# **Electricity Deregulation:** Evidence, Analysis and Public Policy

By

Marilyn Chikaodili Amobi

Thesis

Submitted to City University for the Degree of Doctor of Philosophy in Economics

November 2004

TO THE ALMIGHTY GOD

*'For I know the thoughts that I think toward you, saith the Lord, thoughts of peace, and not of evil, to give you an expected end. Then shall ye call upon me, and ye shall go and pray unto me, and I will hearken unto you. And ye shall seek me, and find me, when ye shall search for me with all your heart. And I will be found of you, saith the Lord: and I will turn away your captivity, and I will gather you from all the nations, and from all the places whither I have driven you saith the Lord; and I will bring you again into the place whence I caused you to be carried away captive' (DAKEs; Jeremiah 29 v 11-14).*

**ACKOWLEDGEMENTS**

I am very grateful to Professor Erik .R. Larsen who did everything possible to support me throughout this research. His commitment is legendary. Also to Margaret Busgith whose exceptional care and concern sustained me. She proof-read the entire thesis.

## Contents

## List of Tables

**List of figures**

# SECTION 1

## Introduction

# 1 Introduction

The electricity industry is the most important utility; it influences the organisation of production and service delivery in all the other sectors in any economy, through that, it influences economic growth and development. That is why its efficiency attracts so much attention. The open question though, is whether unbundling, privatisation and deregulation, are the best policy dimensions that countries should follow to improve electricity sector efficiency. For the countries that choose to unbundle and to introduce competition into generation and supply segments, the regulatory reform is a very complicated process. It usually involves all the other sectors in the economy such as the judiciary, financial, the Treasury and Parliament. The interaction between these institutions during the regulatory reform influences the contractual framework, implementation processes and the initial policies that will be adopted for vesting a new regime; they also determine the successful transition to full competition in retail supply.

Chile pioneered unbundling of electricity systems and the introduction of competition into generation and supply in the late 1970s (Fisher & Galetoviz, 1998). The lack of history and data on markets meant that the public policies, which most of the other markets that emerged shortly after used were based on the results from economic theory and simulation experiments. This thesis uses a historical approach of case studies to examine the outcomes in the privatised England and Wales' electricity industry. It also uses the experience from the regional power integration regime that the Southern African Development Community (SADC) implemented in 1995, as a basis to enhance our understanding of how the initial conditions of a country can determine the methodology and process for its regulatory reform.

This thesis finds some results that allow me to raise some fundamental questions. Is there a guarantee that competition policy will lead to significant efficiency gains in ALL economies? Can price mechanism deliver production and allocative efficiency in electricity markets? Are there other options which some of the emerging and developing countries can use to improve efficiency in their electricity sector? This thesis reveals the likelihood that most Governments' will continue to play an active role in directing the conduct of the multiple agents in the system and the performance of electricity markets. They will use the Sector Regulators to do this; these agencies will use forms of limit pricing (see Joskow & Tirole, 2004) and fair-trading acts, to restrain high prices; and to prescribe codes of conduct between undertakings, in particular with regard to agreements that can inhibit the development of efficient competition. In addition to these, Governments should expect to invest a lot of regulatory input to achieve smooth transition of the regulatory reform, which means significant finance to ensure that competition regimes succeed. I find that the design, price rule and implementation methodologies that England and Wales adopted, exacerbated production and allocative inefficiency in the pool. The results suggest that the same problems might impede the success of competition policy in some of the emerging markets.

Finally, I find that the inherent features of the nation states in sub-Saharan Africa (SSA), such as the unstable socio-politics, lack of institutional framework, poor economic indices and endemic corruption, are barriers to both the development of regional power pools and the entry of foreign investors into the electricity network capacity building. Therefore, I have used the experience from the UK, to develop a theoretic model, which the SSA member states can follow to introduce regional power pools. The model will also help them to develop an environment that might encourage foreigners, as well as some of the indigenous entrepreneurs, to invest in its electricity market. Nonetheless, in the course of this research, I have had discussions with some of the public servants who are directly involved in steering forward the privatisation initiatives in Nigeria. I am aware that the wave of privatisation, which is quite intense in the sub-region, will continue; and despite the issues raised, they confirm that new regional markets are likely to emerge within the next five years.

## 1.1 Purpose of this section

This project systematically investigates the outcomes in the privatised and deregulated electricity market in England and Wales and the factors that might inhibit the success of competition policy in electricity systems in sub-Saharan Africa (SSA). I believe that the objective permits the coverage of industrial organisation, economic regulation and competition policy. Apart from these, the empirical analysis of the England and Wales' pool regime, provides the crucial evidence that is required for policy recommendations on price rule and market design; the areas identified for further research are also founded on that.

This present section introduces the entire research. The technical nature of the industry calls for a broad over-view of its common terminology; therefore, this section starts with a broad review of the industry segments, through which it provides an intuitive description of unbundling and deregulation. As well as that, it exposits what energy balancing on the *national transmission system* (NTS) entails. A glossary of some of the common acronyms that the research contains is included in the appendix.

This section also discusses some of the reasons why Governments choose to unbundle and deregulate electricity systems. The discussion brings out the agency theory issue and its associated debate on whether ownership is the main cause of the inefficient service delivery by the public utilities. It also helps to highlight and to an extent clarify, if there are really significant efficiency gains that countries might earn from unbundling, privatisation and deregulation of their electricity industry. I include a discussion on the strategic behaviour of Generators and the application of competition law into electricity markets. My intention in including these in this section is to help me introduce the reader to the analytical context for some of the conjectures, as well as the inferences that I make in the body of the thesis, particularly in the empirical analysis in section 3.

The rest of this section is presented as follows: 1.2 covers the overview of the industry segments; 1.3 discusses why some countries chose deregulation as the best path to improve efficiency in the electricity sector; and 1.4 reviews the empirical findings on the effect of ownership on the productivity of managers. The conduct of Generators in wholesale trading is summarised in 1.5; and the research questions are set out in 1.6. 1.7 sets out the research framework; 1.8, the structure of the thesis and 1.9 summarises this section.

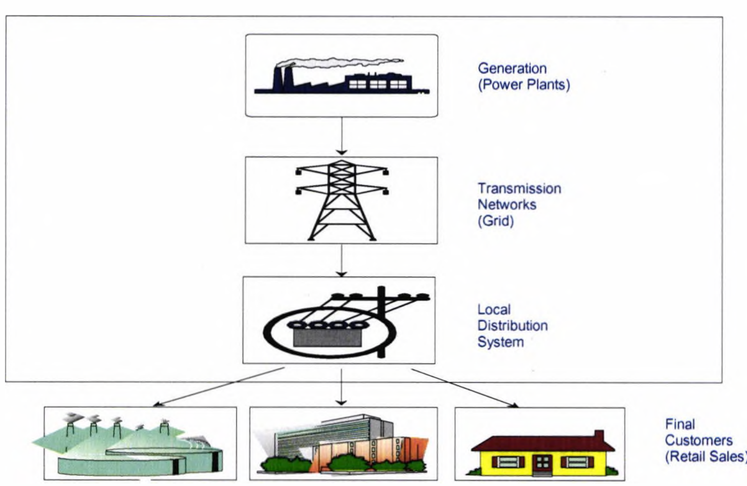
## 1.2 Overview of the electricity industry segments and energy balancing on the *national transmission system* (NTS)

This is an applied economics research for which an understanding of how the electricity industry works is needed; at least from the perspective of generation, it is important. I believe that a discussion of the interaction between the downstream generation and the system operations on the *national transmission system* (NTS) should be part of this introductory section. It will help a reader who is not familiar with the technical requirements in the industry, to follow the issues that I discuss in the rest of the thesis.

Electricity production involves the transformation of inputs such as wind, oil, gas, nuclear, solar, coal, biomass and water, into useable energy. Generators produce electricity as a commodity and consumers consume it as a service. Armstrong et al (1994) identify five stages in electricity production-supply chain as inputs; generation; transmission; distribution and end-user supply. The fuel source for generation: input, is the most important element in producing electricity. Many people consider that input source is a part of generation. Weiner et al (1997) consider that input and generation are one integrated stage; also that the emergence of energy trading created an additional function. They therefore categorise 6 industry segments as generation, transmission, distribution, energy services, power markets and IT products and services. I use Armstrong et al (1994) categorisation of the industry segments, but assume that input is not a distinct function from generation, and in figure 1.1 show four generalised segments in the industry, which are generation, transmission, distribution and retail sales.

**Figure 1.1**

**Production – Consumption Chain**



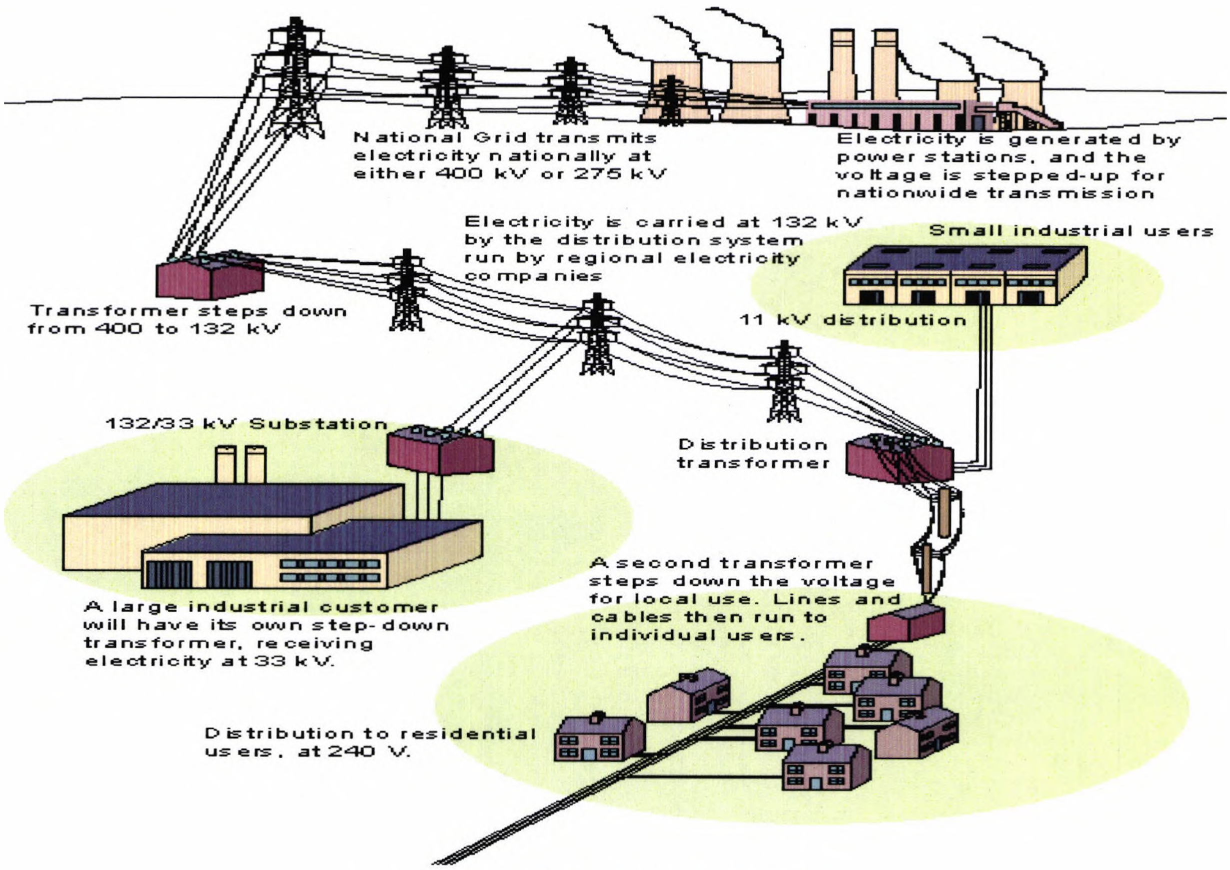

Before deregulation gained prominence over the past 20 years, a vertically integrated utility firm usually carried out the functions in figure 1.1. The exception was in England and Wales: before the industry's privatisation in 1990, the British electricity system was vertically integrated up until the distribution level. From that point, there were monopoly Regional Boards that carried out the distribution and end-user supply services within their regions. Electricity cannot be stored in appreciable quantities; once generated, it is transported on a real-time (minute by minute) basis, through the high and low-voltage wires onto the distribution levels and through to the final consumers. The entire stages in the chain occur as though an electricity system is one integrated vast machine.

Grid injections must always be equal to the offtakes; this equality includes electricity dissipated in the form of heat, which is termed thermal losses that occurs as power is transported from the points of injection onto the Grid, and across the network. The actual volume of thermal loss on any network depends on the distance between generation and load; more losses occur if generation is located far away from demand; however, increases in voltage reduces thermal losses. It is also possible to minimise heat loss by planning the network in a way that ensures that generation is located as close as possible to demand (see Mortlock, 1952, chapters VIII and XV, pages 180 and 343 for discussions on transmission load analysis).

An electricity system usually consists of nodes and other transmission infrastructure, all of which are interconnected by wires and cables. The latter is separated into high and low-voltage wires. Large industrial sites, some of which usually have the capacity to generate their own power, are often directly connected to the NTS. In England and Wales, the threshold for such consumers after privatisation in 1990 was 250MW. In contrast, retail consumers, predominantly households, are only connected onto electricity networks at the *local distribution levels* (LDZs). Figure 1.2 is a simplified representation of an electricity network; it shows the high and low voltage wires of a large industrial customer directly connected onto the NTS on which electricity is transported over 400kV and 275kV lines (see NGC, 2000). It also shows a distribution transformer that steps down the voltage of electricity from the NTS onto the LDZ where it is carried through 132kV lines to a final consumer that is connected to it.

**Figure 1.2**

**A Simplified Electricity Network**




Source: http://www.nationalgrid.com/uk/img/im_generation.gif

*1.2.2 Vertically integrated versus unbundled electricity system*

An electricity system is vertically integrated if a single firm performs all the stages involved in the chain shown in figure 1.1. The monopoly firm decides the source and procurement of input; schedules and despatches all the plants located across the network; co-ordinates capacity and energy balancing on the NTS and manages all aspects of end-user supply services. Most electricity systems were originally organised as vertically integrated 'natural' monopolies. It was more efficient to have a single firm carrying out the whole functions in the production-supply chain because the industry is very capital intensive and requires huge initial sunk investment. Moreover, the benefits of economies of scope and scale meant that the unit cost of production decreases with increases in output (see Scherer & Ross 1990). Given the existence of a single firm, it will reduce its average cost (AC) if it expands its supply density. During the early days, it was therefore thought more economic if only one firm carried out the entire process involved in the production, through the supply chain in figure 1.1 in a geographic area. Because it could do so at a lower cost than if mutually exclusive entities served the same district (see Laffont and Tirole, 1993).

Before rapid privatisation gained prominence in the early 1990s, most Governments did not consider using the private sector to provide electricity. This was partly due to the huge sunk capital, which it required. It was also thought that the possibility of losing the economies of scale and scope might mean that incumbents would use forms of predatory behaviour to deter entry. Or the private firms might not serve unprofitable rural areas. This meant that the equitable distribution of power, which was initially the main objective by most Governments, might be inhibited.

In an unbundled electricity system, the grid functions are separated from distribution, generation and supply. In contrast to a vertically integrated regime, generation and supply functions are assumed contestable. Consequently once an industry is unbundled, it is possible to allow free entry and exit to occur in the segment; the view being that competition would cause the Generators' to behave in a socially desirable way. Consistent with the economics theory of free markets, the net effect expected is that efficient competition would lead to marginal cost (MC) pricing. If that is achieved, Generators would not earn abnormal profits (see Tirole, 1998; Viscusi, et. al, 1995). This cause and effect ideology is based on the *'contestability theory'*, which assumes perfectly free entry and exist into a market that is deemed *'contestable'*. The 'costless' nature of entry and exit means that fixed costs are irrelevant; and a firm can enter into such a contestable market if prices are above costs. The interesting aspect of this expected entry is that it can occur before an incumbent, even if it is a monopoly, can respond. In addition, contestability is also based on the assumption that it is easy to adapt the operating assets in such markets into other productive uses if the lack of profitability induces a firm to exit the industry. Contestability theory is based on the assumption that the invisible hand leads to production and allocative efficiency; firms will make zero profits; there will be Ramsey optimal prices and the firms' will not have the opportunity to systematically cross-subsidise between sub-product markets (see Demsetz, 1986; Baumol, 1982; Shepherd, 1984).

Based on the contestability assumptions that I mention in the last paragraph, once generation is privatised and competition is introduced into generation, multiple Generators are allowed to enter and participate in generating electricity (Pritchard et. al, 2000). The important question is whether competition policy would deliver significant efficiency gains in generation? Or, put another way, is electricity industry contestable?

The electricity industry requires huge 'sunk' costs; and the infrastructure across the sub-segments are specialised, thus they cannot be put to alternative use even when it becomes economic that an agent leaves the market. In generation, plants cannot be put into other productive uses; moreover, the sunk costs are relatively higher in the short than long run (Shepherd, 1984). This makes sense considering that it can take more than 20 years in some cases, to fully amortise the investments made in generation plants. Electricity is not really a contestable market; therefore, like I find in most of the analysis in section 3, production and allocative inefficiency occur when Generators compete to procure capacity for injecting energy on the Grid. Although the price rules may be set with the view to maximise social welfare, the inherent features of electricity, which includes the lack of storage in appreciable quantities, combined with the real-time balancing of injections with off-takes that is required on the NTS, makes pricing difficult. There are usually small number of players that are geographic monopolies, their capacities are limited in the short-run and they engage in repeated interactions, with the knowledge of meeting again in the future, to procure capacity. The structure of the industry plus the design of the commodity auctions enhances tacit collusion. The result is that agents are able to keep prices above competitive levels most of the time; and consequently the aggregate costs of generation in these deregulated regimes are always above competitive levels (see Fehr and Harbord, 1993).

