatility to balance demand variations in the presence of capacity constraints and in the absence of storage; (h) a very small fraction of peak electricity demand can typically respond voluntarily to large sudden price increases resulting from sudden imbalances in demand and supply; (i) except for the largest customers, demand typically cannot be physically controlled on an individual basis in the short run so that any administrative rationing must be accomplished on “zonal” basis, making individual price-contingent “priority rationing” contracts³ infeasible for these customers; (j) as a result, an effective controlled “last resort” involuntary rationing system must be in place to keep the entire network, or a large portion of the network, from collapsing so as to avoid both involuntary rationing on the demand side and idle generating capacity on the supply side).

From a longer run perspective, changes in the physical infrastructure of an electric power system can take a significant amount of time to be realized. New generating stations and new transmission lines can take several years to plan and build. Procedures for environmental reviews to obtain certifications to build new facilities can add significantly to the time it takes to change the physical capacity of the network. Major transmission facilities are often especially challenging in this regard since they typically traverse multiple local and regional government jurisdictions.

³Chao and Wilson (1987) develop the theory of priority rationing for the case where individual consumers can be rationed by the system operator.

Many of these attributes are not unique to electricity. Rather what makes electric power systems special is the intensity of the individual attributes and the combination of so many of them in a single product. So, empty hotel rooms or airline seats cannot be stored. But the short run demand for hotel rooms and airline seats is much more elastic than the demand for electricity. And if a big tour bus filled with passengers demanding hotel rooms suddenly shows up in a city where all of the hotel rooms are full, there is no need for a "hotel system operator" to act quickly to avoid all of the hotels suddenly closing down, thrusting their occupants out on the streets. Indeed, the passengers on the bus are turned away by individual hotel operators and just sleep on the bus while the hotel occupants continue to sleep soundly. If the flights are full when a passenger calls for a reservation she can take the train, drive, travel the next day, or take her chances by flying standby; a "stockout" does not disrupt the operation of the airline network. Of course, in the case of airlines an air traffic controller is required for safety reasons and in this sense there is a similarity to the system operator of an electric power network. Furthermore, air traffic control systems have been criticized for not using any economic mechanisms to allocate scarce takeoff and landing slots and congested airspace. The air traffic control system is quite reliable but not economically efficient.

Unlike electricity, natural gas can be and is stored economically, though storage is costly and its ability to replace current production streams limited. Moreover, once gas in storage is released for sale, it may take a significant amount of time to replenish it, while electricity generating plants that run to meet peak demand one week during a cold snap or heat wave can run again two weeks later if the peak demand reappears, assuming that they don't break down in between. Natural gas networks must also meet physical

operating criteria, in particular maintaining minimum physical pressure in the pipes, but the real-time physical operating constraints are less stringent than with electricity since variations in the pressure or packing of the pipeline can accommodate short run changes in supply and demand relatively easily.⁴ The aggregate demand for natural gas has a larger short-run price elasticity than is the case for electricity, largely due to fuel switching capabilities. Historically, there has been relatively more interruptible demand on the gas side than on the electricity side. Since natural gas can be stored in situ or in storage facilities and demand is more responsive to price spikes, commodity prices for natural gas exhibit much less short-run volatility than do electricity prices.⁵ Designing and building new facilities can also take several years as with electric power infrastructure, though existing facilities can often be expanded quickly at modest cost by increasing compressor capacity or adding short loops around bottlenecks on the network. I would argue that creating well functioning liberalized natural gas markets is less of a technical challenge than is creating well-functioning liberalized electricity markets.

Importantly, liberalized electricity and natural gas sectors also have organizational attributes that are different from those that govern most competitive markets. As already noted, these sectors are composed of competitive segments (electricity generation, natural gas production, wholesale marketing and retail supply) and regulated monopoly segments (transmission and distribution). The performance of the competitive segments depends critically on the performance of the regulated monopoly platforms on which they operate. Actions taken by the monopoly network operator as it balances the system to meet

⁴ Generating plants that provide frequency regulation, reactive power support and spinning reserves on an electric power network play a similar role.

⁵ The intra-year spot prices for natural gas vary by one order of magnitude while the unconstrained intra-year spot prices for electricity vary by as much as three orders of magnitude.

physical operating reliability criteria can affect market prices in the short run and incentives to invest in new facilities in the long run.

In addition, the in the case of electricity, market mechanisms that are relied upon for the physical or near physical operation of the system (day-ahead, intra-day, real time energy markets and operating reserves) are “designed” by regulators in consultation with stakeholders, rather than evolving naturally via the invisible hand. These market mechanisms may have design features (“flaws”) that adversely affect the behavior and performance of the market. So, for example in England and Wales, NETA replaced the Pool because it was thought that the Pool had design flaws. In the New England region of the USA, what was viewed as a flawed single price auction mechanism in the wholesale electricity market was replaced with a locational marginal price (LMP) mechanism that reflects the marginal cost of congestion and marginal losses in prices at each major node on the network (ISO New England (2005a)).

Electricity sector liberalization must deal with another set of network issues in continental Europe and North America. The synchronized AC networks in continental Europe and the U.S. span large geographic areas that include several countries or states and multiple system operators with "control area" responsibilities for specified portions of the larger synchronized network. From a physical perspective it is one network. From ownership, control and regulatory perspectives it is several networks. Moreover, although power is traded both within and between individual network control areas, the market designs may differ between them. Coordinating the physical operation of the multiple system operators and harmonizing the market designs in each of them is

important for achieving supply security and economic efficiency goals for the entire integrated network.

The studies of the 2003 U.S. and Italian blackouts either state or imply that electricity sector liberalization per se played no role in the blackouts. I think that this conclusion is too cavalier. Liberalization in North America and Europe has placed increased stress on the reliable operation of electric transmission networks in a number of dimensions. Essentially the same transmission network that existed before liberalization in continental Europe and the U.S. now supports a much greater volume of trading between countries (in Europe) and regions (in the U.S.) than in the past and, as a result, the electricity network runs close to physical operating constraints more frequently. Vertical and horizontal restructuring that has accompanied liberalization has brought more market participants into the system and complicated coordination issues between suppliers and network operators and between network operators in different countries (in Europe) and different regions (in the U.S.). The harmonization of market mechanisms in different countries (Europe) and regions (the U.S.) efficiently to dispatch generation and to allocate scarce transmission capacity on the synchronized network they all share is still a work in progress and leads to a sacrifice of some efficiency benefits from competition. Finally, horizontal and vertical restructuring to support well functioning competitive electricity markets had been (and still is) only partially implemented in much of Europe and the U.S.

Consider First Energy, the company at the center of the U.S. blackout. It is a vertically integrated utility (generation, transmission and distribution) which, at that time,

was still the system operator in its area.⁶ It is fairly clear that it did not view transmission and network operations as a core business for the company and it appears to have devoted limited resources (labor, training, computer and communications equipment) to its transmission business. Inadequate tree trimming and maintenance were identified as the prime initial causes of the blackout. Neither Ohio nor FERC have adopted incentive regulation programs that would have given First Energy incentives to maintain and operate its system reliably and, indeed, regulatory responsibility for transmission in the U.S. is split between the states and the federal government in such a way as to make good regulation almost impossible. It does not require too much imagination to conclude that incomplete and ineffective liberalization made at least some indirect contribution to this cascading blackout.

Natural gas networks in the United States and Europe have some similar coordination characteristics. There are multiple pipeline owners that must coordinate their pricing, scheduling and balancing protocols to make efficient use of the system for delivering natural gas reliably from dispersed production sources to dispersed consumers. The U.S. gas transmission network relies more on parallel (competing) pipelines than does the European gas network. It is also far more advanced with most aspects of liberalization. However, the coordination challenges are not as great as in electricity if for no other reasons than there is more time to respond to changes in supply and demand conditions on the network than with electricity.