When electricity systems are unbundled and deregulation occurs, the sub-additive nature of the wires and cables mean that the transmission and distribution segment remain 'natural' monopoly's. Therefore, it is economic to vest statutory rights on a single utility to own the transmission infrastructure. This firm is referred to as a *transmission owner* (TO); it will make economic decisions regarding the investments on the network. The other role on the transmission system is the requirement for the co-ordination of the network activities, which is termed the system operations. The utility firm that carries out the role would act in a capacity that is known as the *system operator* (SO). System operations involve the co-ordination of the residual energy balancing on the NTS. Apart from some of the states in the USA, it is common to find in electricity markets, that a single utility firm is given the statutory right to serve in the dual capacity as a TO and SO. Such a firm operates in what is known in the electricity deregulation literature as a non-profit making and an *independent system operator* (ISO). This was the case when England and Wales vested its electricity privatisation in 1990; the *National Grid Company* (NGC) was given the statutory ISO rights.

There are a couple of issues that follow deregulation. For example, given the monopoly rights that the reform confers on the SO, allocative efficiency on the transmission system might be jeopardised; therefore, the firms' tariff is controlled and regulated. There is usually a statutory Regulatory Agency that exists in most markets; it will *'ring fence' the* monopoly business of the SO, set the control price and veto its capacity charging methodology. There are two main tools, which are used in electricity markets, to regulate the monopoly transmission and distribution businesses; these are the *cost of service* (COS) regulation, which some of the states in the USA use. And the incentive based 'X' efficiency tool, the Retail Price Index (RPI) minus an efficiency rate ('X'), which England and Wales uses to regulate its utility industry. The investment decisions of the agents in the COS regulatory environment are more likely to be biased in favour of capital than labour and might not control capital expenditure because they know that they will be allowed their reasonable costs (*Averch-Johnson effect*). In contrast, the firms under incentive-based RPI-X will use best practises to curtail all operating costs that they can control because they will keep the difference between the allowed revenue and the actual expenditure (see Laffont & Tirole, 1993; Viscusi, et. al, 1995). By separating and *ring fencing* the monopoly business of the ISO, it allows it to provide a non-discriminatory tariff and services to all incumbents as well as entrants into the industry. It also helps to curtail any conflict of interest that the ISO may have particularly if it decides to diversify into the competitive segments in the industry; an example is if it enters into generation business.

Finally, regulatory reforms in many cases, plan a phased retail competition because smooth transition is best achieved over time, the design inefficiencies in the initial policies for vesting are identified and the modifications are based on the developments in the industry. Since there are no quick fix solutions to the current issues in electricity markets, it seems to me that these modifications, which by their nature curtail market failure, are easily identified if sufficient time is allowed for the industry to go through its period of hard landing.

The generic pattern for creating electricity markets is first to unbundle the industry segments; thereafter, to introduce competition into generation, as with the wholesale commodity trading. In supply services, it is also possible to introduce competition into metering, billing and customer services. All the regulatory reform processes are carried out consistently, with the overriding objective being always to ensure that the aggregate costs of generation and transportation are economic; a long-term reliability of supply guaranteed and appropriate *long term investments* (LTI) are made. These would ensure that the network is always safe and secure (see Hunt and Shuttleworth, 1996; Sidak & Spulber, 1998).

### *1.2.3 Safety and security of the transmission system*

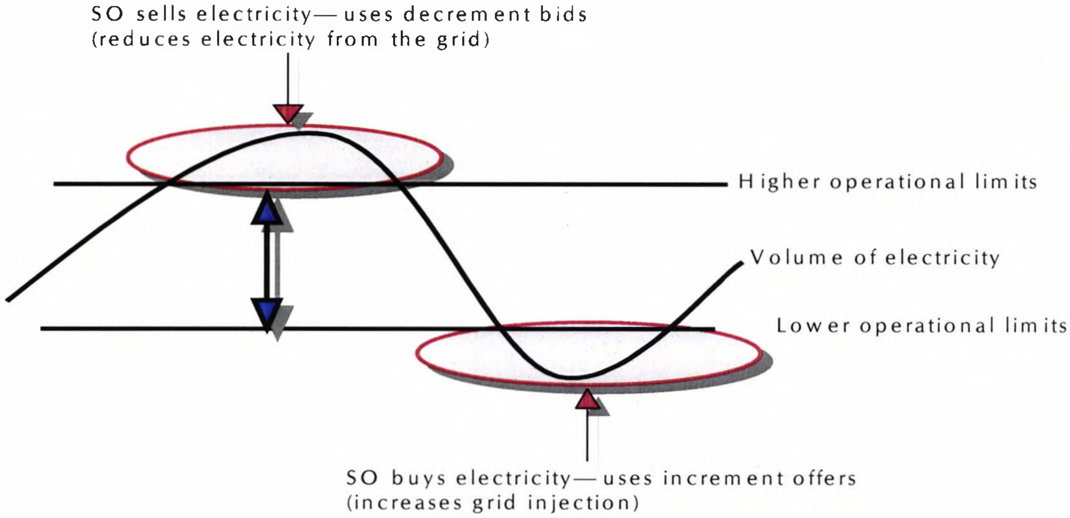
Electricity follows the laws of physics—the path of least resistance and not a contract path. Part of its efficiency enhancing attributes, which includes the safety and security of the NTS, requires that there is always equality between grid injections and the total consumption including thermal losses. Any deviations that occur between injections and *off-takes* can travel instantly across the system; and may cause expensive damage to electricity infrastructure as well as to the end user appliances. It can also lead to a complete shut down of a network, as was the case in August 2003, when 'six states in the USA and one in Canada lost electric power in North America's worst ever blackout, affecting millions of people and thousands of businesses'. In London in the same August 2003, 'two faults in rapid succession in equipment operated by the National Grid Company led to a loss of electricity at 6.20pm in an area of South London between Wimbledon and Hurst in Kent. This was a loss of 20% of the total electricity supply to London at that time, 410,000 customers were affected, with supplies lost to a large part of the London Underground and Network Rail' (Parliamentary Office of Science & Technology, 2003: 2).

In competitive regimes, different agents make independent yet instant decisions that contribute to the real-time processes from generation through to transmission, distribution and supply to the final consumer. There is a notional security equilibrium that changes if the system deviates from its accepted frequency responses, voltages and energy tolerances. Inaccurate demand and weather forecasts, as well as inter-system transmission shocks, which can include unexpected outages—can and do—cause energy imbalances. Traders' commercial strategies may also lead to sub-optimal situations some of the time. It is also possible that flows onto the network from electricity that the *combined heat and power* (CHP) Generators produce from their simultaneous generation of heat and electricity can cause energy imbalance. Since the system is integrated, there are externalities; it is good if excess power comes onto the system when there would have been a shortage, but bad if it comes on when it is not required. It is also a major problem if technical constraints make Generators' unable to inject contracted power onto the grid. For whatever reason deviations occur, the *system operator* (SO) carries out balancing actions to restore the network to its acceptable tolerances whenever it veers towards an imbalance.

Balancing action involves the SO literally *'replacing'* missing energy by calling on Generators that offered to increase generation (called *increment* offers) during such periods. On the other hand, the SO will *'remove'* excess power by asking Generators that bid to decrease generation (called *decrement* bids) to reduce generation. In the latter case, Generators that decrease consumption from the network usually have the capacity to generate their own electricity requirement. The role of the SO as the *residual balancer* means that it is the *'invisible'* and equilibrating hand that restores system tolerance levels in electricity markets. Figure 1.3 is a very simplified illustration of energy balancing by the SO.

**Figure 1.3**

**Energy balancing**




### *1.2.4 Externality and security costs*

Generators pursue their commercial strategies without individual regard to the system security position. The security and safety of the system is threatened if it deviates from its balanced tolerances; and depending on the severity of the energy imbalance, the system may collapse. This is also the case if Generators inject more power onto the network than its maximum physical capacity; or if transmission constraints mean that Generators have contracted positions above the network capacity levels. In the latter case, congestion will occur.

The SO manages the externalities that arise from congestion and energy imbalance; and all the grid users, including final consumers', share the benefits or negative effects of having a balanced system. The negative externality is the increased threat to both the security of supply and the safety of the system; this increases the within the day costs that the SO incurs for balancing the system. If the *ex-post* pool sell price in a residual pool includes the costs of system security; then it is the final consumers that ultimately bear the SOs balancing costs.

In the other oligopoly markets, suppliers use their supply functions to manipulate prices (see Tirole, 1998), in electricity, the Generators use capacity (See Green and Newberry, 1992; Bolle, 1992) to manipulate both the spot (day-ahead and within the day) and the contracts market. For example, this project finds in section 3 that Generators persistently used capacity to manipulate the Uplift costs in England and Wales' pool between January 1994 and December 2000. Uplift was the 'pot' that captured all the necessary costs that NGC spent to balance the system within the day. It consisted of costs such as those incurred to resolve constraints, start-up and for the procurement of balancing services—reactive power, spinning and non-spinning reserves, frequency responses and black start. This cost was charged to the *Regional Electricity Companies* (RECs) in proportion to their *offtake* throughout. The RECs were allowed to pass on 95% of their purchase costs to their captive customers (Armstrong et al, 1994); therefore, in reality, the final consumers indirectly bore the costs that the NGC incurred to maintain the system within its tolerance levels.

It is reasonable to expect that economic efficiency on the transmission system will be enhanced if the Generators' balance their injections with 'offtakes'; in particular, if they operate within their day-ahead *final physical notification* (FPN). If they do this, it will be possible but subject to the level of inter-system transmission shocks that occur within the day, for the SO to operate the network closer to its simulated day-ahead unconstrained schedule. Put another way, because of the network externality that out of balance Generators create, the marginal social benefit derived from operating a balanced NTS is higher than the private benefits that the Generators might earn from running outside their FPNs. Once this efficiency expectation assumption is made, three issues arise that relate to the formulation of economic policies for the efficient balancing of the NTS.

(1) How is access onto the grid allocated and priced? (2) Although the SO plans dispatch on the day-ahead, actual consumption and supply are really uncertain; and (3) given that an SO exists to carry out residual balancing of the grid: how is the cost of system security treated? These issues can enhance market failure; as a result, electricity markets use different policy options to curtail the effects of their threats to the security of the system operations and on the market-derived prices.

In issue (1) mentioned in the above paragraph, what most markets do is to use a competition mechanism such as auctions to allocate transmission access rights in a primary market. They will have a secondary market in place if necessary, by which the Generators' would trade rights that they procure during primary auction rounds. Depending on the nature of the grid (temporary or permanent constraint boundaries), they use either a *Zonal, Nodal or Flowgate* methodology (Hogan, 2000b; Hogan and Harvey, 2000; Hogan, Johnsen, Verma & Wolfram, 2000) to price transmission access. The important point is to ensure that whatever methodology is adopted, can enhance the SOs congestion management (see for example, Green, 1998b). If there were a policy that prevents Generators from hoarding rights, it would curtail the abuse of monopoly power. Most markets pursue incentives that restrain hoarding; one common way is to adopt policies that legalise an agents loss of rights which are not utilised by say a cut off time before 'gate closure'. An example of this is the inclusion of a '*use-it or lose-it*' clause into the market rules.

Issues (2) and (3) relate formulation of economic policies on the NTS but from the perspective of the treatment and initiatives for energy balancing. Policy initiatives that are based on collective energy balancing decisions enhance free riding, which is a known cause of market failure. It might not be socially beneficial to expect that the Grid users would voluntarily balance their individual injections with offtakes. The structure of the electricity industry, plus the associated imperfect attributes means that Generators will always manipulate capacity to earn higher prices. Consequently, policy initiatives that are based on collective decision making, might not be the best way to enhance NTS efficiency (see Cuyler, 1985; Tirole, 1998; Scherer & Ross, 1990). Moreover, the idea of finding the right value discovery for commodity and capacity, through competitive mechanism, means that markets must adopt policies that minimise free riding in regulatory reforms. This suggests to me that imposing a form of taxation on the Grid users may be the only way in which the 'network peace' as well as the internalisation of negative externality will be achieved.

In some residual pools, for example in England and Wales', the commodity buy price, was set *ex-post* and it included system *security costs*[1] (see Electricity Pool, 1997). The demand-side bore this cost in proportion to their 'off-take' throughput. The bids and offers into the pool were not *'firm'*; as a result the Generators' were not directly responsible for the imbalances that they contributed to the network. Since the regime was based on averaging balancing costs, Generators did not really apply due diligence to balance their injections with offtakes.

A cost-targeted arrangement may induce the Generators to strive to maintain balanced positions at the designated 'gate closure' on any system. The knock-on effect of having a better balancing conduct from the Generators, might be a significant reduction in the SOs within the day security costs. To achieve this level of efficiency, electricity markets impose individual responsibility on Generators' to maintain balanced positions by transmitting security costs through the price mechanism. They do this by operating regimes that are based on *'firm'* bids and offers and in which the price rule includes a penal *'cash out*[2]*'* for energy imbalances. In such regimes, out of balance Generators usually earn *'spill'* charges if they inject more power onto the grid than their day ahead FPN specifications; and those that deliver less than their day-ahead notification pay the *'top-up'* prices (See Offer, 1998b).

### *1.2.5 Summary*

[1] The security costs that the SO: NGC, incurred to restore the system to its balanced tolerance limits. It consisted of the costs for procuring transmission (balancing) services: black start, reactive power, reserves (spinning and non-spinning) and frequency response. It also includes costs for resolving constraints: constraining 'on' and 'off' and start-up.

[2] I assume that there is a cut-off time for each trading period, and which is set before real-time. That is, each market will set a time after which traders cannot contract for a particular half-hour (or period, depending on the markets calibrations). This cut-off time is called 'gate closure'. *'Cash out'* is a penalty charge that is levied on a Generator for its energy deviation between its day-ahead notification and what it actually does on that day, calculated at 'gate closure'. Under this regime, a Generator that operates outside its day-ahead physical notification is charged or rebated a system marginal price and a system average charge if it is within its notification.

The last section summarises the interaction between generation and transmission segments in electricity markets. It is a process that is much more complicated than what commodity trading involves in some of the other markets. The features of electricity which limit the application of competition policy into the industry are the same issues that exacerbate the effect of the imperfect nature of the industry plus monopoly power, on prices. They cause arbitrage and add flavour to the traders' ingenuity, which their commercial strategy usually reflects. The traders' opportunistic behaviour are profit maximising strategies, which are allowed in other commodity markets; in electricity, they can lead to sub-optimal situations some of the time. Therefore, they are not particularly acceptable; moreover, as section 2 shows, the Sector Regulator usually keeps close surveillance of the market activities and intervenes if prices are persistently above cost.

### *1.3 Empirical issues: Why deregulate?*

Why do countries choose to deregulate their electricity systems? The reasons are almost always based on a combination of circumstances (Campos & Esfahani, 1996), the most quoted of these are listed below.

- The inefficiency of public enterprises
- Macroeconomic solution to resource waste
- Regional markets: a path to a reduction in the aggregate costs of generation and transportation
- Political ideology
- Directives from the Bretton Woods Institutions

#### *1.3.1 Inefficiency of public enterprises*

One of the main reasons that countries claim to opt for privatisation and deregulation of electricity systems is that the vertically integrated, and usually monopoly utilities, persistently produce electricity inefficiently (Plane, 1992). There is this general notion that compared to the private firms, managers in the public corporations persistently fail to minimise the costs of production, which they can curtail. This behaviour comes from the objectives, incentives and constraints under which the managers of public corporations operate. Public corporations are also seen as extensions of the political party in power; consequently, they remain a means by which the Government grants favours to its supporters.

In relation to their private counterparts, the managers in public corporations also have no incentives to reduce production costs because they are shielded from take-over threats. They are also not subjected to shareholder accountability. The workforce is also heavily overmanned in the public companies. The managers in public corporations tend to receive huge and often open-ended subsidies from the Treasury. Therefore, the managers in the public enterprises do not have a compelling drive to generate their working capital requirement (see Newberry, 2000). Moreover they have low priorities to run the corporations in an economic and profitable manner; it is also common to find that the workforces' goal for engagement excludes all forms of penalty for poor productivity (Domberger & Piggott, 1986).

Since the 1990s, we witnessed many Heads of Government and their policy advisors approach unbundling and deregulation of electricity networks with vigour. They act as though it is a long-term solution to the resource waste experienced by using the public corporations to produce and deliver goods and services. They see privatisation as a substitute to bad management (Bennell, 1997: Suliman & Gebreysus, 2001). This implies that these Governments concede that deregulation is a path towards the right value discovery for capacity and commodity in generation; it can also enhance best practices that might enable the firms in the privatised market to curtail the production costs that they can control.