So far, the UK has been spared most of the challenges of operating within a larger physical network with many hands on the wheel. In electricity, there is only a DC

⁶ The Midwest ISO (MISO), the independent system operator for a large portion of the transmission networks in the Mid-western U.S. had only taken over limited control area responsibilities by August 2003.

interconnection with continental Europe and BETTA has internalized network operations between England and Wales and Scotland. In gas, the UK has relied primarily on gas delivered to the beach from the UK North Sea fields, on a limited volume of storage, and relatively little on interconnectors with other European countries or on LNG imports. This situation is now changing and the UK gas network will become increasingly integrated with the only partially liberalized continental European gas system and with a growing world market in LNG.

The unusual combinations of physical, economic and organizational attributes does not mean that liberalized electricity and natural gas sectors cannot yield good performance from cost, price and supply security perspectives, especially compared to the alternative of vertically integrated regulated monopolies. It does mean that creating the necessary institutional infrastructure is very challenging. If we can get it right then we should expect to see good performance in all dimensions, including supply security dimensions. If we get it wrong there will eventually be serious performance problems.

4. NETWORK REGULATORY FRAMEWORK

The development and application of a sound regulatory framework for transmission and distribution networks is an important component of an electricity and natural gas liberalization program that has good performance attributes from both a cost and security of supply perspective. For the unregulated market segments of liberalized electricity and natural gas sectors to work well a robust transmission network that can respond quickly to changing supply and demand conditions is essential. The attributes of the regulatory framework affect both the short-run operating reliability and long run

resource adequacy. The regulatory framework that has evolved in the UK over the last 15 years is the international gold standard for electricity and natural gas network regulation within a liberalized sector context.

From a short run operating reliability perspective the challenge is to apply a regulatory framework that exhibits a proper balance between incentives to reduce operating costs, capital expenditures and incentives to maintain or improve reliability in the short run and the long run.⁷ Regulators in liberalized electricity and natural gas sectors were fairly quick to adopt price-cap mechanisms (RPI – x) as an (apparently) simple way to provide high powered incentives for cost reduction while, through the periodic reset of the base price level (P_0), conveying the benefits of lower operating costs to consumers (Beesley and Littlechild (1989)). In the UK, the improvement in productivity, measured over the decade following privatization and restructuring, is impressive (Domah and Pollitt (2001)).

One of the well known problems with a pure price cap mechanism is that it may provide incentives to reduce network reliability, in terms of the frequency and duration of network outages and the speed with which new suppliers and consumers are connected to the network, as they stimulate network owners and operators to reduce operating costs (Joskow (2005d)). In order to deal with this potential problem, network regulatory frameworks have been extended to incorporate targets for various dimensions of reliability and financial penalties and rewards for falling short of or exceeding them. The UK has been a pioneer in this regard, though U.S. regulators have begun to implement similar “quality of service” regulatory mechanisms. However, achieving the right

⁷See Joskow (2005d) for an extensive discussion of the theory and practice of incentive regulation for electricity transmission and distribution networks.

relative marginal incentives for cost reduction and service quality changes remains a challenge, especially building consumer valuations for network reliability and other dimensions of service quality directly into the regulatory mechanism. And there continues to be the potential for new technologies to make it possible to “unbundle” some aspects of reliability or service quality so that individual consumers can express their individual preferences for service reliability.

Of course price cap regulation is not as easy to implement in practice as it may appear to be in theory. Mechanisms must be put in place to set and reset P_0 and x . P_0 in turn must reflect both forecasts of efficient operating costs and a budget for capital expenditures to replace ageing equipment and for new equipment to support changes in supply and demand for electricity and natural gas, consistent with achieving both operating reliability and resource adequacy goals. Moreover, once a capital budget is approved, “mundane” issues like the depreciation rate, the debt/equity ratio, and the cost of capital must be resolved. Although rarely discussed in the academic literature on price cap mechanisms (Joskow (2005d)), the capital budgeting process and the determination of the cost of capital and associated allowed rate of return have very important implications for the long run supply security attributes of transmission and distribution networks from a resource adequacy perspective. Imperfections in the regulatory framework here can have serious adverse network security consequences going forward and also undermine the performance of the competitive electricity and gas segments that rely on the network.

Developing a good forward budget for capital expenditures for the transmission network owner is a very challenging problem. It requires that the regulator implement an

investment planning process in which the network owners offer their proposed investment plans and the regulator, with the help of its own consultants and other stakeholders, must ultimately evaluate them. The regulator can know the firm's efficient capital needs over the next five years only imperfectly and the network owner will always know more about its best estimate for future capital needs than does the regulator --- a standard asymmetric information problem. Moreover, even the network owner can know its efficient capital needs only with considerable uncertainty, since future investment needs will necessarily depend on contingencies as they evolve (demand growth, interconnections, environmental and safety regulation changes) over the "fixed price" period. OFGEM adopted a particularly clever "menu of incentive contracts" approach in its recent DNO price review (OFGEM (2004)) to resolve differences in views about future capital needs between the regulator's consultant and the DNOs. The DNOs could accept a capital budget close to that recommended by the consultant and get a higher expected return and a higher-powered cost-sharing formula or a capital investment budget further from the consultant's recommendation with a lower expected return and a less powerful cost-sharing formula (OFGEM (2004)).

However, given the inherent uncertainties about future efficient levels of capital expenditures and the possibility that the networks will underspend in the face of a hard capital budget and a price cap mechanisms, I do not think that there is any way to avoid some kind of ex post review of deviations from the capital budget to determine whether they were efficient and to provide for recovery of efficient overspend and recapture of underspend compared to approved capital expenditure budgets that did not reflect efficiencies. This places a significant but necessary burden on the regulator.

One of the things that always puzzled me about U.S. regulation of electric utilities during the 1970s and 1980s was the amount of time devoted to arguing about whether the net-of-tax cost of equity capital was say 11% or 12% (nominal). The effect on retail electricity prices of any decision within this range is tiny and imperceptible to consumers once it is include with all of the other elements that go into the retail prices that they see. The effects regulatory decision about the cost of capital on consumer prices are even smaller today as the scope of regulation has been reduced to network charges only. However, the effects on the network owner's incentives to invest can be very large. The U.S. Federal Energy Regulatory Commission (FERC) has historically chosen to allow gas pipeline owners allowed rates of return on equity that are at the high end of a zone of reasonableness because FERC has been very focused on stimulating investment, reducing congestion and increasing reliability. This is one reason why investment in natural gas pipeline capacity has preceded reasonably well in the liberalized U.S. market and there is little congestion on the natural gas pipeline network.

FERC has recently proposed new policies that would promote increased investment in electric transmission investment by reducing regulatory uncertainty and increasing the profitability of transmission investments in response to growing concerns about the consequences of inadequate electric transmission investment and obligations imposed on FERC by the Energy Policy Act of 2005 (FERC (2005b)). The Act requires FERC to adopt incentive or performance-based electric transmission pricing mechanisms that benefit consumers by reducing the cost of delivered power and ensuring reliability by reducing transmission congestion. The U.S. provides an unfortunate case study of how a poorly developed regulatory framework for electric transmission can undermine

investment incentives and how insufficient investment can in turn undermine the performance of wholesale power markets and reduce reliability. The existing framework for supporting transmission investment in the U.S. is seriously flawed. Regulatory responsibilities are split between the states and the federal government in sometimes mysterious ways (Joskow (2005b,c)). FERC initially supported what I consider to be a flawed “merchant investment” model for electricity transmission investment (Joskow and Tirole (2005a)) and confused issues of who pays for transmission upgrades with questions about whether such upgrades are mediated through market mechanisms or regulatory mechanisms or a combination of both. Transmission investments driven by reliability considerations and transmission investments driven by congestion cost reductions are inherently interdependent but have been treated by FERC and some system operators in the U.S. as if they were completely separable (Joskow (2005c)). The U.S. does not even collect statistics on transmission investment and transmission network performance that are adequate to evaluate the performance of the network (U.S. Energy Information Administration (2004)).

Accordingly, it should not be surprising that there has been little progress in developing and applying a coherent incentive regulation framework for transmission. Moreover, there has been little if any investment in transmission facilities to increase interregional transfer capability. As a result, as new generating capacity has been added in the U.S. and as wholesale market activity has expanded, there is growing congestion on the network, increased use of administrative transmission load relief procedures, and the system runs much closer to the margins of operating reliability constraints, making it

more susceptible to network failures and involuntary rationing of demand (Joskow (2005b)).

Inadequate investment in electric transmission infrastructure has unfortunate implications for both electricity prices for consumers in constrained import areas and for reliability. For example, the Eastern U.S. has abundant generating capacity at the present time. However, due largely to transmission constraints, the prices for power vary widely across the region. **Table 1** reports data on forward prices for monthly contracts for January/February 2006 at various locations in the Eastern U.S. as quoted in October 2005. The prices vary from a low of \$94/Mwh (Ohio) to a high of \$204/Mwh (New York City). Natural gas, which has a more robust pipeline transmission network, exhibits much smaller locational basis differences under normal winter operating conditions, though the infrastructure in the Northeast is only barely adequate to meet demand under extreme weather conditions.⁸

In addition to the effects of transmission congestion on wholesale power prices and the social costs of congestion, a congested transmission network makes it more challenging to achieve efficient wholesale market performance. Congestion increases market power problems and the use of highly imperfect regulatory mitigation mechanisms to respond to them. Congestion makes it more challenging for system operators to maintain reliability using standard market mechanisms, leading them to pay specific generators significant sums to stay in the market rather than retire and to rely more on Out-of-Market (OOM) calls that depress market prices received by other

⁸ FERC (2005, p. 145). Note that the large locational basis difference between Henry Hub (Louisiana) and New York in January 2004 corresponded to an extremely cold weather event during which imports of natural gas into portions of the Northeast were constrained by limitations on pipeline capacity. This event is discussed in more detail in the section on interactions between gas and electricity markets below.

suppliers. In New England, the amount of generating capacity operating subject to “reliability contracts” with the independent system operator (ISO) has increased from about 500 Mw in 2002 to over 7,000 Mw projected (including pending contracts) for 2005 (ISO New England 2005, p. 80) and will add hundreds of millions of dollars of "uplift" costs to electricity consumers' bills in 2006 .

Another relevant issue that must be addressed is where to draw the line between investments in regulated monopoly network elements and investments in network elements that will be determined by market forces and whose prices will be unregulated. As previously notes, in the U.S., FERC at one time envisioned that electric transmission network investments, aside from interconnection facilities linking generators with the grid, would be guided by market forces. FERC expected to rely extensively on market signals (locational price differences) and unregulated merchant (or voluntary market participant driven) investment to expand the capacity of the grid (Joskow and Tirole (2005a)), turning to regulated investments identified through an open planning process only as a last resort. Unfortunately, this led to almost no investment in the electric transmission network for several years.⁹ In the U.S., the role for merchant investment in electricity infrastructure now seems to be focused on electric transmission interconnectors between regions, though only one has been built and one is under construction. Both have been secured by long-term contracts with a municipal utility which can pass along the associated costs in regulated prices charged to retail consumers.

I have serious doubts about the viability of a merchant investment framework for electric transmission investment in the near term, even for interconnectors, especially in

⁹ Most of the observed transmission investment is for regulated investments required to build interconnections between the network and new generating units.

Europe and the U.S. where wholesale and retail electricity markets in different countries (Europe) and different regions (U.S.) are still in a state of flux. In the case of electricity interconnectors between countries and regions, I think that merchant investors should be given an opportunity to develop projects, but a meaningful regulatory backstop should be available as well to expand interconnector capacity to reflect opportunities to access cheaper power supplies and to increase the reliability of the system, especially its ability to respond quickly to significant changes in supply and demand conditions which may have low probabilities of being realized. In this regard I approve of the analysis applied by the Dutch regulator in its review of the NordNed interconnector project in December 2004¹⁰ and by the policies on interconnector financing being developed by the European Commission.

On the natural gas side there has been much more success with the development of individual long distance pipelines and expansion projects with the financial backing, through contracts or ownership positions, of groups of shippers and large customers. In the U.S., the development of new pipeline capacity depends primarily on voluntary contractual agreements between pipelines, shippers, LDCs and large consumers, operating under the shadow of FERC regulation of maximum transport prices. However, in the U.S., as in the UK, the liberalization process of the natural gas sector is now reasonably mature and the market and regulatory mechanisms are fairly stable. Commercial arrangements are honored, and one rarely hears concerns in the U.S. about the Canadians cutting off their exports of gas during tight supply situations. There have been some questions raised about whether pipeline investment has expanded adequately

¹⁰ *Decision on the Application of TenneT for Permission to Finance the NordNed Cable ...*,” December 23, 2004.

to meet growing demand for natural gas in the electricity sector, especially in the Northeast which historically has a weak gas transmission network. The traditional reliance on long-term contractual commitments with shippers and large consumers is also being tested as liberalization in the U.S. has moved market participants to shorter term contractual arrangements.

An important question for the UK, as it looks forward to relying more on imports of natural gas to meet demand as production from the UK North Sea areas declines, is how it will mesh its liberalized natural gas market with what are only partially liberalized markets in Europe. In particular, will commercial arrangements between market participants in different countries be honored or will governments in Europe try to “capture” gas that might otherwise be exported for their own citizens during supply emergencies using out-of-market mechanisms and government-induced behavior that restricts the transportation of gas when it is most valuable? Partially liberalized gas markets in continental Europe can lead to higher prices and more frequent demand-rationing in the UK as it relies more on imports than would be the gas in a well functioning competitive European gas market.

Finally, we come to natural gas storage facilities. Pipeline transportation and underground storage are complementary components of a natural gas system. While mainline gas transmission lines provide the crucial link between producing area and marketplace, gas storage facilities help to maintain the system’s reliability and its capability to transport gas supplies efficiently and without interruption. The capability to store gas as backup ensures supply availability in downstream markets during periods of heavy demand by supplementing pipeline capacity. Storage also enables greater system

efficiency by allowing more level production and transmission flows. In some instances, development or expansion of the pipeline network is tied inexorably with storage and vice versa.

Are storage facilities properly part of the regulated gas pipeline network or part of the unregulated gas production and supply sector? It depends on what kind of storage we are talking about and what its purpose is. Dispersed “fast response” gas storage facilities are needed to respond to short-lived pipeline network outages or constraints that would otherwise lead to reliability problems, including unacceptable reductions in pressure or service curtailments. These facilities should be under the control of the gas pipeline network operator (the facilities can be owned or contracted) and a component of the regulatory contract.¹¹ It would also be reasonable for a regulated gas transmission or distribution network owner to have the flexibility to evaluate tradeoffs between expanding delivery capabilities by expanding the capacity of the pipeline or expanding local storage capacity instead to meet peak demand. However, as long as the gas market itself is functioning well, and in particular that price-contingent demand response or other types of voluntarily negotiated interruptible contracts rather than involuntary curtailments are relied upon to balance supply and demand, there is no reason why gas storage facilities designed for other than short-term reliability reasons should not be left to the market. Ten years of U.S. experience with deregulation of entry into gas storage and open access to gas storage facilities controlled by interstate pipelines has been quite good.

¹¹ Just as electricity network operators contract for frequency regulation, operating reserves, and replacement reserves with generators.

The current policy focus is on the provision of information to the market about the status of gas storage inventories to improve market performance.¹²

5. NATURAL GAS RESOURCE ADEQUACY

Competitive natural gas production markets work very well when they are allowed to work without government interference. Suppliers respond quickly to price movements by increasing or decreasing exploration and development activity, injecting gas into storage or taking it out, and adjusting physical production where this is feasible. There are liquid financial markets to allow buyers and sellers to hedge risks and express their views on future changes in supply and demand conditions. Rising natural gas prices have led to a huge expansion in the development of LNG export and import facilities in the last few years. Suppliers do have to confront environmental and certification processes for major new production facilities, especially LNG facilities, and at least in the U.S. this is a significant constraint on expanding LNG import capabilities, as well as on extending exploration and development activity to protected off-shore areas and Alaska. Policymakers need to ensure that environmental and related project certification reviews are conducted in a fair and efficient way so that new facilities are not unduly delayed and their costs increased unreasonably.