### *1.3.2 Macroeconomics solution to resource waste*

Most governments' originally provided electricity through publicly owned and vertically integrated electricity utilities; and the Treasury subsidised the daily operations of these firms, in a number of countries. Where there is no formal periodic subsidy, the Treasury will provide 'life lines' to the companies if they experience financial crisis, which threatens the reliability of supply. The Treasury usually uses tax revenue, other income from the economy or borrowings, to augment whatever income the public utilities make from their production, and which they use to meet their working capital requirement. This leads to the case which reformers put forward: they suggest that instead of simply spending money on the public companies, the Governments' can earn significant revenue by selling their shareholding in these corporations. This was the case in England and Wales when it auctioned the 3G licences in 2000. There are also similar success stories recorded in the deregulation of telecommunications in Africa (Shirley; 1992; Gebreab 2002). It seems to me that part of the drive for the rapid moves to deregulate electricity is this record of extra income, which some Governments made from selling telecommunications bandwidth. The implication is that they expect the same type of competition models applied to telecommunications, trucking and the airline industries (see Boreinstein & Bushnell, 2000), to mention but a few, which are perceived to have worked, and can deliver comparable efficiency savings, when applied to electricity. But the evidence is that the opposite has been the case (see for example, Sweeney, 2002; Fisher & Galetoviz, 1998; Kaiser, 2000).

In summary, a number of Governments across the world, expect that privatisation and deregulation will be economically beneficial because of reasons such as (1) the Treasury will earn income, which the Government can re-direct towards the development of other sectors in the economy. (2) By removing the funding of the public utilities from the Treasury's books, there is an added macro benefit that comes from the reduction of the Governments' PSBR (see Plane, 1992). (3) Privatisation can also induce rapid economic growth and development. The Gross Domestic Product (GDP) after the reform will reflect this; compared to Africa and Latin American countries, privatisation was the primary cause of the rapid economic growth in the Asian market (Dollar, 1992).

Regarding point two in the paragraph above, and in practise, I observe that Government might not always earn any significant revenue from the sale of its public utilities. Citing the case in England and Wales, Vickers and Yarrow (1988) note that the huge savings recorded in Britain, is a result of the accounting and valuation methodology that was used for the transfer. This suggests therefore that the earnings in England and Wales' privatisation might have been overstated; and implies that some of the emerging markets that anticipate such gains should treat it with the caution that it deserves (see also Wamukonya, 2003).

### *1.3.3 Regional markets: a path to reduction in the aggregate costs of generation and transportation*

Since the 1990s, we have witnessed the birth of regional markets; the examples include the power integration regime in *Southern African Development Community* (SADC), Scandinavian Nordel / Nord pool, the Western Europe's UCPTE, and before the restructuring of the US electricity market in 1996, the Mid-Continent Area Power Pool (O'Leary, 1998). These regimes were created partly because the participating member states expected that they could earn significant reductions in the AC of generation and transportation of electricity within the region. This can be achieved by integrating the separate utility systems into a larger regional network with the size of the market served by multiple Generators' (Weiner, 1997; Wolf, 2001). This is possible because the natural resources available in the region would be optimised if generation were carried out in the places with greater comparative resource advantage. Subject to the availability of the right transmission infrastructure, it may be possible to wheel electricity from the places of excess supply and lower costs to those of surplus demand and higher cost. I must point out that the reduced AC might not be achieved unless the right generation and transmission infrastructure exists. In particular, the lack of interconnectors will limit efficient power wheeling across the regional market. Lack of the appropriate generation and transmission infrastructure is an important constraint in the development of the power-pooling regime in SADC.

### *1.3.4 Political ideology*

Political ideology and not necessarily the economics of the privatisation projects has been cited as a factor that also drives some countries to privatise and deregulate their electricity industry (Joskow, 2001; Percebois, 1999; Weiner et al, 1997). For instance in the UK, the wave of privatisation started in the 1980s under the Conservative Government of Lady Margaret Thatcher. It was in 1987 that the Conservative Party included its intention to privatise the electricity industry in its manifesto (Bunn, 1994). After the Second World War (1930s and 1940s) most Governments thought that equitable distribution of goods and services was guaranteed if they owned the factors of production. In contrast in the 1980s, it seems that the Conservative Government in Britain felt that the private sector had developed sufficiently. They might also have thought that the institutions, which foster cohesion in free markets were sovereign and matured enough to support the emergence of independent and transparent competition police regimes. Indeed, it appears that they felt that equitable distribution might also be achieved through the private provision of goods and services. In this case, Government can pursue social-orientated goals by including such objectives into franchise licences (see Shelifer, 1998; Laffont & Tirole, 1993). It seems that they considered that privatisation might also be a democratic way to get a significant number of the population to own factors of production; thereby giving them the opportunity to play active roles in the different as well as important sectors in the economy (Bos, 1994).

It appears that the success in Britain has influenced the political ideology in other parts of the world. In the past decade, the UK has provided a 'world laboratory' in teaching how the less advanced nations can turn their state-owned inefficient public companies into efficient private firms (The Economist October 13, 2001).

### *1.3.5 Bretton Woods Institutions: behind the shadows*

Governments in some of the developing countries cannot afford the huge finance that they require to develop their electricity sectors; therefore, in such economies, privatisation is seen as a way to overcome organisational inertia (Wilson, 2001) and to attract private and foreign equity into electricity network capacity building.

Some developing countries unbundle and deregulate their electricity systems based on the directives from the Bretton Woods Institutions (Plane, 1997; Stiglitz, 2002). It is possible that the perception that the World donor agencies have about the success, which the Lady Margaret Thatcher's Conservative Government made in the UK by liberalising markets such as its foreign exchange and some of the earlier utility privatisation's (Telecommunications and in natural gas) in the 1980s, influenced their sectoral lending policies. As a result, in the past 15 years, these banks have worked closely with the Heads of Government in places like the sub-Saharan Africa (SSA), to steer forward projects to substitute the public for private ownership of electricity networks. They induce competition policy regimes directly under different forms of austerity and / or *structural adjustment programmes (SAP)*; and indirectly through dialogues. Either way, the institutions focus on selling to these governments the unequivocal guarantee that liberalisation of markets is the best option for them to reduce both their PSBR and huge accrued international debts ( see Gibbon et al, 1992; Harrigan et al, 1991; Due, 1993).

The *Less Developed Countries* (LDCs) particularly those that fall within the World Banks *Heavily Indebted Poor Countries* (HIPCs) initiative in SSA are the worst hit. Compared to their counterparts in Latin America and Asia, they depend more on the funding from these donor and aid agencies to survive (see Plane, 1997). But section 4 in this research conjectures that competition policy will be a high transaction cost, which would yield little benefits to some of these countries in SSA. Nevertheless, their Heads of Government apparently have absolutely no choice but to implement the World Banks SAP because they depend highly on the international agencies to survive. For example, the Economist Magazine (1 – 7 June 2002) reports the Zambian experience with seeking funding; 'donors insisted that aid to Zambia worth around $1 billion a year in the mid-1990s, be made conditional on governments privatising the mines, which it had owned and run since 1970. In 2000, a debt package of $3.8 billion was promised to ease the sale, making Zambia among the biggest recipient of official aid in sub-Saharan Africa' (page 83).

What is the basis for the World Banks deregulation-based lending policy? Stiglitz (2002) reveals that this aspect of the bank policy is not based on the economics of these projects. He goes further to say that 'the IMF and the World Bank [have become] the new missionary institutions, through which ideas [are] pushed on the reluctant poor countries that often badly [need] their loans and grants (page 13). He also expresses concern that during his period at the World Bank, he observed that 'decisions were made on the basis of what seemed a curious blend of ideology and bad economics, dogma that sometimes seemed to be thinly veiling [on] special interest...' (page xiii). It appears that the wrong foundation for these privatisation drives by the World Bank is the reason why most of the deregulation projects that were carried out in Sub-Saharan Africa since the 1980s have persistently failed (see for example, Sulaiman & Ghebreysus, 2001; Bennell, 1997).

### *1.3.6 Summary*

The neo-classical free markets mechanism and agency theories underpin the reasons on which some Governments unbundle and deregulate electricity systems. Outside the academic economics circle, politicians expect that competition policy and price mechanism will enhance production and allocative efficiency. They also expect that the significant reduction in the role of the State will address part of the macroeconomic problems of resource allocation. This suggests a question about the role of government. The History of Economic Thought extensively documents discussions about the appropriate role of the State, begun long before Keynes and the 1930s Depression (See for example, Adam Smith, 1776, 1994 & 2000; Sunderland (eds.), 1998; Galbraith, 1988). Some of the issues raised includes whether Government or the private sector should provide 'essential services' such as education, electricity, water, housing, refuse collection and sewerage services; and what proportion of the service charge should be borne by the citizens.

## 1.4 Does ownership of a firm determine managers' productivity? An overview of empirical findings

This research examines the outcome in an electricity market whose regulatory reform was based on the inefficiency of public enterprises and the ideology that greater productivity can be achieved through the private provision of electricity. It is therefore necessary to review what exists in the literature on the comparative efficiency of private and public enterprises.

The debate about the relative effectiveness of private rather than public ownership of the factors of production dates back to Adam Smith (1776). Smith upholds there is greater productivity when factors of production are transferred into private hands (see Sunderland (eds.), 1998). In the past 20 years, although comparative empirical studies provide conclusive evidence in some cases, the truth is that economists are divided about the effect of ownership on the performance of firms'. Some studies report better performance by the private than the publicly owned firms do (see for example Crain & Zardkoohi, 1978); others find the opposite to be the case. There are others who rule out any relationship between ownership and performance (see Frank & Mayers, 1997). The latter school of thought argues that ownership structure is not an issue; instead that it is the absence of competition, which limits managers from using best practices to minimise costs (see for example Shelifer, 1998).

Most comparative efficiency studies on the performance of private, public and mixed ownership firms have been carried out in the United States of America (see Parker & Raffiee, 1994; Crain & Zardkoohi, 1978; Boyer (reviewer), 1995; Bowen, Moncur & Pollock, 1991). Relative to electricity, there are sufficient numbers of observations when the water and sewerage firms are divided into private or public firms; the technology used is more straightforward, concepts easier to understand and the output measures are very easily determined. Crain & Zardkoohi (1978) examine the behaviour and calculate the percentage difference between the observed and optimum cost minimisation by the public and the regulated privately owned water utilities. They found private and regulated firms more efficient than their public counterparts. In contrast, Bhattacharyya et al (1994) assumes that an unobserved shadow price reflects the regulatory environment in the water industry. They tested for cost minimisation by deriving shadow prices as functions of market prices. And found that the efficiency between the private and public utilities is more widely depressed between the best and the worst practices; but on average, they report that the publicly owned are more efficient than the private water utilities. Kamshad (1992) found that the growth of French co-operatives and their American counterparts, as well as their respective survival relationships were independent of firms' ownership structure (page 2). Similarly, Frank & Mayers (1997) did not find evidence of ownership contributing to poor performance by some firms in Germany. In similar research, Whitley et al (1995) reported no significant difference in the management and employment policies between the private and publicly owned firms in Hungary.

The above studies consider performance in isolation from any form of competition policy, whether between *business units* (BU) in the same organisation or one, which the operating environment induces. They also ignore the impact of incentives on managers' ability to improve efficiency. There are studies that uphold the positive impact of these on managers' productivity. Ruggers and Leslie (1991) found managers' were improving efficiency in the Hungarian public utilities when they were subjected to greater shareholders accountability. There are also reports that the threats of an imminent change in regime induced managers' to improve their efficiency (see Pinto & Van Wijnbergen, 1995). In section 3, this thesis finds that the system marginal price (SMP) which is a near proxy to the Generators commercial strategies being lowest in 2000, was the run up to the implementation of the New Electricity Trading Arrangements (NETA). This result suggests that the threat of an imminent change in the trading arrangements had a positive effect on the Generators bidding behaviour (see OFGEM, 2001). Referring to Yarrow (1989), Pollitt (1995) reports that British Steel achieved its largest productivity gains towards the run up to the industry's privatisation. These positive effects of the threats of a potential change in regime on managers' performance may be because they expected that the good reputation of productivity that they made before the change would guarantee that they were employed in the reformed industry.

Estrin & Perotin (1991) found that ownership does not matter; instead it is the environment within which a firm operates that determines innovation and best practices. Shelifer (1998) upholds the same view; he contends that ownership does not drive productivity; and concludes that firms' perform better in competitive environments. He goes further to reiterate that Government does not need to own factors of production to foster social goals; instead, they can include enforceable welfare-orientated clauses in franchise or contracting licences (see Viscusi et al, 1995, for discussions on franchising).

On average, I find that a number of the studies that I reviewed and which propose the efficiency enhancing and innovative attributes of competition policy paid little attention to its limited effectiveness if applied in some highly concentrated industries. They do not consider that it might not yield significant efficiency gains in cases where the rudimentary development or the sub-additive nature of production means that it is economic to have a natural monopoly responsible for production and distribution of goods and services (see Tirole, 1998). This is the situation in the electricity industry; therefore, it raises the question of whether the peculiar industry structure can inhibit the development of the desired contestable capacity and commodity regimes that are expected post-deregulation. Put another way, is ownership of electricity utility firms a primary driver for efficiency gains which can be passed on to the final consumers in the form of reduced prices for services? In the next section, I turn to the empirical evidence from some electricity markets?

### *1.4.2 Competition policy and network utilities*

Like the ownership-performance discourse, there is also great controversy between economists as to whether competition policy is the primary driver of managers' performance in all types of industries. I do not think that this lack of consensus is evidence that political affiliation is an influence on their views. Instead I believe that because people are prone to seeing the world from their own prism, it would be normal to see that an individuals area of specialisation would be a strong influence on their ideology. For example, I would not be surprised to find that a pure macroeconomist would be favourably disposed to privatisation and deregulation. He / she is likely to argue that it would help to address resource waste. In this case, the macroeconomist is possibly by omission isolating all the other issues in the economy and considers that the only relevant problem requires the optimal allocation of scarce resources.

On the other hand, another extreme approach, may be the view by a pure industrial economist and an anti-neo-classical theorist who might oppose privatisation bills. Here the case might be made about the sub-additive nature of transmission and distribution; the lack of contestability in generation and supply and the fact that Generators would always behave opportunistically. It is likely that based on these industry imperfections, this opponent might argue that *laissez faire* might lead to inefficiency in electricity. His / her additional support might be along the lines that marginal cost (MC) pricing will be difficult, if not impossible, to achieve and sustain even with a lot of regulatory intervention. Monopoly profit might lead to over capacity, which although might be good in some commodity markets, in utility networks, is socially inefficient. Whose viewpoints matter and to which side can one attribute the evidence from some of the electricity markets? In essence, which economist should we believe? Next, I review the experience recorded so far in some of the electricity markets.

England and Wales achieved most of the objectives with which it set out to deregulate its electricity by the beginning of the year 2001; it earned significant efficiency savings and relative to the period before privatisation, the average and real electricity prices had decreased (Littlechild, 2001). Nonetheless, pool prices failed to reflect the full changes in the industry structure, capacity mix and the reduced costs of input. That is, despite the huge regulatory input that went into the reform combined with the Regulators daily market surveillance, the price mechanism was inefficient. Newberry and Pollitt (1997) also found that the large industrial consumers benefited more from the industry's privatisation than their retail counterparts. Nonetheless, the success of the regime, particularly given that the lights remained on in England and Wales (Green, 1999) and the system met growth in demand, throughout the 1990s, suggests that competition policy can lead to significant efficiency savings in capacity constrained network utilities.

In unpublished research, Rochino (2001) finds that the lack of the economies of scope leads to higher transportation costs in the privatised electricity markets in 19 OECD countries. She also finds that the vertically integrated electricity firms can improve efficiency by stimulating competition between business units. This suggests that ownership does not matter in electricity, since both the private and the public electricity utilities can do this. This is similar to what Newberry (1995b) finds when he compared the experience between the pre and post-deregulated regimes in Europe, Scandinavia, Argentina and Chile. He reports the regulatory regime and not the ownership structure of the utilities affected the managers' performance. This suggests that the environment within which an electricity utility operates can induce managers to use best practises in production. I conjectured from this finding that ownership may not matter in electricity, instead that efficient public policy is what is required for both the private and public electricity utilities to improve efficiency. Nonetheless, Newbery

(1995b) recommends that where feasible when economies deregulate electricity systems, the monopoly transportation business should be left under public ownership. Indeed, public ownership may lead to lower costs of production particularly where 'user-oriented tasks and services' have attributes, which may be difficult to specify contractually (Kwoka, 2002).

Most of the considerations above are both relevant and applicable in some of the mature economies in Europe and the USA. The dimensions change when consideration is given to the less developed countries (LDCs). This research shows in section 4 that the case for private provision and deregulation of electricity generation is ambiguous. If political alliance and endemic corruption, which are barriers to private and foreign investment into capacity building in a number of the developing countries in places like SSA, are assumed, ownership will not matter. This is because both the public and private firms can be very inefficient (Sulaiman & Ghebreysus, 2001).

I think of other constraints when I consider some other institutional ingredients that enhanced the success of the regulatory reform in Britain. One is the existence and influence of a mature capital market on the managers' productivity. It appears that the pressure from the interaction between the commodity and the capital market is important to sustain the impact of the threat of take-overs on managers' efficiency. This cause and effect relationship is easily seen in the mature economies where information about companies is incorporated instantly into their share prices on the Stock Exchange. One of the benefits of competition when it occurs in mature economies and where the institutions that enhance cohesion between the agents in the system are independent and sovereign, is that the capital market usually restores the inefficiency created in the product market (Domberger & Piggott, 1986).