Dry gas production in the lower 48 states of the U.S. has or soon will peak and supply and demand are forecast to be balanced by growing imports of LNG (and possibly with gas from Alaska and proximate areas in northern Canada if the necessary pipelines

¹²FERC (2005a)

are built to move the gas down into the lower 48 states).¹³ The UK is in a similar situation, looking for increased imports from Europe through expanded interconnectors and expanded LNG imports to meet future demand as production from the UK fields in the North Sea continue to decline. The growing reliance on imports by the U.S. and the UK does raise some issues. First, the UK will be relying much more on the natural gas supply markets in continental Europe which are only partially liberalized. The big question here is whether pipeline suppliers controlled by other EU countries will honor commercial commitments during supply emergencies or divert supplies to their own citizens using out-of-market mechanisms and the "socialization" of recovery of the associated costs. Such policies can lead to higher prices and more rationing of demand in the UK during tight supply conditions than would be the case if there were a well-functioning competitive European gas market. The UK clearly has an interest in promoting full liberalization and getting government out of the business of allocating gas supplies or controlling prices when supplies are tight.

Second, the primary sources of LNG exports are also the primary sources of petroleum exports (International Energy Agency (2005)). They are concentrated in politically unstable areas of the world. To the extent that one is concerned about "oil supply security," relying on the same suppliers for LNG is not a move in the right direction. Since LNG trades in world markets, these concerns are global and not limited to individual countries. Increasing global diversification of LNG import supply sources can help to mitigate politically motivated supply disruption. I will leave others to discuss potential diplomatic and military responses to these concerns.

¹³ U.S. Energy Information Administration (2005). Exports to the U.S. from the conventional production areas in Alberta also appear to have peaked.

Finally, the increasing reliance of many countries on LNG imports, combined with expanded storage capacity, is already leading to international linkages between natural gas markets in different regions that historically been thought of as being isolated regional markets. These changes will eventually lead to a world market price for natural gas determined by movements of LNG serving to arbitrage locational price differences. This is not a "problem," but requires recognition in traditional "supply security" evaluations. Gas will move around until the arbitrage opportunities are exhausted. This should help rather than hurt on the supply security front as long as governments do not interfere with market pricing and the associated allocation of scarce supplies across locations and consumers.

6. ELECTRICITY GENERATION RESOURCE ADEQUACY: MARKET AND REGULATORY IMPERFECTIONS

I turn now to the investment incentives provided by liberalized electricity sectors to stimulate efficient investments in generating capacity at the "right" times and in the "right" places to balance supply and demand at minimum cost in the long run. Unlike, the situation for natural gas supplies, at the moment, there is considerable concern in the U.S., Canada and other countries that competitive wholesale electricity markets do not provide adequate incentives to stimulate adequate investment in generating capacity to meet reliability standards. Various proposals for regulator determined capacity obligations, capacity prices, and long-term contracting obligations placed on retailers and/or on the system operator are now being considered (Cramton and Stoft (2005), California Public Utility Commission (2005), Joskow and Tirole (2005b)).

At first blush, it is puzzling that policymakers should be worrying about investment in new electric generating capacity. Between 1990 and 2002 there was 26,000 Mw of new generating capacity added in the UK, or about 40% of the initial stock, while nearly 20,000 Mw was retired (United Kingdom Department of Trade and Industry (2002)). In the U.S., roughly 200,000 Mw of new generating capacity entered service between 1999 and 2004, and increase in total U.S. generating capacity of about 30% (Joskow (2006)) from the level in 1998. There has been little new investment in generating capacity in the Nordic countries or Northern Europe in the last few years, but these countries entered the liberalization era with excess capacity. The drying up of investment has sensibly been viewed as a sound market response to a system that persistently had excessive investment in generating capacity (Von Der Fehr, Amundsen and Bergman (2005)).

Despite the enormous quantity of new generating capacity that entered service between 2000 and 2004, and the existence of excess capacity in most regions of the country, U.S. policymakers are now very concerned about future shortages of generating capacity resulting from retirements and inadequate investment. Many of the merchant generating companies that made these investments subsequently experienced serious financial problems and several went bankrupt. The liberal financing arrangements available to support these projects during the financial bubble years are no longer available and project financing for new generating plants is difficult to arrange unless there is a long term sales contract with a creditworthy buyer to support it. Rising natural gas prices have changed the economic attractiveness of the combined-cycle gas turbine technology that has dominated the fleet of new plants. The quantity of new generating

capacity coming out of the construction pipeline is falling significantly, and most of the new capacity under construction now in the U.S. is either being built under traditional regulatory arrangements or benefits from various subsidies and contractual benefits available to renewable energy, primarily wind.

Very little investment in new merchant generating capacity is being committed at the present time in the U.S. System operators in the Northeast and California are projecting shortages and increases in power supply emergencies three to five years into the future, recognizing that developing, permitting and completing new generating plants takes several years.

On the one hand, a market response that leads prices (adjusted for fuel costs) and profits to fall and investment to decline dramatically when there is excess capacity, is just the response that we would be looking for from a competitive market. For 25 years prior to the most recent market reforms, the regulated U.S. electric power industry had excess generating capacity which consumers were forced to pay for through cost-based regulated prices. The promise of competition was that investors would bear the risk of excess capacity and reap the rewards of tight capacity contingencies, a risk that they could try partially to reallocate by offering forward contracts to consumers and their intermediaries. At least some of the noise about investment incentives is coming from owners of merchant generating plants who would just like to see higher prices and profits.

On the other hand, numerous analyses of the performance of the organized wholesale electricity markets in the U.S. indicate that they do not appear to produce enough net revenues from sales of energy and operating reserves (ancillary services) into

the market to support investment in new generating capacity in the right places and consistent with the administrative reliability criteria that are still applicable in each region. Moreover, while capacity obligations and associated capacity prices that are components of the market designs in the Northeastern U.S. wholesale electricity markets produce additional net revenue for generators over and above what they get from selling energy and ancillary services, the existing capacity pricing mechanisms do not appear to yield revenues that fill this “net revenue” gap. That is, wholesale prices have been too low even when supplies are tight.

The experience in the PJM Regional Transmission Organization¹⁴ in the U.S. is fairly typical. **Table 2** displays the net revenue that a hypothetical new combustion turbine (CT) would have earned by selling energy and ancillary services in PJM's spot markets if it were dispatched optimally to reflect its marginal running costs and market spot prices in each hour for the years 1999-2004. In no year would a new peaking turbine have earned enough net revenues from sales of energy and ancillary services alone to cover the fixed costs of a new generating unit and, on average, the scarcity rents contributed only about 40% of the capital costs of a new peaking unit. Based on energy market revenues alone, it would not be rational for an investor to invest in new combustion turbine or combined-cycle gas turbine (CCGT) capacity in the PJM region. PJM has capacity obligations that are imposed on load serving entities (LSE)¹⁵ and there is a market where qualifying capacity entitlements are traded. Capacity obligations in PJM were carried over from its origins as a centrally dispatched power pool into its

¹⁴ PJM is the system operator for the transmission networks in the Mid-Atlantic states and portions of several Mid-western states.

¹⁵ A load serving entity (LSE) is equivalent to a retail supplier in the UK. However, in PJM the bulk of the retail supply is still provided by regulated distribution companies which procure power to serve retail customers.

competitive market design. Sales of capacity entitlements provide another source of revenues for generating units. However, even adding in revenues from sales of capacity at market prices, the total net revenues that would have been earned by a new plant over this six year period would have been significantly less than the capital costs of an investment in new peaking capacity.

This phenomenon is not unique to PJM. Every organized market in the U.S. exhibits a similar gap between net revenues produced by energy markets and the fixed costs of investing in new capacity measured over several years time (FERC (2005), p. 60; New York ISO (2005), pages 22-25). There is still a significant gap when capacity payments are included. The only exception appears to be New York City where prices for energy and capacity collectively appear to be sufficient to support new investment, though new investment in New York may be much more costly than assumed in these analyses (FERC (2005), page 60). Moreover, a large fraction of the net revenue there comes from capacity payments rather than energy market revenues (New York ISO (2005), p. 23).

How can we explain the empirical observation that if investors in new generation expected to rely only on revenues from sales of energy and ancillary services it would not be profitable to invest in new generating capacity? Obviously, there is something about these markets that keeps prices too low. There are three attributes of electricity and electricity networks that I discussed earlier that interact with imperfections in existing wholesale market institutions to cause this "net revenue gap problem." First, individual generating plants are needed to run, and are economical to run in a well-functioning wholesale market, for widely varying fractions of the year to meet demand; from 8760

hours during the year (base load) to perhaps 100 or fewer hours per year (peaking). Second, there is relatively little price-contingent demand response and related quick-response interruptible contracts available in most regional U.S. electricity markets. **Table 3** provides estimates of the share of demand response as a fraction of peak demand available in each of the U.S. regional reliability regions. Due to transmission constraints, the fraction of peak demand that can be managed with demand response in some sub-regions is much smaller. Moreover, some of the interruptible contracts were negotiated under the assumption that there would be no interruptions. Customers who thought that they could get a price discount with no interruptible pain either do not respond to curtailment notices or quickly cancel their contracts. Third, in order to operate the network reliably, system operators purchase frequency regulation services and operating reserves to allow the system to continue operating reliably if there is a sudden failure of a large generating plant or major transmission line and to respond to short-term variations in demand. Operating and related replacement reserves typically amount to about 10% of peak demand. When operating reserves fall below a certain level (e.g. 7% of peak demand), the system operator will take actions to reduce demand to keep operating reserves from falling further in order to avoid a network collapse (Joskow and Tirole (2005b)). Thus, one can think of the capacity constraint as being binding when generating capacity falls below about 110% of peak demand.

The first attribute means that the marginal investment that in generating capacity that just balances supply and demand efficiently in the long run will run only a few hours on average in any year. Moreover, the number of hours this “peaking” capacity runs will vary widely from year to year depending on variations in the supply/demand balance

from year to year. As a result, for investments in peaking capacity to be financially attractive, the associated capital costs must be recovered from market revenues during a relatively small number of hours. Prices during these hours must rise above the marginal operating cost of this peaking capacity if investment is to be economical on a total cost basis. For example, if the annualized fixed costs of a peaking turbine are \$80,000 per Mw/Year and it runs only 50 hours on average in a year than it must be expected to sell its output at prices that yield, on average, a margin (net of operating costs) of at least \$1,600/Mwh over the generating unit's life. However, prices will rise above marginal operating costs in a competitive market only when a supply constraint is reached and "scarcity prices" are set by demand bidding for the opportunity to have access to generation from this scarce generating capacity.

Figure 1 illustrates a wholesale electricity market equilibrium where supply and demand are cleared at a quantity that is below the maximum generating capacity on the system. The price reflects the marginal cost of operating the last generator that economically supplies to clear the market. Generators with lower operating costs earn some competitive market rents at this price and these rents help to pay for their capital costs. **Figure 2** depicts a situation in which generating capacity is fully exhausted and price is determined by demand response actions. At this equilibrium price, competitive market "scarcity rents" are produced to help to pay for the capital costs of the inframarginal capacity as well as for the capital costs of the peaking capacity. In reality things are a little more complicated because demand and supply are stochastic and system operators must hold operating reserves to ensure that they can respond very quickly to, for example, unplanned outages of generation or transmission equipment. This creates an

additional set of demand contingencies where increases in demand are accommodated by reducing operating reserves below target levels. When the network operates with lower levels of operating reserves the probability of a network collapse increases and the marginal social cost of further reducing operating reserves increases as well. At some point the system operator will no longer allow operating reserves to fall further and instead will implement rolling blackouts to avoid a more costly system collapse (Joskow and Tirole (2005b)).

An electric power system with traditional levels of reliability would find itself in "scarcity conditions" due to either operating reserve deficiencies or rolling blackouts only a few hours each year (e.g. 50). When these contingencies arise, the efficient market clearing price can be quite high, reflecting the value of lost load or unserved energy; as much as \$10,000 - \$15,000/Mwh. In a well-functioning wholesale electricity market a large fraction of the net revenues earned by generating units that run for only a small fraction of the year are realized during these high-priced "scarcity" hours. If prices are not right during these hours it will distort investment incentives, lead to underinvestment in peaking capacity and increase the probability of rolling blackouts or a network collapse. Accordingly, investment (and retirement) incentives on the margin are extremely sensitive to price formation during a relatively small number of high demand hours. Unfortunately, absent adequate demand response, the short-run aggregate demand curve eventually becomes vertical and cannot be relied upon to balance (vertical) supply and demand when generating capacity is fully utilized or to determine a clearing price. Under these "scarcity" conditions, administrative allocation and pricing rules, as well as small changes in behavior by the system operator can have a very large effect on prices.

If we look at the distribution of spot prices for New York and New England this is clear (New York ISO (2005); ISO New England (2005a)). A peaking plant that runs 50 hours a year needs net revenue of about \$1,600/Mwh to cover its capital and other fixed costs. The highest hourly price in New York between 2002 and 2004 was \$1000 and the average price for the highest-priced 50 hours about \$300/Mwh. Wholesale prices for electric energy are too low to attract adequate investment to balance supply and demand at conventional levels of reliability in the long run. Why are prices so low? The primary contributing factors are:

(a) All of the organized markets in the U.S. have wholesale market price caps and other market power mitigation mechanisms in place. The price cap is typically \$1000/Mwh (except in California where it is \$250/Mwh). (It is \$AU10,000 in Australia!) These caps are too low to allow prices to rise high enough during operating reserve deficiency conditions or when involuntary demand rationing is being imposed. However, it can't just be the price caps because the price caps are rarely hit even during operating reserve deficiency conditions.

(b) The absence of adequate demand response to be called upon (or relied upon) to reduce demand so that operating reserve and capacity constraints are not violated also ends up depressing prices. Appropriate price formation requires the existence of adequate demand response that there is relatively little short-term demand response in most U.S. electricity markets. There is not enough active demand to bid up the price to market clearing levels.

(c) Because system operators do not expect that rising prices will balance the system they take various "reliability" actions when they anticipate that the system will be

running close to its limits. These actions all have the effect (though not the goal) of suppressing prices. The system operator may start calling on replacement reserves before market prices rise to reflect the scarcity conditions. They may issue emergency appeals before prices rise to signal that there is a supply emergency. They may reduce voltage by 5%, effectively reducing demand and reducing prices. However, voltage reductions are not really free. Equipment runs less efficiently and savings now will partially be compensated for by increased consumption later. By reducing demand in this way under operating reserve deficiency conditions, market prices are depressed and provide incorrect price signals. They may use Out-of-Market (OOM) calls on selected generators that have the effect of paying some generators premium prices but depress the market prices paid to other suppliers. They make up-front payments to small emergency generators to allow them to be called upon during supply emergencies, but these payments are restricted to a small subset of generators and are not available to all generators providing supplies at the same time. They will begin to curtail demand involuntarily before prices rise to levels consistent with the value of unserved energy.

(d) Reliability standards have been carried over from the old regulated regime. They have not been reevaluated as part of the liberalization process. Even if demand were fully represented in the market, there were no price caps, and system operators allowed prices to rise to clear the market before taking administrative reliability actions, it is not clear that the market would yield the levels of reliability that are reflected in current reliability rules. For example, the reliability council that covers New England and New York has a requirement for installed reserve capacity margins that is consistent with an implicit value of unserved energy of about \$300,000/Mwh. This is at least 10

times higher than the highest estimates of the value of unserved energy that I have ever seen. A well functioning market would be satisfied with a lower reserve margin. On the other hand, absent adequate demand response, involuntary rationing is likely to impose a much higher average cost of unserved energy than would voluntary price response.