My experience from working as a banker in Lagos, Nigeria and which is applicable to most of the other developing countries in SSA is that less than 40% of the cash in circulation in the economy was routed through the banking sector. Less than 10% of the first Municipal Bond that my bank structured for the construction of a market complex for the Lagos Island Local Government (LILG) was subscribed. This is because the population conducts a very limited amount of business transactions through the finance sector; and they do not have a culture of long-term savings. Apart from the Stock Exchanges in Ghana and Lagos, the rest of the countries in West Africa do not have well developed stock markets. This means that there will be a very limited interaction between the commodity and capital market post-deregulation. It therefore shows that these economies might in fact have no social welfare gain by substituting public with private provision of electricity. Moreover and usually, the assets of the utilities that are earmarked for privatisation are virtually non-operating and obsolete; significant numbers of the staff would have either left the company or those that remain, would not have the expertise to steer forward sustainable operations. This is my personal experience with the electricity sector in Nigeria, where there is a vertically integrated monopoly and publicly owned electricity utility, the National Electric Power Authority (NEPA). Entry for example into the Nigeria power sector requires that a foreign firm invests huge sunk costs to resuscitate the utilities, in fact it will need to re-Grid the network. The socio-political risks involved, combined with the level of endemic corruption are just part of the reason why foreign firms are not 'very keen' to enter into the market or even to buy out NEPA.

The arguments about the efficiency of public electricity companies will continue. On a general basis, the same issues that concerned Chadwick and Mill in their debates in the 1850s over nationalising certain British Industries and to Von Mises and Hayek versus Lange and Lerner in their well-known exchanges about the efficiency of centrally planned economy also continue to dominate literature on the (in)efficiency of public electricity utilities (Crain & Zardkoohi 1978: 395). On the other hand, most Governments seem to have accepted that unbundling and deregulation is the only way to improve the efficiency of electricity utilities. Nonetheless, we have seen that since the 1990s, Government has increased the use of the Sector Regulators in monitoring the daily market operations in residual pools. The objective is to protect the interest of the final consumers, which they expect to achieve by ensuring that they are not charged *unfair* prices for their electricity consumption. Sections 2 and 3 show that this type of monitoring which involves price caps, forced divestment of plants and strongly worded communications with Generators, occurred throughout the pool regime. An additional concern is seeing that the supply is reliable and it is done over safe and secure Grids. Electricity markets also use competition (anti-trust in the USA) laws, which are usually stipulated in fair-trading acts to set the guidelines about agreements between undertakings. The reliance on public policies and not market forces means that price mechanism does not determine the structure, conduct and performance in deregulated electricity systems.

Government also intervenes in electricity markets to curtail the operation costs to the taxpayer that internalises all the costs including over-runs and market failures from regulatory reforms.

Since the rise of electricity deregulation, we have seen private utilities announce and internalise huge profits. But they do not bear fully the associated costs of market failure. Instead, what happens is that the Government always steps in to provide financial 'life lines' to these firms that are operating in 'competitive environments', once exposed to situations that threaten the reliability of supply. The implication is that the costs of market failure are socialised to the taxpayer (see Wamukonya, 2003). This is what happened with Rail Track in Britain (see for example Economist, October 13, 2001). The company was making losses; was not in any financial position to make appropriate network investments and its continued operation was clearly against the public interest. This is where the question of who bears the cost of network investments became quite relevant. Government had made several lump sum payments to Rail Track, which did not help its situation. What was quite interesting was that the shareholders expected to continue to earn returns on their investment. That is, they looked to the Government to subsidise both the company's working capital and their dividends.

Some of the current issues that dominate industry forums in England and Wales electricity industry include seeking answers to questions such as whether the shareholders, final consumers or taxpayers should bear the costs of acts of god? Should the Government continue to subsidise the ISOs network investments in the privatised regime; that is, should the burden of making appropriate *long term investments* (LTI) be placed solely on the transmission owner?

The issues, which I mention in the paragraph above underlined most of the initiatives that the UK's Office of Gas and Electricity Markets (OFGEM) steered forward towards the run up to the implementation of the New Electricity Trading Arrangements (NETA) in 2000. The worrying thing is that the solutions that policy advisors give to some of these problems that are potential threats to the system operations keep changing. They therefore make it very difficult to rely on existing markets for the knowledge about the right policies that might minimise market failures. On the other hand, in section 2 this research finds that the UK put in huge and expensive regulatory input to achieve a smooth transition to retail competition; it also reveals that the independent and mature institutions supported the success of the regulatory reform. Nonetheless, the advocates of electricity deregulation play down the fact that the regulatory reform-related costs as well as those that the Government incurs to 'police' daily market operations are quite high. Moreover, these are necessary expenditures required to facilitate part of the objectives of any competition reform.

### 1.5 *Strategy, price and capacity determination in electricity markets*

This section provides an insight into the way that I expect the Generators to behave in electricity markets. My intention is for this sub-section to give the reader some intuition and background on some of the points of argument that I make in most of section 3.

Industrial organisation (IO), game and auction theories provide intuitions behind traders' conduct in electricity markets. The interaction between the traders fits properly and is best described as a *'supergame'*, which Friedman (1971) defines as 'playing of an infinite sequence of 'ordinary games over time' (page 1). Traders engage in repeated interactions to procure capacity, knowing that they will meet again in the future. Their strategies are simply directed at maximising profits; since market prices cannot clear aggregate supplies, Generators' strategically determine price as well as capacities against profit function payoffs (see Brock & Schienkman, 1985; Bolle, 1993; Bower & Bunn, 2000). Section 3 of this project finds that capacity is the main tool, which Generators used to manipulate prices in the England and Wales' pool (see also Bolle, 1992). This is consistent with the theory of suppliers behaviour in oligopoly markets (see Tirole, 1998). In theory agents in a capacity constrained industry will also use tacit collusion to keep prices above competitive levels (see Tirole, 1998). Coding and signalling are some well known practices that agents have used in other capacity constrained industries where auctions, a mechanism that involves the repeated interactions between participants', were used to allocate resource (see for example Porter and Zona, 1993; Schmalensee, 1979; McMillan, 1991; Crampton & Schwartz, 2000).

In electricity, Generators can exhibit tacit collusion by submission of mutually reinforcing offers into the residual pool (see for example Wolfram, 1998). They may also rely on using capacity to manipulate within the day balancing costs (see section 3 in this research). Such anti-competitive practices are easy to carry out because with repeated interaction Generators' build expectations about competitors potential strategies. The information in the public domain also enables the Generators' to formulate opportunistic behaviours, which might help them to maximise their marginal private benefits. Some of this information includes: *Generators Registered Capacity* (GRC); some elements of the cost structure of competitors, fuel type and plant size as well as the variables that go into the SOs forecast of the gross demand. Moreover, job mobility between firms also enhances knowledge transfer about the competitors potential strategies.

Price and capacity setting in electricity markets is analogous to the Kreps and Schienkmans (1983) two-stage capacity constrained and non co-operative oligopoly game model. First, the Generators' build plants before registering their trading capacities; thereafter they decide the capacity to bring to the market. This happens in Stage 1 of the game, which is where quantity setting is similar to 'Cournot' competition.

The market is Stage 2 of the game; and where the realisation of demand as well as price competition occurs. The allocation of the Grid capacity is done in a 'Betrand' fashion and it is based on the Generators offer prices, but with the proviso that a trader cannot change the Cournot quantity that it brings to the market from stage 1. Since capacity brought to the market is binding, and price does not clear aggregate supply, the important and profit maximising requirement is for a trader to learn the market rules quickly and to acquire the skills with which to maximise profits. This is achieved once the trader is able to estimate the industry's aggregate load. Once it is achieved, traders can determine the residual demand, which his/her firm will supply to the SO with some accuracy, and offer such residual and inframarginal capacities at prices that exceed its avoidable costs.

The traders' face conflicting objectives when they arrive in the market, which derives from the desire to maximise profits and to be in-merit. Whilst their commercial strategies might satisfy their marginal private benefits, it could lead to system sub-optimal situations some of the time. The general rule though is that prices in oligopoly markets will always be lower than monopoly levels but be higher than the Betrand equilibrium (Tirole, 1998). In electricity, the mixed strategies that Generators adopt when they offer capacities between the load regimes leads to the aggregate costs of generation being above competitive levels (see Fehr and Harbord, 1993). This raises questions about the social benefit of using competition mechanism to allocate generation capacity in network utilities.

Using competition mechanisms, such as auctions, to allocate scarce resources is considered a success in some of the other short run capacity constrained networks. The Internet, business to business e-commerce, transport, railway capacity franchises, construction, telecommunications, natural gas, finance and investment are examples of some of the sectors that have used it to allocate resources. One argument made in the literature is that auctions helps to reveal information about the private agents' valuations of goods to a buyer; thereby enhancing its right value discovery. However, it is thought appropriate that the structure of the industry is such that concentration and rudimentary development limits efficient resource allocation (see Aron, 1998).

Nonetheless, using auctions to allocate commodity and capacity in network utilities such as energy markets has led to inefficient outcomes. This was the case when prices rose by over 1000% during the second six-monthly allocation (March 2000) of entry capacity through the beach[3] terminals into the England and Wales' natural gas system (see OFGEM, 2000). In section 3 of this research, I find that the system marginal price (SMP) in England and Wales' pool failed to reflect demand and supply situations. A number of the earlier empirical studies on production and allocative efficiency in the pool, which I reviewed in section 3 part 3.2, including the Regulators' price inquiry (for example OFFER, 1991; OFFER, 1994; OFFER, 1999) attribute some of the shortcomings in the regime to the design inefficiencies of the auctions. However, I find that a robust design which is combined with the right implementation methodology may curtail inefficient outcomes, when network utilities use auctions to allocate scarce resources (see Klemperer, 2001, Klemperer, 2000b).

### *1.5.2 Competition Law: Application to electricity market*

Public policy is the main determinant of market performance in deregulated electricity regimes. This section provides a broad summary of the foundations for the implementation of competition (anti-trust) law into electricity markets.

In electricity markets, the guidelines and procedure for agreements between undertakings are usually spelt out in competition and fair-trading acts, both of which provide the codes of

[3] St. Fergus, Thedlethorpe, Bacton, Barrow, Easington, Teeside, are the six beach terminals that are linked into the natural gas transmission system that Transco owns and operates. Shippers land gas from the off-shore fields like the North Sea at the terminals and depending on their contracts with Transco, it will transport the gas through the transmission lines down to the local distribution zones (LDZs).

conduct between utility firms. Where regional markets have emerged and a centralised Regulatory Body exists, it will usually set out the guidelines for fair-trading across member states within Directives. The member states' would often ensure that they harmonise internal competition and fair-trading act(s) with those set out by the regional regulatory agency.

Most electricity markets that emerged after 1990 adopted the England and Wales' licence-based regulatory reform. Articles 85—collusive practices that inhibit competition—and 86—abuse of dominant position—guides the organisation and flow of goods within the common European market (Kora, 1997). It also influences the fair-trading acts and competition policies in Britain, where the Chapter 1 and 2 prohibitions integrate the provisions of articles 85 and 86 to specify the guidelines that undertakings should follow when entering into agreements. It is also common to find that these types of documents will specify concerted practices that are considered to restrict competition within the internal market. Since each utility has its inherent features, restrictive practices are interpreted and tailored to meet the technical and structural arrangements of each network. In the energy sector, The Competition Act 1998 includes the consideration of *'pre-emptive[4]'* behaviour when the Sector Regulator investigates anti-competitive cases (see OFGEM, March 2001). Despite the distinct features of each utility, the criteria for defining a relevant market and assessing dominance are consistent with the Commissions' application in the regional market.

The guidelines for and the evaluation of anti-competitive practices involve the definition of the relevant market plus an assessment of the dominance of a firm within that geographic market. Definition of the relevant geographic market includes consideration of some of the following: the features of the product, its use, temporal characteristics, homogenous conditions of competition between member states and the effect of the economic powers of the firm. An important aspect of the assessment are the barriers that can prevent similar and potentially efficient firms from entering the market because that is what will limit the development of efficient competition (Kora, 1997).

[4] 'Pre-emptive behaviour' describes abuses by dominant incumbents in markets that are being opened to competition or are newly opened to competition, which are designed to adversely affect the development of competition' (OFT, March 2001:10).

In United Brands Co and United Brands Continental BV v. Commission, the Commission considered UBCs structural network and its competitive advantages within the relevant market. It defines dominance as:

> 'a position of economic strength enjoyed by an undertaking, which enables it to prevent effective competition being maintained [in] the relevant market by giving it the power to behave to an appreciable extent independently of its competitors, customers and ultimately of its consumers' (Bergeron & Kallaugher, undated: 47:65).

The Commission uses similar criteria to define the relevant market and to assess dominance in Hoffman-LaRoche and Co. AG v. Commission; Hugin Kassaregister AB and Hugin Cash Registers Ltd. v. Commission; NV Nederlandsche Banden-Industrie-Michelin v. Commission; British Broadcasting Corporation & BBC Enterprises Ltd v. Commission and in European Court of Justice (ECJ), in Radio Telefis (RTE) and Independent television Publications Ltd (ITP) v. Commission (see Bergeron & Kallaugher, undated).

In OFGEM v. British Energy and AES (2001), the Competition Commission (CC)—formerly the Monopolies and Mergers Commission (MMC)—considers the demand-side, temporal aspects, generation issues (especially flexibility of operation of gensets) and institutional / regulatory arrangements to define the relevant electricity market. CC acknowledges that there is no alternative to electricity that is available to the final consumers; it upholds that the relevant market is 'electricity from whatever source'. It acknowledges that technical inflexibility limits switching demand from peak to off-peak; and that there is very limited opportunity for the Generators to compete between peak-demand and base-load plants. Therefore, it upholds that each load regime is a separate market. This was a familiar argument in an earlier investigation, which MMC carried out in 1996 into the—then—proposed mergers between National Power plc and Southern Electric Plc and PowerGen and Midlands Electricity Plc (see MMC, 1996a; MMC, 1996b). The implications of this in this current research and particularly in Section 3 is that where possible they should treat each period, load regime and table indicators as mutually exclusive markets. This is also the basis on which this research conducts structural estimations.

## 1.6 *Research Questions*

This sub-section starts with the justification of the broad research questions that I deal with in the rest of this thesis. Before the late 1980s and early 1990s, there was relatively limited data and history on the performance of some of the deregulated electricity systems, which could form a firm foundation for empirical analysis. As a result, economists, engineers and industry analysts started using game theory and simulation as a tool to build strategic models, which represented real-life electricity systems in *laboratory* experiments (see Gary & Larsen, 1998). These simulations are usually micro worlds of what real markets ought to be; therefore they approximate the capacity mix and strategic behaviour of traders (Nicolaisen et. al, 2001).

Most of these experiments are directed at (i) observing the interaction(s) between the multiple agents in a system. (ii) Learning how such interactions affect market-derived prices and (iii) gaining an understanding of the impact of transmission constraints and monopoly on market-derived prices. They, therefore provide organisations with the opportunity 'to learn without experience' (Lomi & Larsen 1999). This also gives policy advisors the basis on which to formulate best economic, technical and engineering policies for the efficient operation and regulation of real electricity markets.

Simulations are usually conventionally designed with multiple profit maximising agents that are free to enter into bilateral contracts to procure the capacity for injecting and 'off-taking' power from locations across a hypothetical grid. The participants (players) are allowed to contract above their FPN and there will be an imbalance settlement set up to deal with energy deviations. Bunn et al (1993), use optimisation and system dynamics to simulate the effects of privatisation on long-term investments (LTI) in an electricity system. Gary & Larsen (1998) use system dynamics modelling to examine the effect of the decisions that the companies make regarding investment in gas-fired plants in the evolution of the energy trading regimes in England and Wales. These two studies predict in different ways that fragmentation will lead IPPs to invest massively in gas-fired generation during the 1990s; this would lead to excessive capacity on a relatively unconstrained Grid. In theory and based on the level of excess capacity that they anticipate, pool prices ought to follow a constantly decreasing pattern. Nonetheless, they found that the price rule creates the incentive for Generators' to manipulate capacity so as to earn higher rents; therefore they concluded that prices will not decrease in proportion to that which would be comparable to the capacity increase (Bunn, Dyner & Larsen, 1998).

Bower and Bunn (2000) examine production and allocative efficiency in the British electricity market if Generators trade outside the pool arrangement. They find that under the discriminatory pay-as-bid (PAB) rule that Generators' will segment the market into load regimes and the periods of low demand, they would offer higher prices for their residual peak capacities. They conclude that prices would be more volatile under PAB than if the Generators earn uniform system marginal price (SMP). In addition, trading outside the pool, which is based on a PAB, residual-balancing mechanism (BM) would deliver inefficiencies worse than what occurred under the uniform SMP pool regime.

Over twenty years since Chile pioneered wholesale trading, there is now sufficient data and history from a number of markets, which show that simulation is no longer the only way for policy analysts to make inferences about efficiency levels in deregulated markets. It is also possible to use case studies to formulate theories about the expectations for emerging markets. This research seeks answers to the following questions:

- Is there a relationship between institutions and successful regulatory reform?

- What is the impact of market structure, design and price rule on production and allocative efficiency in an electricity market?

- Given the context of developing countries, what is the best way in which they can introduce competition policy into their electricity systems. Can the experience from Europe be used to prescribe a path for the successful implementation of unbundling and deregulation of electricity systems?