There are at least three kinds of approaches that can be taken to resolve these perceived electric generation resource adequacy problems. One approach is to fix the imperfections in the spot market. That is, get rid of the price caps, work hard to develop more market-based demand response, require system operators to integrate their reliability actions to movements in market prices and to rely on generally available market mechanisms to balance the system, and to reevaluate reliability criteria to ensure that they reflect reasonable measures of the value of reliability to consumers. U.S. policymakers are not enthusiastic about lifting the price caps. It has proven difficult to convince system operators in the U.S. to rely more on market mechanisms during extremely tight supply contingencies. Efforts are underway to expand demand response, but progress has been slow.

A second approach is to impose (better designed) forward capacity obligations on Load Serving Entities (LSEs) (or retail suppliers in UK parlance) that require them to contract forward for generating capacity to meet their peak demand plus an administratively determined reserve margin. Load serving entities would then have to contract in advance for installed generating capacity that is expected to be available to meet their peak demand (plus a reserve margin) as well as to buy energy (forward or in the balancing market) to meet realized demand. The installed capacity reserve margin obligation is set to reflect the system operator's reliability criteria (e.g. 14% to 18%

reserve margin at a "competitive" peak demand level). The price for this capacity acts as a sort of "safety valve" to produce the net revenues required to attract investment in generation consistent with engineering reliability criteria (Joskow and Tirole (2005b)). All generators get the benefit of the associated capacity prices and, in principle, the demand side should see this price as well. Variations on capacity and forward contracting obligations have been tried in the U.S. and are now being refined (California Public Utilities Commission (2005)).

A third approach is to require the system operator to solicit bids for long term contracts with new generators only in order to acquire additional reserve capacity that is not being provided by the market. This approach effectively involves price discrimination between existing generators and new generators that have access to the long term contract option. In light of the pricing issues identified in the organized U.S. wholesale markets, this approach is likely to lead to ISO procurement crowding out market-based entry of new generating plants. This in turn is likely to distort the mix of generating technologies that end up being built. There is presently a lot of enthusiasm in the U.S. for this approach, largely because it makes it unnecessary for customers to pay the full market value of capacity to existing generators.

There is another set of deterrents to investment in new generating capacity in the U.S. and some other countries. The future regulatory and market environments are very uncertain. Some states have embraced liberalization reforms and others have not. There is a continuous set of reforms of the previous reforms of wholesale market institutions. Infirmities in retail competition programs have delayed the development of a strong competitive retail supply segment that could do more long term contracting (or

acquisitions) with generators. Given all of this policy uncertainty and the continuing “reform of the reforms” it is not hard to figure out why potential investors in new generating capacity are unwilling to commit capital unless they can get long-term contracts with credit-worthy buyers.

The UK's electricity sector is in much better shape from a generation resource adequacy perspective than is the electricity sector in the U.S. It has a reasonably stable and well developed set of market institutions. Retail competition has evolved nicely. There are no price caps. The system operator must go to the markets first to balance the system. It is incentivised to balance the system efficiently. The market and regulatory frameworks for electricity in the UK are by far the best on earth. This suggests that the UK should be cautious about making significant additional changes to its wholesale electricity market institutions. There is significant value in market and regulatory stability. If it's not badly broke my advice is not to try to fix it.

7. IT'S THE DEMAND SIDE STUPID

In both electricity and natural gas markets, operating reliability and resource adequacy issues are significantly reduced when there is a large active and credible demand side that is available to allow supply and demand to be balanced with market mechanisms, associated price movements, and with a minimum of administrative intervention by the system operator to meet reliability standards. Reliable network operations can be maintained with traditional market mechanisms, the consumers who place the lowest values on service will be curtailed (voluntarily) first, market prices will

rise to reflect the value of unserved energy, and investors can be more confident that they will be paid the true competitive market price once they enter the market.

Increasing demand response in competitive electricity markets has proven to be difficult. Obviously, customers who cannot see the real time price because they do not have a real time meter and are billed based on load profiles are not going to show any demand response during power supply emergencies, aside from feeling like they are doing their civic duty by responding to emergency appeals. And real time meters and associated data processing are still relatively expensive compared to the savings that smaller customers are likely to realize from installing them, taking account of the reality that they are not going to sit and watch the meter all day and night. New technologies are likely to be required to make it easier for customers to pre-program their appliances and equipment to respond to changes in prices or system conditions. Efficient levels of demand response will only be stimulated if prices are allowed to rise to levels far above the wholesale market price caps that regulators have imposed in the U.S. during "scarcity conditions" in order to give consumers appropriate incentives to participate in demand side programs. Moreover, the ability of the distribution or transmission network operator to physically control consumption of individual customers on behalf of their retail supplier may be more important than relying on real time meters (Chao and Wilson (1987)). The customer would then simply contract to have her load reduced or curtailed or certain appliances turned off or cycled when pre-specified wholesale market price levels or system contingencies are reached. Day-ahead price-signals and curtailment notices make these types of priority rationing contracts more appealing to consumers. The key here is to have the communications and control capability linking the network

operator with individual consumer locations. In the U.S., air conditioner cycling programs, relying on signals sent over the electric distribution lines (but it could just as well be the internet or a radio signal), have been especially popular with consumers. A lot more can be done here.

The natural gas industry has always had a lot more demand response capability than has the electricity industry. Dual fuel capabilities in power generation and industry made it feasible and potentially economical for consumers to switch to petroleum when supplies of natural gas were tight. This capability allowed them to contract for less costly interruptible pipeline capacity and to switch to cheaper petroleum products when natural gas prices rose during cold snaps due either to limitations on the volume of commodity gas available or to regional or local network deliverability constraints. Increased reliance on daily metering and pricing of gas also increases price responsive demand.

It is my impression that traditional demand response capability in the natural gas industry have declined in the U.S. and probably in the UK (National Grid (2005)). This decline is due to changes in industry composition, the retirement of conventional steam generating plants that could burn either oil or natural gas, the construction of some combined-cycle gas turbine (CCGT) units without dual fuel capabilities or only limited oil storage, and environmental constraints on burning oil rather than gas (ISO New England (2005b)). This suggests that increased demand for natural gas during extreme weather conditions will be more difficult to manage with demand response, require higher prices to balance supply and demand, lengthen the duration of curtailments of interruptible customers, and increase the possibility of involuntary curtailments. If this

impression is correct it also suggests that this will change the economics of investment in gas storage capabilities and increase incentives to expand storage capacity.

8. INTERACTIONS BETWEEN GAS AND ELECTRICITY MARKETS

This leads directly to a brief discussion of supply security issues that may arise as a consequence of the interaction between liberalized gas and electricity sectors, in the context of increased reliance on CCGT capacity. In a region like New England where the spot electricity market clears with gas-fired capacity about 85% of the hours, the economic relationships are straightforward. CCGT generators transform gas into electricity. If the wholesale price of electricity is high enough to make it profitable they will buy the gas and produce the electricity. If the price of electricity is not high enough to make manufacturing electricity with natural gas profitable, generators with gas and pipeline contracts will sell them to others who value them more highly. For gas fired generating capacity to be in the bid-based dispatch in the electricity sector, electricity prices must be greater than or equal to the marginal cost of manufacturing electricity with gas. Basically, a set of simple arbitrage conditions will simultaneously determine the allocation of gas between the end-use sector and the electricity generating sector and the prices of electricity and natural gas.

What could happen that might restrict the efficient arbitrage of natural gas between these two markets? The New England cold snap of mid-January 2004 provides an excellent example and also further illuminates the issues surrounding generation adequacy discussed earlier (ISO New England (2004) and FERC (2005a)). The third week in January 2004 was the coldest week in New England in about 50 years. Demand

for natural gas was at an all-time high due to high heating demand and both the import capacity of the pipelines and the delivery capacity of the distribution networks were stressed. Pipelines implemented strict balancing rules and imbalance penalties. Natural gas in New England traded as high as \$75/mmbtu (\$7.50/therm) on the peak day (ISO New England (2004) and FERC (2005a)). Local gas distribution companies relied on gas drawn from local “peak shaving” storage facilities and gas released from the electricity sector to meet the unusually high gas demand from residential and commercial consumers with gas space heating. In the end there were no involuntary curtailments on the gas side, though a longer cold snap would have been problematic as local storage could have been exhausted.