This project uses the England and Wales' electricity privatisation and the regional power integration regime in the Southern African Development Community (SADC) and sub-Saharan Africa (SSA), as case studies to answer these three broad questions.

## 1.7 Research Framework

The research follows Edward Mason's 1930s conceptual model of industrial organisation represented in figure 1.4 (Scherer & Ross, 1990). The figure shows that some of the factors that affect supply and demand are the basic conditions in any industry. For example, improvement in production technology can cause an outward shift in the supply curve, meaning that more goods would be supplied at a given price than what occurred before the innovation. An example of this type of innovation is the introduction of *combined-cycle gas turbine* (CCGT). In relation to the larger sized thermal plants, CCGTs are more thermally efficient, smaller and require less capital; they also take less time to construct. The loss of load probability on a system will reduce if it uses more thermally efficient and operationally flexible plants to meet demand.

The figure also shows that the consumers' ability to respond to changes in price is an important basic condition in a market, because it determines the quantity of the goods or services that customers will continue to buy if there is a marginal change in price. Therefore, firms have to assess policies to increase product costs from the perspective of the final consumer as well as the substitutes that they have available.

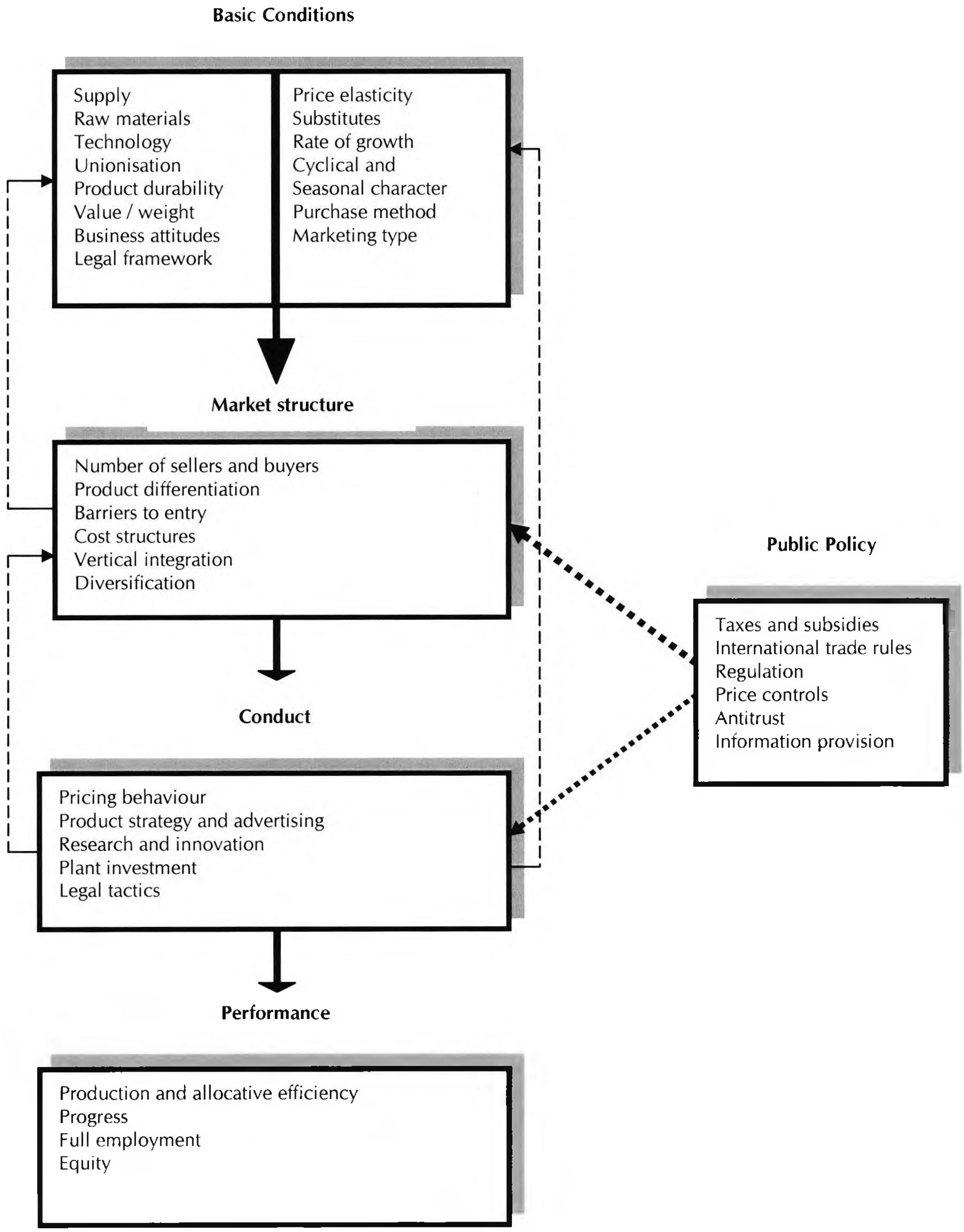
The basic structure feeds into the market structure. This is where the number of sellers and the ease of entry or exit define the level of competition that may exist in the market; it is also at this stage that the influence of concentration on firms strategies are felt. In the electricity industry, the market-derived price does not clear aggregate supplies; therefore the strategies that Generators' adopt when they compete to be in-merit reflects their desire to get called as well as to maximise profits. This market structures and conduct feed into the industry's performance, which then determines the way that price patterns, emerge.

The feedback arrows in the diagram show that all markets do not follow the structure-conduct-performance strictly. Once the issues that create imperfections, such as the nature of a good (being a merit good) and the industry structure (example a network utility) are assumed, public policy will become the primary determinant of the structure, conduct and performance of the market. An example in electricity is where regulatory initiatives such as forced divestment of the capacity of a dominant Generator alter the market concentration. Once the number of the Generators that own the marginal plants increase, greater competition to be in-merit will influence the commercial strategies and could lead to prices reflecting avoidable costs. Similarly, limit pricing will restrain how high Generators can offer their capacities and the fair-trading acts will influence horizontal and vertical mergers as well as the trading conditions between undertakings (see Kora, 1997; Whish, 1993). This was the case in the England and Wales' pool where the Regulators close surveillance of the market operations, helped to curtail higher prices than what was observed. The forced divestment of National Power (NP) and PowerGen (PG) mid-merit plants limited their control of the industry's residual demand; and the CC guidelines meant that the mergers were perceived to be against the public interest and were not allowed (see MMC, 1996a & b).

The above argument shows that there is a weak link in the structure-conduct-performance paradigm represented in figure 1.4. It is a standard paradigm, which seems biased towards the neo-classical theory of perfect competition (See Culyer, 1985; Sloman, 2000) since it makes sense to expect that in such regimes performance will follow through from the structure of the industry and the conduct of the agents in the system. It is also in such markets that the price mechanism can, in theory, lead to production and allocative efficiency; in short that Pareto-optimal results are achieved (see Tirole, 1998). This paper shows that public policy determines market outcomes in electricity markets; therefore, Mason's flow path is weak. Nonetheless, this research uses that framework, and covers all aspects of the flow because it permits a critical evaluation of the market and the derivation of policies that may improve market outcomes.

**Figure 1.4**

**The Structure-Conduct-Performance Paradigm**




Source: Scherer & Ross (1990:5)

## 1.8 Structure of the research project

This research focuses on a generation segment; it investigates all aspects of the interaction of public policy and the influence of market structure on the conduct of Generators; and their effect on market outcomes in the England and Wales' pool. It also evaluates the prospects for the successful privatisation and deregulation of electricity systems in SSA.

It is divided into sections; and structured as follows: section 2 investigates the effect of regulation on the evolution of the England and Wales' pool regime. It focuses on the regulatory input that went into the different stages in the evolution of the privatised industry. It finds that the regulatory reform was not planned in a hastily convened manner, moreover, that the sovereign institutions that influenced the success of the regulatory reform in the UK were mature before the 1980s. The findings suggest that it is important to have appropriate governance institutions, which would support the emergence as well as the evolution of a market reform. As a result, I conclude that it might be better for some of the emerging markets that do not have a mature institutional framework, but wish to improve the efficiency of their vertically integrated and monopoly electricity utilities, to use other efficiency benchmarking techniques to stimulate competition in electricity.

Section 3 examines the market conduct in the pool, from the perspective of its price trend. It consists of five related parts that use quantitative techniques to investigate patterns in the components of the pool selling price (PSP). Collectively, the section finds that the rule for setting prices, the initial policy at vesting that led to the creation of a duopoly market that the National Power (NP) and PowerGen (PG) dominated, and the Grid Codes definition of operating plant availability, exercebated anti-competitive practices. This finding is consistent with some of the results from earlier studies that Fehr and Harbord (1993), and Wolak and Patrick (2001) carried out.

Section 4 investigates how the interaction between market structure, corruption, socio-politics and economics, might inhibit the success of competition policy in sub-Saharan Africa (SSA). It finds that the inherent features of the region will remain a barrier to foreign investment post-privatisation. This is consistent with the experience in the Republic of Armenia where similar issues inhibited the success of the electricity reform (see Kaiser, 2000).

Section 5 concludes by summarising the contributions that I have made to the literature on deregulated electricity markets; discusses the areas for further research that the study has opened and concludes the thesis.

### *1.9 Summary*

This chapter has provided an introductory overview of the technical and industry issues that electricity generation and its interaction with the NTS entails. It shows that generation is not contestable, and provides an insight into the difficulty encountered by treating electricity as a 'tradable' product. This section also reveals that there is no conclusive evidence on the impact of ownership on the performance of managers. I also tried to show why some academic industrial economists argue that the neo-classical model of laissez-faire plus contestability theory will not lead to significant efficiency gains when applied to electricity generation. I introduced the way that Generators might behave in electricity markets, which opened up the summary discussion of the need to use public policy to direct the performance of electricity markets; and through competition law, to control the agreements between undertakings. The framework that I use in the research is Edward Mason's 1930s conceptual model of industrial organisation. Although it has some inherent weaknesses, I limit my ambition in its use to the critical evaluation that it enables me to carry out this research.

# SECTION 2

## Regulatory Reform and Restructuring of an Electricity Industry: An economic regulation and competition perspective

*Abstract*

*Unbundling and deregulation of electricity systems gained prominence in the past 20 years. However, since 1990, the England and Wales' licensed-based regime has influenced the reforms, which many other countries have implemented. I review some of the economic regulation and competition issues that facilitated a smooth transition to full retail competition under the pool era in Britain. I find that its regulatory oversight approximated price mechanism; therefore, it changed as the market evolved. I also find that the existence of mature and appropriate institutions including contractual arrangements enhanced the evolutionary regulatory oversight. This study highlights the implication for emerging markets; and concludes that they must have the right structures, which includes planned regulatory reforms as well as mature and sovereign institutions, for their competition policy regimes to succeed.*

*Key words: Competition, Capacity Payment (CP), Deregulation, Electricity, England and Wales' Pool, Regulation, Pool purchase price (PPP), Pool selling price (PSP), System Marginal Price (SMP).*

## Introduction

This section reviews the economic regulation and competition issues that emerged in the development of the privatised England and Wales' pool. It provides a better understanding of the regulatory input that emerging markets should expect to deal with once they privatise and deregulate electricity systems. It covers the following: 2.2 present the methodology for the study. The brief summary of the background to British electricity industry and its march towards privatisation is in 2.3. 2.4 discuss how the markets worked post-deregulation; the implications for emerging markets is in 2.5 and 2.6 concludes

### *2.2 Method of study*

The England and Wales' electricity privatisation is considered a great success; hence this paper uses a historical economic regulation case study approach to discuss the transition to competition under the pool. The approach involves the critical evaluation of the economics, regulation and competition processes in the development of the privatised electricity market. Therefore, this paper provides a better understanding of the regulatory and competition challenges that deregulated electricity systems may face in their development.

Vickers & Yarrow (1988) provide a comparative and forward looking analysis of the economic regulation and competition issues which the British telecommunications, transport, water and the energy sectors may face post-deregulation. Green (1999) uses the same approach to examine the problems that inhibit the development of efficient competition in the pool. And Littlechild (1994) evaluates the requirements for a full transition to competition in supply services in the British electricity market. This paper adds to the literature by focussing on the main regulatory input that went into the privatised market during the pool regime and highlights the implications for emerging markets.

### *2.3 Towards unbundling and privatisation*

Before March 1990, the *Central Electricity Generation Board* (CEGB) managed the electricity industry in England and Wales, which was—then—organised as vertically integrated up till the distribution segment. The CEGB was a result of the recommendations made by the Herbert Committee, which was constituted in 1956 to analyse the operations of the electricity industry in England and Wales (Vickers and Yarrow, 1988). CEGB served as a public utility that centrally co-ordinated, scheduled and dispatched all the plants used to meet demand. It owned and operated the power stations as well as the *national transmission system* (NTS); and sold approximately 95% of its production to the regional area boards, which served as the distribution utilities. It used comparative efficiency processes to monitor the performance between those regional area boards. Summarising, CEGB carried out two broad roles: (1) generation and transmission of power and (2) maintenance of efficiency standards, which it achieved by using efficiency benchmarking to stimulate competition between the regional area boards.

### *2.3.2 Initial policies for reforming the British Electricity Industry*

The reform in the UK was directed at stimulating competition in generation and supply through the entry of privately owned *independent power producers* (IPPs). In contrast, to use efficient economic polices to regulate the monopoly businesses in transmission and distribution. The White Paper: *'Privatising Electricity'* (1988) stipulated the industry's regulatory reform and the Electricity Act 1989 provided the regulatory framework upon which the industry was privatised. Generation and supply services were to be separated from transmission and distribution. The government rejected initial proposals to create 4 or 5-generation firms from CEGBs non-base load capacity (Green, 1996); instead, for the CEGB to be divided into three private generations and one *independent system operator (ISO)* utilities. Competition was expected to evolve in generation through the entry of *independent power producers* (IPPs). The governments realised after publishing the White Paper that the sales valuation of the nuclear plants, which accounted for approximately 20% of the system capacity, was much lower than their worth; consequently, it resolved in November 1989 to keep them under public ownership (Green & Newberry, 1998).

Fifty-seven percent (57%) of the non-base load coal capacity (MMC, 1996:84) was divided between two companies: National Power (NP) and PowerGen (PG) in the ratio of 3:2, whilst a third company Nuclear Electric inherited the base load nuclear plants excluding Magnox, which was later sold as British Energy. The shares for NP and PG were floated in two tranches: 60% in March 1991 and 40% in February 1995. Nuclear Electric operated as a publicly owned company until July 1996 when it was floated. 'Nuclear Electric's plant was split: the newer *advanced gas-cooled reactor* (AGR) and *pressurised-water reactor* (PWR) stations, which accounted for about two-thirds of nuclear output, became part of the privatised British Energy and the older Magnox stations became Magnox Electric, which remained in government ownership' (Littlechild, 1998:195). The compulsory pool was created as a mechanism to facilitate wholesales commodity trading. The ISO was required to use a computer mechanism analogous to CEGBs, to centrally co-ordinate and dispatch all the plants across the grid. There was also a provision for a system settlement administrator who would co-ordinate and processes the settlements of the transactions between its members.

In transmission, the 12 *Regional Electricity Companies* (RECs)[6] were joint owners of the grid; and they were allowed to enter the generation market if they wanted to. National Grid Company (NGC), which was the fourth company created at vesting served as the ISO; consequently, it had statutory ownership of the NTS plus the monopoly to transport power over 400KV and 275KV voltage lines. It became the *system operator* (SO) with the responsibility to schedule and dispatch plants centrally, facilitate competition and to set prices. It also determined the capacity and commodity charging methodology subject to the veto rights of the *Director General of Electricity Supply* (DGES) in the—then—*Office of Electricity Regulation* (OFFER). And from 1999, it became the remit of the Director General of Gas and Electricity (DGGE), under the energy regulatory *Office of Gas and Electricity Markets* (OFGEM[7]). NGC was required to publish it use of system charges. In the *local distribution zones* (LDZs) supply services was separated from distribution. The initial polices in distribution were focused on encouraging and developing multiple but privately owned companies (Littlechild, 2001). The RECs shares were floated on the Stock Exchange in December 1990, but they operated as monopoly distribution companies within their LDZs. They operated separate accounts for their monopoly and competitive businesses;

[6] East Midlands Electricity PLC, Eastern Electricity PLC, London Electricity PLC, Manweb PLC, Midlands Electricity PLC, Northern Electric PLC, Norweb PLC, Seeboard PLC, South Wales electricity PLC, South Western PLC, Southern Electric PLC, Yorkshire Electricity Group PLC.

[7] OFGEM was created in 1999 from the merger between OFFER and the Office of Gas Supply (OFGAS).

and the regulator used an incentive-based regulatory tool, the R*etail Price Index* (RPI) minus an efficiency rate ('X'): 'RPI-X', to regulate their tariffs.

In Supply, the White Paper provided that transition to full competition was implemented in phases; it was structured to be phased over 8 years: 1990, 1994 and 1998 (Littlechild, 1994). The first phase occurred in 1990 when approximately 50,000 customers who consumed more than 1MW of power were allowed to choose a supplier from any of the RECs. These types of customers usually had metered sites with load profiles. The second phase began in 1994; it was then that those consuming approximately 100kW of electricity got the right to choose their supplier. Full competition in supply occurred in June 1999, when the remaining over 26 million retail consumers, predominantly households that consumed less than 100KW, got the franchise to choose their suppliers.

In Scotland, the Scottish electricity supply industry (ESI) was also restructured. In June 1991, Scottish Power (SP) and Scottish HydroElectric 'were privatised as geographically distinct and vertically integrated privately regulated utility but Scottish Nuclear remained as a publicly owned utility; it had long-term contracts with the two private utilities (Newbery, 1995:42).

### *2.3.3 Summary*

The White Paper set out the initial proposals for privatising the industry in 1990. Those proposals were based not only on introducing competition in generation and supply, but also for *'ring fencing'* and regulating the monopoly businesses of the transmission and distribution companies. It also provided scheduling and dispatching of all the plants used to meet demand was centralised in a similar manner to the practise under the CEGB regime.

## *2.4 How did the industry segments work post-deregulation?*

This section discusses the key milestones in the operation of the different segments during the pool regime. Sub-section 2.4.1 covers the pool's wholesale trading arrangements that the bilateral contracts market complemented and 2.4.2, transmission, distribution and supply.

### 2.4.1 Organisations of wholesale trading

Wholesale trading is usually organised through competitive arrangements such as a pool and / or a power exchange. In theory, the market maker for a pool is a non-profit making entity that organises, administers and carries out all the settlement of the transactions between its members. Like other commodity markets, electricity pools have buy and sell prices both of which can be set before physical delivery occurs. A price that is set before real time is said to be *ex-ante* whilst the one set after physical delivery occurs is termed *ex-post*.

In other commodity markets, geographical price differences are usually attributed to transportation costs whilst the difference between the buy and sell price is the *market makers'* income. In electricity, the pool market maker does not earn any income; instead, the difference between commodity buy and sell price, represents the cost for the *use of the system*. The latter is also part of the cost that the SO incurs to maintain the security of the *national transmission system* (NTS). It includes costs such as those for resolving constraints: constraining *'on'* and *'off'* plants, and for procuring balancing services such as reserves (spinning and non-spinning), black start and reactive power. Some pools set the purchase price before physical delivery; and add the use of the system charge to it to derive the sell price. This was the rule in England and Wales' pool (see The Electricity Pool, 1997).

#### 2.4.1.1 The Pool Arrangement

The England and Wales', pool was created in March 1990. It was an unincorporated association of all the agents' licensed to use the grid; and included the on-shore generators, IPPs plus the imports from France and Scotland. Being a re-creation of CEGB, it used computer mechanism to schedule, plan and dispatch all the plants across the grid; and to set commodity prices. It operated a dual pricing system with a buy price which was determined *ex ante* and on a day ahead basis, whilst the sell price was set *ex-post*.

*2.4.1.1.1 Governance*

The *Pooling and Settlement Agreement* (PSA) was the document that governed the pool. It set out the exact procedures for its daily operation, interaction of members, administration, guidelines on how payments were calculated and specified the contractual obligation required of each pool member (OFGEM, August 1999). Apart from the PSA, there was also the 'Grid Code', which set out the interaction and contractual obligation between the grid users and the NGC. It also provided guidelines on how NGC could optimise resources to carry out its SO roles.

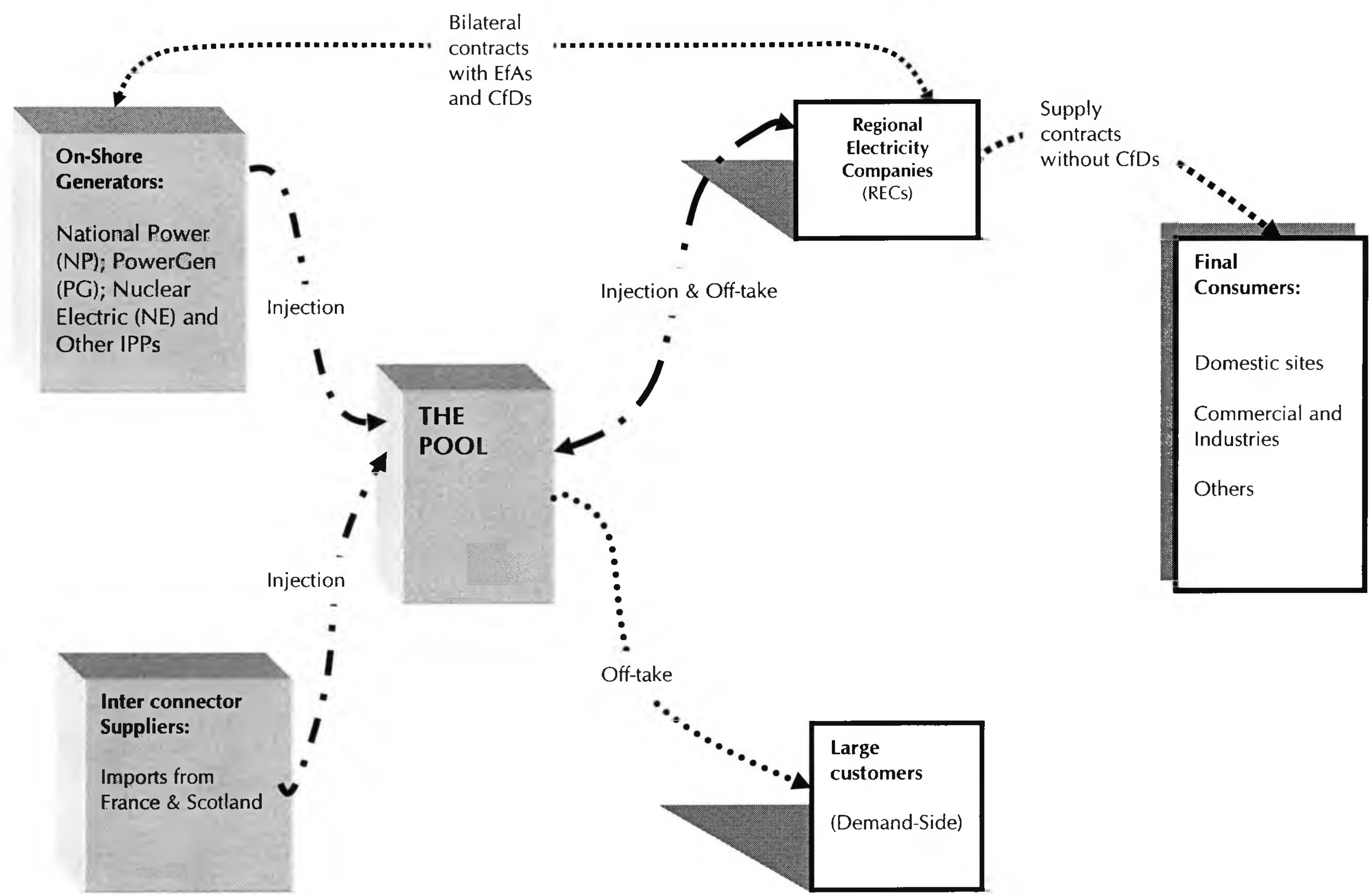
*2.4.1.1.2 Daily operations and administration*

The daily operations and administration of the pool was much more complex than that of any of the other known spot markets. It functioned through two mechanisms: a physical day-ahead pool and a bilateral financial contracts market. Figure 2.4.1 is a simplified representation of the interaction between these two sub-markets. The physical flow involved generators injecting power onto the grid; having contracted with NGC, it transported the electricity through the high voltage lines onto the distribution lines. In the LDZs, the distributions companies completed the transportation to the final consumers. Injection onto the grid is represented as an inflow into the pool from the on-shore generators, which included the RECs that have generation plants and the interconnector imports from France and Scotland. The demand-side that purchased electricity in bulk are labelled *'off takes'*; they were mainly metered sites, had load profiles and alternative means for generating their own power and thus could enter into interruptible contracts with the SO. The final consumers usually entered into contractual agreements with their electricity suppliers. In the bilateral financial contracts, generators' and the demand-side used agreements that were limited to only financial settlements to hedge against volatile movements in the pool prices.

A discussion of the operation and how prices were set including the highlights of the regulatory oversight in the pool and the bilateral contracts market follows.

**Figure 2.5.1**

**A simplified structure of the physical—inflows and off-takes—and bilateral contracts in the pool**



### 2.4.1.1.3 *Physical day-ahead pool*

The organisation of the day-head pool involved the centralised planning of an unconstrained schedule for dispatching plants on the day and the setting of ex-ante prices.

*Central control and dispatch*

NGC started the processes of price and capacity determination by using historical information on *seasonal normal demand* (SND) to forecast aggregate gross consumption by the pumped storage, large loads that consumed about 250MW and the *non-daily metered* (NDM) customers. It also estimated and inputted a value for reserve. Thereafter it offered half-hourly capacity slots for injections and *off take* from the grid to generators and the demand-side.

By 10 am on the day-ahead, generators' submitted pairs of half-hourly price and quantity offers to the pool. In theory, they submitted supply functions (Green and Newberry, 1992). The demand-side also submitted combinations of half-hourly price and quantity bids at which they were willing to decrease—decrement—consumption or increase—increment—consumption if the SO required such services on the day.

NGC used its *Grid Ordering and Loading* (GOAL) model to simulate an unconstrained schedule that it relied upon for dispatching plants on the day. This process involved stacking generators' offers in an ascending order of marginal cost. It also took cognisance of the flexibility of the plants declared available to run in accordance with the Grid Codes definition of operating plants' availability. For example, the least cost and most inflexible plants such as a nuclear plant was scheduled first and used to meet base load demand. As demand peaked, it brought on the more expensive but relatively most flexible plants such as the hydro or *open-cycle gas turbine* (OCGT), which were predominately used to meet demand in the peak segment of the *Load Duration Curve* (LDC). This cost-based scheduling of plants is called *merit ordering*.

*Price Rule and Prices setting*

The pool was modelled as a last price auction market in which the entire England and Wales was a single price zone (Klemperer, 2000; Elmaghraby, 1998; von der Fehr and Harbord, 1993); consequently one generator supplied power at the margin. There was a buy, a *pool purchase price* (PPP) and a sell, *pool selling price* (PSP), both of which were based on a *system marginal price* (SMP) that was set on the day-ahead. Unlike what happens in some of the other commodity markets, the difference between PPP and PSP was not NGCs income; instead, it represented the cost of keeping the NTS safe and secure within its tolerance limits.

The SMP was derived at the point of intersection of the supply and NGCs forecasted inelastic demand curve. It was the cost of the last MW of power and from the most expensive generating set that was used to meet demand at each of the 48 half-hours within a day. Hunt (1995) provides a decomposed definition for SMP as:

> "System" because all plants are bidding; "Marginal" because it is the last plant that counts; "Price" because it is the generator's price that counts—its cost are not in question (although the system is designed to ensure that generators have every incentive to bid at their own marginal cost)' (page 7).

*Pool Purchase Price (PPP)*

All the generators' received a *pool purchase price* (PPP) for any MW of power they sold into the pool. PPP was calculated as:

$$PPP = SMP * (1 - LOLP) + VOLL * LOLP$$

$$= SMP + LOLP(VOLL - SMP)$$

$$= SMP + CP$$

Where CP, capacity payment, is the payment made to generators for any MW of power declared available, even if within day inter system transmission shocks prevented such plants from being dispatched on the day.

The $LOLP$ , the loss of load probability, was calculated 24 hours ahead and based on a scheduling programme. It ranged between 'one' and 'zero'; and provided an indication of the systems capacity to meet increments in demand on the day. It was designed to be highest when the network was under demand stress and lowest during the periods of excess capacity relative to demand. The LOLPs role made it the most sensitive component in the calculation of CP.

$VOLL$, the *value of lost load*, was imputed 'from utility planning models', and measured the social cost of the systems failure to meet demand (Kwoka, 1997). VOLL was designed to be greater than the equilibrium SMP. The government set it at £2 / kWh (£2000/MWh) in 1990 / 91, but it was subject to annual increases based on the *retail price index* (RPI) (OFFER, 1991).

*Pool Selling Price (PSP)*

The demand-side paid the *pool-selling price* (PSP) for any MW of power they took from the grid. PSP was calculated as:

$$PSP = PPP + Uplift$$

Uplift was the cost that NGC incurred to maintain the transmission system within its acceptable energy, frequency responses, and voltage tolerances. Although NGC simulated an unconstrained schedule, which it used to meet demand on the day, there were circumstances when maintenance of system security required it to re-direct the flow of power across the NTS. There were also times when errors in forecasting demand combined with inter-system transmission shocks, led to deviations between forecast and actual volumes demanded and / or supplied. At such times, NGC dealt with power re-directions by constraining *'on'* or *'off'* plants as required. It also procured balancing services, to resolve locational nodal imbalances, by contracting directly with the generators' that had the facilities to provide such services. All of these costs that NGC incurred to restore the system to its balanced tolerances were recovered from the demand-side through the *uplift* charge.

Summarising, the pool was analogous to a commodity spot market; it differed in that all the licensed grid users were compelled to sign onto it; and as members, bought or sold energy only through the mechanism. As a result, it facilitated two things (1) physical contracting and (2) the setting of prices. Its buy price: PPP, which was based on an unconstrained schedule of declared availability and gross demand, was set *ex-ante* and on the day-ahead. And its sell price: PSP was calculated after establishing an *ex-post* cost for keeping the system safe and secure.

*2.4.1.1.4 Evolution of prices*

Tables 2.4.1 and 2.4.2 tabulate the time-weighted and demand-weighted average pool prices from 1990 to 2000. It shows SMP being the highest component of both the PPP and PSP. It also suggests that there was a slow emergence of the level average value of the pool prices. On average, the lowest mean prices after 1994 / 95 occur in 1999/2000 and suggests that there might have been an influence of some cumulative factors on pool prices towards the run up to the implementation of the *new electricity trading arrangements* (NETA) that *'went live'* on March 27, 2001.

**Table 2.4.1**
**Time-Weighted Annual Average Pool Prices (CC, 2001:125)**

*£/MWh, September 2000 prices*

| *Year* | *SMP* | *CP* | *PPP* | *Indexed PPP (1999/2000) = 100* | *PSP* |
|---|---|---|---|---|---|
| 90/91 | 23.15 | 0.06 | 23.22 | 98 | 24.44 |
| 91/92 | 24.85 | 1.65 | 26.51 | 112 | 28.56 |
| 92/93 | 27.94 | 0.24 | 28.14 | 119 | 29.85 |
| 93/94 | 29.33 | 0.34 | 29.67 | 125 | 32.32 |
| 94/95 | 24.54 | 3.77 | 28.31 | 120 | 31.12 |
| 95/96 | 22.19 | 5.10 | 27.29 | 115 | 29.61 |
| 96/97 | 22.86 | 3.60 | 26.47 | 112 | 28.56 |
| 97/98 | 26.42 | 0.92 | 27.34 | 116 | 27.77 |
| 98/99 | 24.76 | 1.05 | 25.81 | 109 | 26.14 |
| 99/00 | 20.76 | 2.89 | 23.64 | 100 | 24.36 |

**Table 2.4.2**
**Demand-Weighted Annual Average Pool Prices (CC, 2001:125)**

*£/MWh, September 2000 prices*

| *Year* | *SMP* | *CP* | *PPP* | *Indexed PPP (1999/2000) = 100* | *PSP* |
|---|---|---|---|---|---|
| 90/91 | 23.87 | 0.08 | 23.95 | 94 | 25.25 |
| 91/92 | 25.39 | 2.09 | 27.48 | 108 | 29.77 |
| 92/93 | 28.50 | 0.26 | 28.76 | 113 | 30.57 |
| 93/94 | 29.95 | 0.43 | 30.38 | 120 | 33.17 |
| 94/95 | 25.89 | 4.65 | 30.55 | 120 | 33.64 |
| 95/96 | 23.41 | 6.33 | 29.74 | 117 | 31.99 |
| 96/97 | 24.00 | 4.48 | 28.48 | 112 | 30.83 |
| 97/98 | 27.76 | 1.19 | 28.95 | 114 | 29.45 |
| 98/99 | 26.10 | 1.31 | 27.41 | 108 | 27.81 |
| 99/00 | 21.81 | 3.57 | 25.38 | 100 | 26.25 |

Table 2.4.3 shows the trend and growth rate in the weekly average of the half-hourly pool prices at the end of each sub-sample period.

**Table 2.4.3**
**Trend and annual growth rates of weekly average pool prices (at the end of the period of each sample period)**

| *Component* | *January 1994 to December 2000* | *January 1994 to December 1998* | *January 1999 to December 2000* |
|---|---|---|---|
| SMP | £19.08 (-3.13%) | £22.12 (2.45%) | £19.14 (-16.20%) |
| CP | £9.10 (162.37%) | £0.04 (-106.95%) | £17.62 (195.13%) |
| PPP | £22.09 (-2.35%) | £23.76 (4.28%) | £21.53 (-5.51%) |
| PSP | £23.04 (-1.70%) | £26.47 (507.44%) | £22.46 (-4.91%) |

Note: The growth rate at the end of the period is in parenthesis.

Table 2.4.3 reveals that PPP and PSP follows a decreasing trend when examined on a full sample basis. The decrease in the trend value of SMP is the primary cause of the decreases in PSP and PPP between January 1999 and December 2000. Increases in competition which resulted from the second phase of the divestiture of NP and PGs mid-merit plants; changes in capacity mix and the knock on effect of the changes in the gas trading arrangements, were some of the factors that led to the significant reductions in the SMP after 1998 (see Evans and Green, 2003; Bower, 2002; OFGEM, 2002a & b).