While electricity demand in New England peaks in the summer, the winter peak demand is not too much lower. In January 2004 the peak electricity demand was about 10% below the peak summer demand. Although there was a significant surplus of installed generating capacity to meet the peak electricity demand, a large fraction of the gas-fired capacity was unavailable either due to declared mechanical problems or decisions to sell the gas to the end-use market rather than use it to produce electricity. The independent electricity system operator for New England, ISO New England, struggled to schedule enough generation to meet demand to avoid rolling blackout or a network collapse and had to implement emergency protocols as generation supply deficiencies loomed. While the system ran for a time with insufficient operating reserves, there were no involuntary outages of customers (and very little demand response) in the end. There are several observations from this experience (FERC (2005a)):

a. The spot gas market worked well to allocate scarce supplies of gas and pipeline capacity, though gas prices did rise to extraordinarily high levels for a brief period of time.

b. There were incompatibilities in the “time-lines” with which the spot gas and spot electricity markets operated and this hindered efficient arbitrage of gas between electricity and end-use sectors. For example, gas supply arrangements had to be made several hours before the day-ahead market and the specification of day-ahead schedules for the electricity market operated. The gas market also had much less intra-day flexibility than the electricity market.

c. The gas delivery infrastructure was barely adequate to meet peak day demand. The gas pipeline and distribution network infrastructure was the most constrained and the source of the gas supply constraints. Gas demand had expanded much more quickly than pipeline delivery capabilities into the region as nearly 10,000 Mw of new gas-fueled generating capacity was installed between 1999 and 2004. The generators’ reliance on non-firm transportation arrangements and their poor credit ratings reduced pipelines’ incentives to expand capacity. Distribution networks operated at more than 100% of design capability. Limited electricity import capabilities also reduced the ability of the region to access electricity from neighboring regions.

d. The electricity spot market worked poorly and prices were too low. The independent system operator’s (ISO) protocols for managing supply emergencies, especially its reliance on out-of-market mechanisms and non-market orders to generators, depressed spot prices below competitive levels. Despite that fact that the ISO declared that there was an operating reserve deficiency and struggled to balance supply and

demand, spark spreads for CCGT capacity were zero or negative (FERC (2005a)). This distorted incentives for allocating gas between the electricity and end-use markets and provided inadequate “scarcity rents” to signal new investments.

e. During extremely cold weather, generator availability declines due to everything from frozen coal piles to frozen valves and hoses. This suggests that assumptions about equipment availability during extreme weather conditions need to be carefully tested.

The growing linkages between gas and electricity markets means that the performance of both markets can be affected adversely by market and regulatory imperfections in the other market as well as its own market. Market design, regulatory and reliability policies therefore need to be compatible across both liberalized gas and liberalized electricity markets.

9. CONCLUSION

Where electricity and natural gas sector liberalization has followed the right path and has had the opportunity to mature and stabilize, as in the UK, there does not appear to be a significant supply security problem. On the gas side in the UK, the growing need to deal with what are only partially liberalized European gas markets, to facilitate expansion of LNG import facilities, and to expand the interconnector and internal pipelines infrastructure to accommodate these new gas supply sources, appear to be the greatest challenges. The institutions and understanding to confront these changes successfully appear to be in place, though the regulatory process will have to adapt to significantly increased capital budgets and the need to provide adequate investment

incentives. OFGEM appears to have met this challenge in the 2004 electricity distribution price review. The UK will have to be a strong advocate for continued liberalization of European gas markets for reasons of self-interest rather than simply ideology.

While I am reasonably optimistic about electric generation resource adequacy in the UK, I am also of the view that the jury is still out on the issue. To be sure, the structural features that plague the organized electricity markets in the U.S. are not present in the UK. However, I believe that there is a tendency to take too much comfort from the experience with entry of new generating capacity in the UK during the 1990s. The conditions that led to the big influx of new generating capacity in the UK during the 1990s were particularly attractive. The legacy fleet of generating plants included a lot of old inefficient coal fired generating capacity. With low gas prices it was economical to build CCGTs to replace them. High electricity price-cost margins resulting from generator market power made entry even more attractive (Wolfram (1999)). There was also significant exit and a meaningful amount of mothballed plant remained that can be returned to service relatively quickly if forward prices are high enough. These attractive conditions for entry not longer prevail. Carbon prices and other environmental constraints are likely to lead to more retirements of old coal plant in the next several years. The renewable energy program creates additional supply-side uncertainty. The ability to finance merchant generating plants has changed. So, we shouldn't draw too much comfort from the investment patterns of the 1990s. At this point the best strategy is to keep the government's hands off of the generation market and closely to monitor

developments to determine whether there are market failures that adversely affect generating capacity investment incentives.

In the U.S., the primary challenge on the gas side is to facilitate a speedier and less costly process for expanding LNG import authority. On the electricity side there is a lot of work to do to improve wholesale and retail market design, unbundling, horizontal and retail restructuring, and incentive regulation of transmission and distribution networks. If there are going to be serious supply security problems, the areas of the U.S. that have tried to liberalize the electricity sectors but have done so incompletely or incorrectly are the places where supply security problems are most likely to emerge in the next five years.

REFERENCES

- Beesley, M. and S. Littlechild (1989), "The Regulation of Privatized Monopolies in the United Kingdom," *Rand Journal of Economics*, 20:3 454-472.
- Borenstein, S., J. Bushnell and F. Wolak (2002). "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Markets." *American Economic Review* 92(5): 1376-1405.
- California Independent Systems Operator (2001), *Operations Economic Report 2001*, December.
- California Independent System Operator (2002), *Third Annual Report on Market Issues and Performance*, January 2002.
- California Public Utilities Commission (2005), "Capacity Markets White Paper," August 25.
- Chao, H. and R. Wilson (1987). "Priority Service: Pricing, Investment and Market Organization." *American Economic Review* (77): 899-916.
- Cramton, P. and S. Stoft (2005). "A Capacity Market that Makes Sense." Working Paper, University of Maryland. March, <http://www.cramton.umd.edu/papers2005-2009/cramton-stoft-a-capacity-market-that-makes-sense.pdf>
- Domah, P.D. and M.G. Pollitt (2001). "The Restructuring and Privatisation of the Regional Electricity Companies in England and Wales: A social cost benefit analysis," *Fiscal Studies*, 22(1):107-146.
- United Kingdom Department of Trade and Industry (2002), *United Kingdom Digest of Energy Statistics 2002* .
- International Energy Agency (2005), *World Energy Outlook*. Paris.
- ISO New England (2004). "Final Report on Electricity Supply Conditions in New England During the January 14-16, 2004 Cold Snap." October. <http://www.iso-ne.com>.
- ISO New England (2005a). *2004 Annual Market Report*.
- ISO New England (2005b). *Power Generation and Fuel Diversity in New England: Ensuring Power System Reliability*, August.
- Joskow, P.L. (2001). "California's Electricity Crisis." *Oxford Review of Economic Policy* 17(3): 365-388.

Joskow, P.L. (2005a). "The Difficult Transition to Competitive Electricity Markets in the United States." *Electricity Deregulation: Where To From Here?*. (J. Griffin and S. Puller, eds.) , Chicago, University of Chicago Press.

Joskow, P.L. (2005b). "Transmission Policy in the United States." *Utilities Policy* (13): 95-115.

Joskow, P.L. (2005c) "Patterns of Transmission Investment." Mimeo.
http://econ-www.mit.edu/faculty/download_pdf.php?id=1133

Joskow (2005d), "Incentive Regulation in Theory and Practice," NBER Regulation Project, http://econ-www.mit.edu/faculty/download_pdf.php?id=1220

Joskow (2006), "Markets for Power in the United States: An Interim Assessment," *The Energy Journal* (forthcoming).

Joskow, P.L. and E. Kahn (2002). "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000." *The Energy Journal* 23(4):1-35.

Joskow, P. L. and J. Tirole (2005a). "Merchant Transmission Investment." *Journal of Industrial Economics* 53(2): 233-264.