What is surprising and important for public policy is the rate at which capacity payment grows after 1998. The only driver for the higher CP values is increases in the LOLP; this is because SMP follows a constantly decreasing trend and the VOLL was constant. Given the volume of capacity that was available on the network combined with the thermal efficiency as well as the operational flexibility of the plants that NGC used over the same period to meet demand, the price pattern suggests that generators manipulated capacity. This raises a question of whether paying generators to make their plants available to run was an appropriate policy on a system that had excess capacity far above the notional 20% plant margin (NGC, 2000) required for keeping a system safe and secure? Paying generators to make their plants available in electricity markets should be forward looking, robust and easy to modify as soon as the plant margin exceeds the highest peak demand based *on average cold spell conditions* (ACS). And once removed public policy should be directed at steering forward initiatives that will enhance efficient competition between generators to be in merit.

*2.4.1.1.5 Emergence of Price Trend and manipulations*

NP and PG started to manipulate capacity and prices ten months into the operation of the pool (see Fehr and Harbord, 1993). In September 1991, the *DGES* set up his first inquiry into the cause of an approximately twenty-nine percentage (29%) increase in the average pool prices over the previous year. In his decision document, which he published in December 1991 (OFFER, 1991), he reports that NP and PGs commercial strategies contributed to the higher prices. He also provides an insight into some of the capacity manipulation strategies that the incumbents might have used to earn higher profits. For example, he finds generators making indiscriminate use of the G*reater than or Equal to* (GE) flexibility marker, a tool included in the Grid Code to help them signify operational flexibility of plants. NP particularly offers its out of merit plants that are located behind short or long term constraint boundaries, at very high prices, because they know that NGC will need them to maintain the security of the system. They manipulate the threshold for setting the *table indicators* and profile availability within the day. The entire capacity manipulations increase capacity and uplift payments (OFFER, 1991; see also OFFER, 1999). The DGES indicates that it might be necessary to steer forward initiatives to withdraw capacity payment in the future if the type of capacity manipulations that he identified continues. Nonetheless, he did not take this forward before implementing *trading outside the pool* (TOP).

It was clear in 1991 that efficient competition might be very difficult to achieve within the duopoly industry structure that was created at vesting. High prices persisted during the early days, leading the DGES to introduce demand-side participation in 1993 with the expectation that it would help to restrain high SMPs. It allowed customers that consumed approximately 1MW of electricity to bid directly into the pool (Wolak and Patrick, 2001). But the demand side participants did not really curtail high peak prices; their bids were often above SMPs and greater than £50/MWh (see Bunn, et al 1997). Consequently, the NGC did not include them in the SMP determination process.

*2.4.1.1.6 Towards divestment & limit pricing*

The DGES had no powers vested in him through the Act to regulate pool prices, presumably because it was anticipated—at the design stages—that the mechanism would be self-equilibrating. Instead, the Act provided for him to make references to the *Competition Commission* (CC)—formerly the Monopolies and Mergers Commission—about any participant whose conduct inhibited the development of competition. Tables 2.4.1 and 2.4.2 show that the SMP is the primary cause of the increases in PPP between 1990 and 1993. Since the generators cost of production decreased over the same period, the only possible reason for the higher prices is that they made offers, which did not reflect their true costs. Because more thermally efficient plants came on line and there was also a significant reduction in the costs of entry.

In February 1992, the Energy Select Committee raised concerns that NP and PGs *pre-emptive behaviour*[8] kept prices above the competitive levels. Consequently, they recommended that the DGES referred the two generators to the Monopolies and Mergers Commission (MMC) (OFFER, February 1994b). As a result of the high prices, which persisted, in April 1993, the DGES decided to reduce NP and PGs control of the industry's residual demand.

He secured an undertaking from NP and PG that they would respectively divest fifteen percent (15%): 4000MW and ten percent (10%): 2000MW of their mid-merit plants. NP and PG negotiated and sold off the capacities to the largest REC at that time: Eastern

[8] 'pre-emptive behaviour' describes abuses by dominant incumbents in markets that are being opened to competition or are newly opened to competition, which are designed to adversely affect the development of competition' (OFT, March 2001:10)

Electricity PLC (Eastern) (now part of TXU Europe) in April 1996. The divestment was initially organised as a sale and lease back with £6 / MWh advance rental fee. The DGES capped prices between the time that he expected it to take for NP and PG to conclude the transfer of ownership to Eastern. The DGES secured an additional undertaking from NP and PG that they would use reasonable endeavour to ensure that prices remained at £2.4p/kWh time-weighted and £2.55p/kWh demand-weighted, between April 1 1994 and 31, March 1996. The price cap was based on the October 1993 prices (see OFFER, 1994).

Following the caps, tables 2.5.1 and 2.5.2 shows a very slight decrease in PPP between 1994 / 95 and 1995 /96. But pool prices exceeded the average time and demand weighted caps in October 1995 when they were £25.35/MWh and £26.94/MWh (CC, 2001). This suggests that despite the divestment that NP and PG still had a strong control in the setting of the marginal price. Eastern's lack of effective mid-merit competition may be because they bought the incumbents older and relatively inefficient plants. Moreover, the percentage of the plants that NP and PG sold were a very small proportion of their total mid-merit portfolio; therefore they still had a strong influence on the industry's residual demand. On the other hand, given that Eastern paid a £6 rental fee, they might have offered their mid-merit capacity at very high prices, to enable them earn sufficient income to cover their rental plus some income for augmenting working capital requirements. These meant that the divestment was not a success. The DGES did not extend the cap when it expired.

The regulator also used persuasive pressure some of the time to keep pool prices low (Green, 1999). But after 1998, much more than strongly worded communications with the generators was required before the opening up of full retail competition combined with the harmonisation of gas and electricity trading arrangements. Without reference to the CC to conduct an investigation on the potential adverse effect on public interest that might arise from a merger between the incumbent duopoly generators and the RECs, in 1999, the DGES agreed that NP and PG could vertically integrate with any REC. In return, he requested that they divested more portions of their mid-merit plants. This marked the second phase of divestment in the pool. In July 1998, PowerGen bought East Midland Electricity and sold 2GW of its Ferrybridge and FiddlersFerry to Edison Mission Energy Limited (Edison) in July 1999. Similarly, National Power bought Midlands Electricity in July

1999; in November 1999, it sold its 4GW Drax plant to AES (OFFER, 1999; CC, 2001) and in February 2000, it sold its Eggborough plant to British Energy.

*2.4.1.1.7 Towards a market abuse licence condition (MALC)*

By the mid-1990s, natural gas and electricity markets converged, with the impact of the convergence felt more in electricity than in gas (see Larsen & Gary; 1998). There were over 20 active shippers[9], major gas and electricity retailers and electricity generation licences in Britain. The utility firms were also gearing up to OFFER retail consumers bundled gas and electricity products once full competition in retail commenced.

One of the ways by which the government acknowledged the convergence between natural gas and electricity was the merger of the Office of Gas Supply (OFGAS) and Office of Electricity Regulation (OFFER) into a single energy regulator: the Office of Gas and Electricity Markets (OFGEM). This happened in 1999 (OFGEM 2000b). In addition, OFGEM and the Department of Trade and Industries (DTI) began steering forward initiatives to harmonise the capacity and commodity regimes in natural gas and electricity. In October 1999, OFGAS directed Transco[10] to implement the *New Gas Trading Arrangements* (NGTA)[11]. Whilst consultations continued on modalities to implement the new electricity trading arrangements (NETA) and the integration of Scotland into a *GB-wide Electricity Trading and Transmission Arrangement* (BETTA) (see OFGEM, 2000a).

[9] '[A] Shipper – a company with Shipper Licence buys gas from producers, sells it to the suppliers and employs the [gas transporter a] GT to transport the gas to consumers. It may also store gas with a Storage Operator to help it manager the balance between its suppliers and the consumer's demand. Its licence requires it to be reasonable and prudent in the way it uses the GTs pipeline Network' (Transco, b: 6).

[10] Transco was the ISO for natural gas

[11] The NGTAs process included the introduction of an on-the-day commodity market (OCM), which EnMo, and independent firm operated; introduction of auctions for allocation of *firm* capacity rights from the beach terminals onto the gas NTS and the energy and capacity regimes.

The second six-monthly auction of the *Monthly System Entry Capacity* (MSEC) from the *six beach* terminals[12] onto the gas NTS occurred in March 2000. The realised prices at the auction rounds rose by over 1000% to what was obtained during the pioneer auctions in October 1999; Transco also over recovered its allowed annual revenue by approximately £84M. Thereafter, the industry participants increased their concerns about the potential impact of changes in gas usage as well as within day swings, on capacity regimes in gas, which they expected, might follow after the implementation of the NETA. The foundation for the concerns, which was based on the differences in the balancing period between the gas and electricity markets, seemed justified.

Generators faced half-hourly whilst shippers were expected to balance their injections with off-takes at the *end-of-day* (EoD); and technical constraints prevented the possibility of harmonising balancing periods between the two markets. The situation presented a scenario in which the interaction between NGTA and NETA might be biased in favour of some grid users. The industry participants that we spoke to at that time as well as the colleagues at the Association of Electricity Producers (AEP) workshop in year 2000 expressed concerns that *combined-cycle gas turbines* (CCGTs) might become baseload plants after the implementation of NETA. NETAs interaction with NGTA might be more favourable to portfolio generators as the dual fuel plants might be in vantage positions to maximise their private marginal benefits by arbitraging, which might allow them to burn gas and sell as electricity.

Apart from the grid users, the DGES was also very concerned that Transco might be subjected to taking increased number of balancing actions within the day. The anticipated indirect effect of the changes in gas usage and swings would be increases in Transco's costs for balancing the gas network; and shippers were more likely to pass through increases in their wholesale prices to the customers that had cost pass through clauses in their contracts. The feeling at OFGEM was that if the DGGE failed to secure a contractual undertaking from the generators' to 'behave well' in the interactive NETA / NGTA market, the development of contestable capacity and commodity regimes in both markets might be jeopardised.

[12] St. Fergus, Bacton, Barrow, Theddlethorpe, Teeside and Easington.

Finally full retail competition was initially scheduled to commence in 1998; this was expected to coincide with the expiration of the IP contracts as well as emergence of the NGTA. Post NETA, the industry participants expected that mergers would mark the era of a fully competitive and harmonised energy market in England and Wales. The main concern about the prospects of mergers was whether it would improve social welfare; in particular, if vertical integration reduces the number of players, it might mean the creation of pockets of monopolists, which would restrict competition and increase prices.

In summary, Mr. McCarthy—the then—Regulator felt that there were threats to the success of the interaction between NGTA and NETA and the security of the gas NTS; with the knock on effect being increased gas costs to the final consumer. Consequently, and relying on the powers vested in him through the *Act*, in 2000, he sought—but failed—to include a *Market Abuse Licence Condition* (MALC) into the generators' licences.

Some of the industry participants that I spoke to at that time, expressed concern that MALC was an open-ended as well as an arbitrary clause that the DGGE intended to use to deal with above costs pricing and anti-competitive practises that had plagued the pool since its inception. AES and British Energy refused to allow that MALC be included into their generator licences'; and the dispute was taken to the *Competition Commission* (CC).

Both in his oral and written submission, Mr McCarthy's—the then DGGE—warned that NETA might not solve the problems that inhibited the development of efficient competition in the privatised industry. He supported his conviction by citing the views expressed by some academic economists (see for example, Rudkevitch, 1997; Wilson, 2001, Gordon, 2000), which includes the effect of the lack of storage in appreciable quantities on prices. The lack of substitutes to the final consumer combined with the relatively low price elasticity limits the demand side participation in markets. Transportation can only be done over existing cables and wires; that is, it is not possible to have more than one transmission system serving a geographic boundary. There are also requirements to maintain a balanced and synchronised grid activity. A capacity and transmission constraint affects the safety of the NTS; moreover it is difficult to find an efficient methodology for pricing access on to the grid and managing congestion. Entry and exit into the industry is not free meaning that the idea of developing contestable capacity and commodity regimes is best an illusion. In sort, electricity industry does not fit the prototype of a contestable market. The Oligopoly nature of the industry combined with the peculiar features of electricity that makes it difficult to trade commodity as what obtains in other markets enhances anti-competitive practises. On the other hand, generators can easily collude to keep market-derived prices above their true costs; this means that the aggregate costs of generation will always be above competitive levels (see Fehr and Harbord, 1993; Tirole, 1998). It was based on the above grounds of argument that Mr. McCarthy suggested and actually reiterated the importance of having 'special measures' in place, which he could use to curtail generators from abusing their monopoly power. He explained that such a measure was particularly necessary to steer forward a successful interaction between the NGTA and NETA (OFGEM, 2000c).

The CCs investigation focused on how MALC would facilitate the promotion of efficient competition; in particular if it was really a pre-requisite for restraining high prices in the industry. It concluded that MALC was not a necessity and upheld that AES and British Energy operating without MALC would have no adverse effect on prices (Competition Commission, 2001).

#### *2.4.1.1.8 Criticisms about the pools prices rule and capacity determination process*

The pool faced a lot of criticisms right from its inception. The criticisms made it very hard to assess how the regime might have worked, but it seems that the assessment was a crucial part of a better understanding of what the regulatory reforms were really meant to achieve (Green, 1999: 516). Generators manipulated pool prices because of some of the inherent inefficiencies in its design and rules. Its governance was rigid; consequently, it was difficult to process modifications to market rules in a timely manner. It was a 'mechanism'; its simulated prices were derived from NGCs very complicated stochastic model and which was subject to technical constraints. Therefore, it delivered inefficient prices sometimes (see von de Fehr & Harbord, 1998; Kwoka, 1997). The price rules inhibited the level of transparency that was required to enhance the rapid development of the futures market, which could OFFER a wide variety of products, which the participants could use to hedge against movements in prices (Littlechild, 2001).

Based on my discussions with colleagues at the Office of Electricity and Gas Markets (OFGEM), it was clear that the DGGE was also concerned that the generators lacked the incentive to make economic and efficient decisions because the offers and bids that they made into the pool were not *'firm'*. The methodology of averaging imbalance costs to the demand-side in proportion to their system throughput also provided no incentive to generators to use reasonable endeavour to maintain balanced positions at all times. Those were some of the reasons why it was very easy for the generators to maximise their marginal benefits from pool simply by manipulating capacity. The DGGE believed that the costs that NGC incurred to balance the system within day would reduce significantly if the generators were directly responsible for the imbalances that each one contributed to the network. He felt that the price mechanism was inefficient in so far as it appeared not to reflect the true costs of producing electricity and the security situation on the system. The grid was a relatively unconstrained one; the costs of entry for CCGTs reduced as competition evolved and there was further decreases in the fuel costs especially National balancing Point (NBP) gas prices. His view was that a better cost-targeted market other than the pool mechanism was the way forward and would deliver low balancing costs.

*2.4.1.1.9 Reform of the Electricity Trading Arrangements (RETA)*

Professor Stephen Littlechild, the first DGES, considered options to *trade outside the pool* (TOP) in 1994 and concluded that the transaction costs might be prohibitive; there was no evidence that it would yield significant efficiency gains. Instead, it seemed that it might inhibit the development of efficient competition.

> … There [was] little tangible evidence of gains likely to be secured, and it would be time-consuming and costly to make the necessary arrangements. There [were] potential detriments to competition and new entry, both from a thinner and less transparent market and from placing new entrants and smaller competitors at a disadvantage in securing rights to dispatch. [Those] detriments [had] to be given particular weight in view of the present positions of market power on both sides of the market. Any significant change might also have adverse effects in terms of perceived market stability. [He] conclude[d] that … a sufficient case [had] not been made for significant changes to existing arrangements so as to allow widespread trading outside the pool.

> [He] should be prepared to reconsider the matter in future, if circumstances warranted it. This might particularly be in the events of a change in market conditions or a lack of progress in pool reforms, or the identification of more tangible benefits from trading outside the pool, or the establishment of less wide-ranging or less costly arrangements to affect it (page iii).

In 1997, the Regulator revisited TOP and launched the RETA. The DGES concluded that TOP was the way forward for the market and published the model for the NETA in 1998 (OFFER, 1998). In the summer of 1999, London Economics in conjunction with OFGEM used both the existing industry participants as well as students from different Universities in the UK, to simulate real life trading positions with the objective to maximise profits. London Economics (LE) report concludes that NETA would be liquid and changes in the price rule which would see a move from the uniform SMP to a discriminatory pay-as-bid (PAB) in the residual Balancing Mechanism (BM) would deliver an efficient market (OFGEM October 1999). And NETA was implemented on March 27, 2001.

*2.4.1.1.10 Summary*

The British pool was a compulsory competition-based mechanism through which electricity was traded. The whole of England and Wales was a single price zone, hence all generators earned a uniform SMP, which was derived at the intersection of the generators declared availability and NGCs forecast of gross demand. At its design stage, the consultation parties felt that the combination of the SMP setting process, rules for deriving PPP and PSP combined with the methodology for deriving the *Transmission Network Use of System Charge* (TNUoS) on the NTS would lead to production and allocative efficiency (Littlechild, 2001). Nonetheless, pool prices failed to reflect the generators true costs of production; it did not also reflect the demand and supply as well as the system security situation across the network. Anti-competitive practises persisted in the pool; and it seemed that it would worsen in the interaction between NGTA and NETA. As a result, the DGGE decided—but failed—to secure a contractually binding way to ensure that the changes in gas usage combined with the increases in within day swings, which were anticipated after the implementation of NETA, did not jeopardise the development of competition in gas and electricity markets. The pool was closed and NETA started on March 27, 2001. The next sub-section evaluates the pools complementary contract market.