Joskow, P.L. and J. Tirole (2005b). "Reliability and Competitive Electricity Markets." September, (revised).
http://econ-www.mit.edu/faculty/download_pdf.php?id=917

Laffont, J-J and J. Tirole (1993), *A Theory of Incentives in Regulation and Procurement*, Cambridge, MA: MIT Press.

National Grid (2005), "Winter Outlook Report 2005/06."
http://www.ofgem.gov.uk/temp/ofgem/cache/cmsattach/12493_214_05.pdf

New York ISO (2005). "2004 State of the Markets Report," prepared by David Patton. July,
http://www.nyiso.com/public/webdocs/documents/market_advisor_reports/2004_patton_final_report.pdf

Office of Gas and Electricity Markets (OFGEM) (2004), "Electricity Distribution Price Control Review: Final Proposals," 265/04, November, London.
http://www.ofgem.gov.uk/temp/ofgem/cache/cmsattach/9416_26504.pdf?wtfrom=/ofgem/whats-new/archive.jsp

U.S. Energy Information Administration (EIA (2004). "Electricity Transmission in a Restructured Industry: Data Needs for Public Policy Analysis." December, DOE/EIA-00639. <http://www.eia.doe.gov/oss/TransmissionDataNeeds-DH.pdf>.

U.S. Energy Information Administration (EIA 2005), *Annual Energy Outlook 2005*. <http://www.eia.doe.gov/oiaf/aeo/index.html>

U.S. Federal Energy Regulatory Commission (FERC 2005a). *State of the Markets Report*. <http://www.ferc.gov/EventCalendar/Files/20050615093455-06-15-05-som2004.pdf>

U.S. Federal Energy Regulatory Commission (FERC 2005b). "Promoting Transmission Investment Through Pricing Reform," Notice of Proposed Rulemaking, Docket RM06-4-000, November.

Von Der Fehr, N-H, E. Amundsen and L. Bergman (2005), "The Nordic Market: Signs of Stress," *The Energy Journal*, 71-98.

Wolfram, Catherine (1999). "Measuring Duopoly Power in the Deregulated UK Electricity Market," *American Economic Review*, 89: 805-826.

TABLE 1

Forward Wholesale Power Prices
Monthly Contracts for January/February 2006
(On October 7, 2005)

<u>Location</u>	<u>Price (\$/Mwh)</u>
Boston	\$194
New York City	\$204
Buffalo, New York	\$130
Pennsylvania (West)	\$115
Ohio	\$ 78
Ontario, Canada	\$120

Source: *Platt's Megawatt Daily*, October 7, 2005

TABLE 2
 THEORETICAL NET ENERGY AND ANCILLARY SERVICES REVENUE FOR A
 NEW COMBUSTION TURBINE PEAKING PLANT
 PJM
 \$/MW-YEAR

<u>YEAR</u>	<u>Net Energy and Ancillary Service Revenue</u>
1999	\$64,445
2000	18,866
2001	41,659
2002	25,622
2003	14,544
2004	<u>10,453</u>
AVERAGE	\$ 29,265

Annualized 20-year Fixed Cost ~ \$70,000/Mw/year

Source: *PJM State of the Market Report 2005*

TABLE 3

Demand Response in Various U.S. Regional Reliability Areas
(North American Electric Reliability Council --- NERC)

2004 NERC-Reported Demand Response				
	Estimated Peak Demand Response (MW)	Actual Peak Demand (MW)	Estimated Demand Response as Share of Peak Demand	Peak Demand Reponse Growth from 2003 (MW)
ECAR	2,643	95,300	3%	-313
ERCOT	892	58,531	2%	173
FRCC	2,822	42,243	7%	27
MAAC	1,082	52,049	2%	-191
MAIN	3,191	53,348	6%	-18
MRO	544	34,852	2%	-1,052
NPCC	2,115	98,454	2%	2,115
SERC	5,781	157,678	4%	221
SPP	990	39,893	2%	-430
WECC	2,561	141,100	2%	740
TOTAL NERC	22,621	773,448	3%	1,272

Source: Derived from NERC Summer Assessments for 2003, 2004, and 2005. Note: "Estimated peak demand response" includes interruptible demand and direct-control, demand-side management.

Source: FERC (2005a)

FIGURE 1
Wholesale Market Equilibrium Below Capacity Constraint

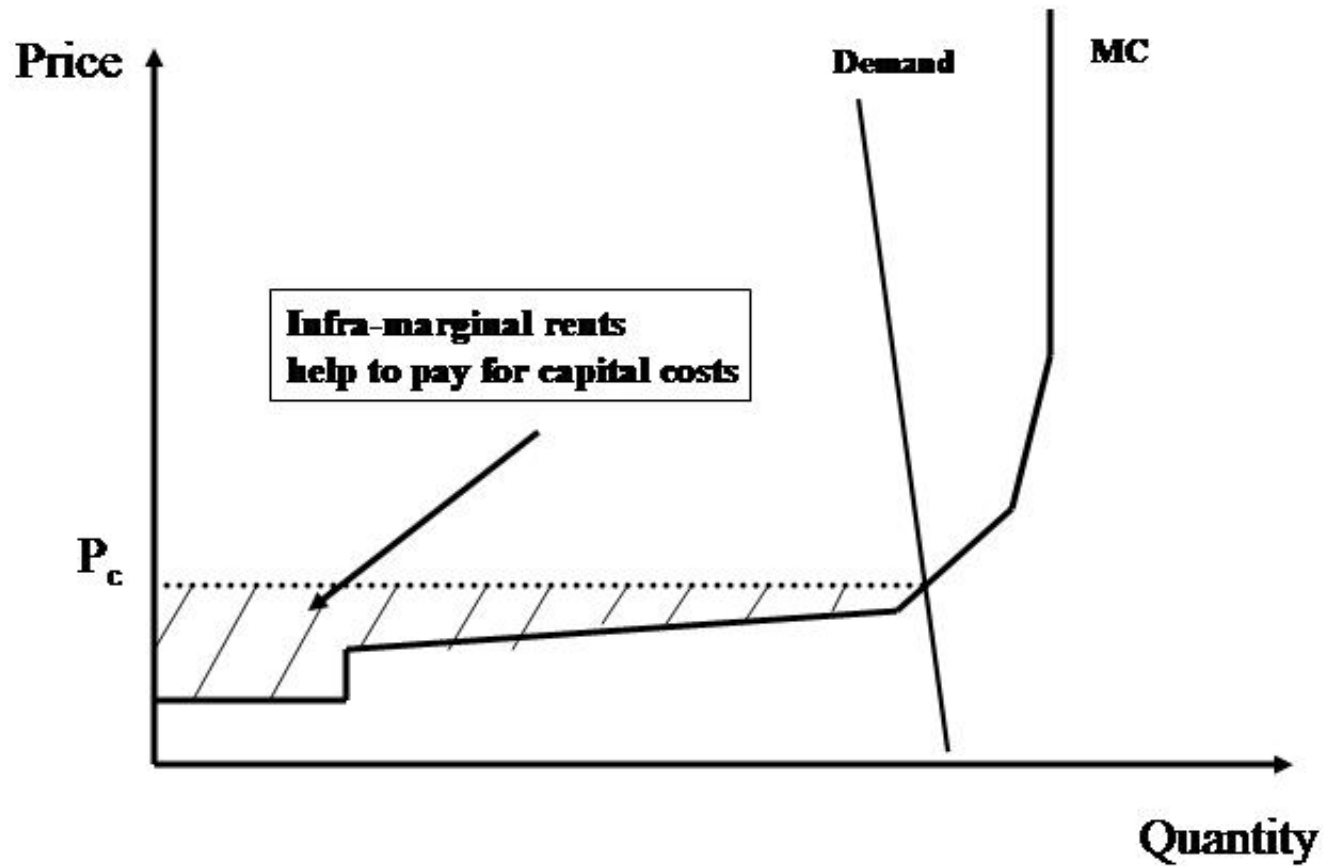


FIGURE 2
Wholesale Market Equilibrium at Capacity Constraint