### 2.4.1.2 *Bilateral Contracts Market*

Prices volatility is one of the generic problems in electricity markets; they can occur in the absence of transmission or capacity constraints (see Hogan, June 3, 1998; Hale et. al, 2000). However, when *demand stress* occurs, which can result from changes in weather conditions, plant outages and transmission or capacity constraints, price spikes can be worse. Unexpected constraints, which can make it impossible to transport power easily across the network, worsen the uncertainties about physical deliveries. The other risk that the participants face in spot markets relates to the opportunistic behaviour of competitors, which can result in higher commodity prices.

One of the ways by which the utility firms can curtail volumetric and operation risks is by planning and organising supply weeks, months and years ahead of real time. They do this by locking in physical deliveries into long-term contracts through the financial markets; thereafter, they rely on the spot market to fine-tune positions closer to real time. Financial markets offer physical contracts, which are prepared, as binding agreements and are collectively known as financial options. Most electricity markets trade variations of these types of agreements called: *swaps* and / or *options* in financial markets.

#### 2.4.1.2.2 *Advantages of financial markets*

The products that financial markets offer gives risk averse utilities the opportunity to standardise their risks in monetary terms (Mork, 2001); thereby enabling them to diversify away those risks to agents that may be more disposed to undertake them, subject to an agreed premium. Therefore, financial markets facilitate planning by giving utility firms certainty about physical supplies. If an agent locks in 100% of its physical deliveries into long-term contracts, it will not have any incentive to manipulate spot market prices because its revenue will depend on the agreed price which it made with the contracting party on the deal day. Since its actions in the spot market will not have any effect on its revenue, it will not have any incentive to manipulate prices. If after locking in 90% of its deliveries in long term contracts, a generator prices it residual 10% above competitive rates, it will not be called. It will lose revenue because it will have to purchase the outstanding 10% capacity at the system *top-up* rate. Because its dominant strategy will be to get called, it will have the incentive to OFFER its 10% capacity at prices that will be close to its avoidable cost. Finally, assuming that there is no threat of entry and there is a monopoly generator that locks in only 10% of its deliveries into long term contracts. Its profit maximising strategy will be directed at offering its 90% capacity at its highest possible price. Prospective entrants into an electricity market that has a developed financial market sub-segment can contract ahead for inputs as well as supplies. This opportunity enhances debt-financing arrangements because lenders can make decisions that would be based on partially confirmed cash flow.

The illustration in the previous paragraph shows that the quantity of physical delivery that is locked into long-term contract determines the severity with which generators can manipulate spot market prices. Based on this explanation, contract markets enhance stability both for prices and output; it allows forward contracting for input and supplies before entry and they provide reliable signals about the industry's *long run marginal costs* (LRAC). It is the aspect of the industry that closely fits the prototype of contestability (see Powell, 1993, Bunn, Day, Larsen & Vlahos, 1997; Newberry, 1995).

*2.4.1.2.3 Types of Financial Agreements*

The two types of financial agreements that exist in electricity markets are:

- Futures and
- Forwards

*Futures Contracts*

Futures contracts are non-physical agreements between two or more industry participants that are limited to only cash settlements. They are usually standardised and with transparent price quotes (Powell, 1993); and based on the understanding that a supplier (usually a buyer) and a generator (usually a seller) compensates each other for changes above contract or strike price. Therefore, a futures contract requires a party in a bilateral transaction or parties in a consortium, to undertake to pay deviations that arise between a strike and an on a day price, to another party or consortium. For example, if a supplier and a generator enter into a bilateral agreement at a strike price of £10 and the market price increases to £15, the generator will pay the supplier £5. Similarly, if the price decreases to £8, the supplier will pay the generator £2.

It is possible to use futures contracts on the transmission system; in such a case, it serves as congestion contract, which contracting parties can use to hedge price differentials between Nodes and Zones. Timing and duration of the contracts are usually the issues for the parties to the agreement to address. But once agreement is executed, the parties can organise and plan dispatch on relatively stable prices across locations because their operating decisions will be based on strike prices that are fixed (see EMF, 2000).

*Forward Contracts*

A forward contract occurs when two parties or a consortium enter into a bilateral agreement for physical delivery at a future date. These types of agreements often have adjustment clauses or covenants such as delivery point, acceptable voltages and frequency tolerances. A participant who hedges its position through forward contracts is only exposed up to the fixed volume of the 'physical quantity' and the 'strike price' agreed on the day the deal is made.

#### *2.4.1.2.4 Financial market in the pool*

There was a clear recognition at the design stages of the privatised British electricity market that the participants would not be happy to be exposed to the risks associated with movements in pool prices (Littlechild, 2001). Therefore, a bilateral supply contract segment was included in the regulatory reform, with the expectation that a financial market that would provide a variety of products, which the participants could use to manage operations and volumetric risks, would emerge as the market evolved.

The pools financial market developed and offered two products: *contracts for differences* (CfDs) and *Electricity Futures Agreement* (EfA). The CfDs were different from the IP contracts. They were structured as one-way or two-way contracts; they were predominately used for long-term contracts that lasted from as little as six months to as long as 15 years. 'Under a one-way CfD, a generator paid a supplier the difference between PPP and the strike price of the contract, whenever PPP exceeded the strike price in return for an up-front payment. A contract with a strike price was equal to the marginal cost of the station, and an up-front payment equal to its fixed costs (plus a profit margin), meant that the contracts buyer was buying the station's output at a price close to cost. The contract was triggered whenever the pool price was above the strike prices, which were the times when the station was running, if it bid at marginal cost. The generator replaced the uncertain stream of pool prices with payments that were equal to its costs, as long as it generated when required. Two-way CfDs required the buyer to pay the seller whenever the pool purchase price was less than the contracts strike price. They acted as a straight hedge on the price, and were normally traded with a strike price that was equal to the expected pool purchase price (plus any margin for risk hedging) and no up front premium' (Green, 2001: 6).

The contracts market also offered the *Electricity Futures Agreement* (EFA), which was used mainly to cover deals that lasted one week to six months. The contracting parties could tailor it to suit their preferences.

Comparing the two products, generators used long term CfDs to lock in over 80% of the industry's physical delivery (OFFER, 1994; CC, 2001). In contrast, specification preferences limited the tradability of EFA, leading to lack of liquidity in that sub-segment (Bower and Bunn, 2000).

*2.4.1.2.5 Summary*

CfDs and EfAs were the two products traded in the pools complementary contract market. It was the segment of generation that made entry into the industry contestable and it provided stability in products and prices. It also provided reliable signals about the industry's LRAC.

### 2.4.2 *Transmission, distribution and supply*

The Regulatory reform in the UK was based on *RPI-X* incentive regulation. Because the utility firms kept the difference between the allowed revenue and their actual costs, they have the incentive to use best practices to curtail the operating costs, which they could control. The 'X' efficiency which was re-set between periods was calculated based on estimates of the growth in demand and productivity levels; value of existing assets and the progress of competition. The RECs were regionally based so it was possible that they might have had heterogeneous differences in their cost structures. But the Regulator did not incorporate yardstick elements into the 'X'; however, it might have been recognised during the price reviews (Armstrong et al, 1998). Merely setting 'X' does nothing to ensure that the regulated utilities will maintain the quality of service standards; consequently, the Regulator put in place quality of standards incentive projects and enforceable penalties were put in place to support compliance.

NGC got financial incentives to carry out its SO roles (see for example, OFFER, 1998a); this included the incentives to make appropriate *long term investments* (LTI) into network capacity. The DGES also introduced the Uplift Management Incentive Scheme (UMIS) in 1994, which he later modified as the Transmission Services Project (TSP) on October 1, 1995, to enable NGC to reduce the costs of system security. NGC also introduced contracting ahead for reserve capacity, which it used to alleviate constraints (CC, 2001).

Full retail competition was structured as a phased project; as the market evolved, it became clearer that achieving a smooth transition would require the close monitoring of market operations and the modification of aspects of the market rules (Littlechild, 2001). There was a lot of regulatory input and expertise that went into the setting of Codes of Service that stipulated quality standards, which the utility firms should maintain, how the firms might deal with customer complaints, metering, meter reading and profiling. Weiner et al (1997) discusses the importance of those measures in facilitating transition to full competition in supply. There were also guidelines that were put in place regarding the separation and allocation of costs, particularly those relate to common services and the regulated and competitive businesses. There was regard about the costs that firms might incur to renegotiate contracts whenever changes were made to the regimes; and the proportion of the costs for purchasing electricity in bulk that the RECs could pass through to their captive customers. The latter was decided but with a proviso that the RECs purchased electricity economically. This economic purchasing requirement was included as Condition 5 in their operating licence. Guidelines and penalties were also laid out for timing, compliance and changeover to new systems as well as the quality of information requirements. Navarro (1996) points out that the lack of a harmonious changeover to systems between firms when a regime changes, can inhibit the speed and efficiency of service with which they deliver services. It can also affect the quality of information that the SO receives with the knock on effect being a limitation of timely an efficient remedial management. Finally at each price review, the regulator had to decide the *regulatory asset base* (RAB) upon which the utility firms earned a *fair* return. All of the issues, which this paragraph highlights, changed as the industry developed and competition was phased in.

Table 2.5.2.1 tabulates the price controls that were applied to transmission, distribution and the down stream supply segment. 'Y' denotes the ninety-five percent (95%) of costs that the RECs were allowed to pass through to their customers; it consisted of the costs for wholesale purchases, for transmission and distribution charges and for fossil fuel levy. This meant that only 5% of their costs were subject to regulatory control.

**Table 2.5.2.1**

**Price Controls in England and Wales (Oxera, 1999:71)**

| *Sector* | *Company* | *Period* | *Control* |
|---|---|---|---|
| Transmission | NGC | 1990-93 | RPI - 0 |
| | | 1993-97 | RPI - 3 |
| | | 1997-98 | RPI – 20 |
| | | 1998 - 01 | RPI – 4 |
| Distribution | RECs | 1990 – 95 | RPI + 0 to 2.5 |
| | | 1995 – 96 | RPI – 11; -14 or – 17 |
| | | 1996 - 97 | RPI – 10; - 11 or – 13 |
| | | 1997 - 00 | RPI –3 |
| Supply | RECs | 1990 – 94 | RPI – 0 + Y |
| | | 1994 – 98 | RPI – 2 + Y |
| | | 1998 – 99 | RPI – 6 |
| | | 1999 - 00 | RPI - 3 |

The table shows that setting of 'x' was evolutionary, with the firms being given tighter efficiency standards as the reform industry evolved; and suggests an expectation that competition would lead to lower costs of production. Using incentive regulation was a great success in Britain (Byatt, 1999); and in electricity, the Regulator gave NGC and the RECs appropriate incentives for them to reduce their operation costs that they could control. As a result, by 1998 / 99, the costs for transmission and distribution had reduced from 10% (in 1991 / 2) to 0.9% (Green, 1998:10)

*2.4.2.2 Transitionary arrangements*

*Initial Portfolio (IP) Contracts*

The viability of the RECs depended on curtailing the risks between purchasing power from the pool and the revenue they earned from sales in the regulated market (Armstrong et al, 1998). One of the measures that the government took to help reduce that risk was to place a set of two-way three-year and five-year contracts termed *initial portfolio* (IP) between NP and PG and the RECs in 1990 and 1993. These contracts 'were purely financial contracts, which resembled options and cash-settled futures contracts (Powell, 1993: 445). They were structured as take-or-pay agreements against British Coal, with the strike prices based on an estimate of the generators costs (OFFER, 1991:33) and associated off-take agreements with the RECs (CC, 2001). The IP contracts seemed a reasonable policy initiative, but it was not solely focused on the privatised electricity industry. The British coal industry had a large stock of relatively expensive coal compared to alternative imports available from Russia, Poland, South Africa and the United States. Therefore the IP contracts were also designed to help protect the coal industry (Armstrong et al, 1998).

*Price plan*

The price rule was designed to help the RECs sustain their off-peak night-time tariffs; in particular the viability of *economy seven—storage heating*—facilities (see OFFER, 1991). *'Table indicators: A and B'* was used to designate peak and off-peak demand. Tables 'A' half-hours occur during the periods of scarce capacity, these were predominately during the day. The PSP during table 'A' indicated half-hours were derived by adding uplift charge to

PPP. In contrast, table B indicated half-hours usually having significant surplus capacity and lower demand. Uplift was not included in the calculation of PSP during table B.

*2.4.2.3 Summary*

The Regulatory reform in the UK was based on the progress of competition in the industry. The DGES gave NGC incentives to enable it to carry out its SO roles efficiently; and used the incentive regulatory tool: the RPI-X, to regulate the monopoly businesses in distribution as well as on the transmission system. The IP contracts were a transitional measure that was put in place to enable the participants minimise operational risks during the hard landing period particularly immediately after vesting. Discriminatory pricing was adopted for the peak and the off-peak periods within the day to enhance the viability of the RECs operations particularly for their nighttime tariffs. The regulatory polices that were applied to transmission and distribution helped the Regulator to meet his dual statutory roles of protecting the final consumers by ensuring that they were charged unfair prices for the services that they consumed. And promoting competition in generation and supply as well as ensuring that the RECs remained in the business and met their *universal service obligations* (USO) whilst earning a *fair* return on their shareholders investments.

## *2.5 Implications for some of the emerging markets*

This sub-section pulls together the UKs experience and highlights the lessons learned.

By the beginning of the year 2001, England and Wales had achieved most of the objectives by which it privatised its electricity industry. For example, wholesale costs had decreased by 25% to 35%; there was approximately a 50% increase in annual network investment and about a 50% decrease in the average minutes lost on the system. In addition, the number of complaints that customers made had reduced by approximately 60%. In the generation segment, the largest company share had reduced from 48% to 12%. The share of the market that NP and PG controlled had decreased from 78% to 26% and the number of times that they set the marginal price reduced had from 90% to 41% (Littlechild, 2001). And wholesale electricity prices had decreased by approximately 40% between the time the Government announced the imminent implementation of the NETA in 1998 (OFGEM, 27/03/02; 2002:48a; 2002:48b).

The figures in the last paragraph suggest that de-integrating and competition policy can lead to significant efficiency gains. Whilst this is the evidence from Britain, there is no guarantee that it will work in other economy; that is, it is unbundling and deregulation is not one size fits all. There are some basic requirements before competition policy can lead to any efficiency gains. It is important to note that Britain had a mature institutional foundation upon which it built is regulatory reform. It also had—and still has—professional expertise that understands the technical, engineering and economics of electricity production through to supply services. As a result, they provided and supported the policy initiatives that were required to steer forward an efficient regulation of the NTS as well as the stimulation of competition into generation and supply segments, during the pool regime.

The UK also has efficient as well as a sophisticated financial market. The DGES relied on the Stock Market to obtain information on reference firms, which he used to calculate the risk-free rate, market risk premiums and beta factor for the Capital-Asset-Pricing Model (CAPM) that was used for calculating the regulated companies cost of capital.

The judiciary is sovereign. Compared to some of the other parts of the world, Britain has a functioning rule of law and enforceable property rights. It is possible that political ideology may influence public policy sometimes, but on average the judiciary processes are independent and subject to less corruption than what exists in some other parts of the world (see Mauro; 1995). Once vesting occurred in 1990, the Regulator allowed the participants to steer forward modifications to the industry Codes. All the consultations to policy issues as well as the legal re-drafting of sub-sections of the governance documents were also open to all the participants. There is a lot of confidence in the regulatory reform and processes because they were transparent, timely and independent of government influence.

### *2.5.2 Willingness to change*

Modifications of aspects of the pool regime were on going throughout the life of the regime. In1991, the DGES undertook to modify it price rules if further manipulations of capacity and prices persisted (see OFFER, 1991). He carried out his first official consultation on trading outside the pool arrangement in 1994 (OFFER, 1994). This was a very good idea and shows that flexibility is necessary requirement in regulatory reforms because it is the way through which the regimes may be aligned with market developments.

However, the DGES did not exhaust the options that he could have used to deal with the anti-competitive practices in the pool before steering forward options to trade outside the pool. For example, although he identified that the two incumbents were using capacity to manipulate capacity payment and warned that it might be necessary to remove it from the price rule, he did not take this forward. In essence, the impact of the withdrawal of capacity payment on commodity prices was not tested throughout the consultations with the industry on options to curtail anti-competitive practises that plagues the regime. The new regime: NETA, is simply a replication of the pool but one in which there is no capacity payment; there is also a penal imbalance cash-out regime. NETAs residual *balancing mechanism* (BM) trades only approximately 5% of physical deliveries. Consistent with the situation in the pool, the participants continue to use long-term contracts to hedge against volatile movements in the spot prices for over 80% of their contracted positions. This evidence suggests that the exorbitant transaction cost that that Government spent to implement NETA could have been avoided (Newberry, 2001). And the pool pricing rules could have been modified by removing the capacity payment (Bower, 2002), changing onto a *'firm'* bids and offers and a penal cash out regime for energy imbalances.

### *2.5.3 Transitory IP contracts*

The initial policy was to stimulate competition through the entry of IPPs but NP and PG were sold off in fully contracted positions both for sales as well as coal input. It seemed a good initiative since it curtailed the effects of hard landing on the generators operations and volumetric risks. As a result, it was possible for generators to focus more on learning the market rules, retraining staff, adjusting and completing compliance to new systems requirements during the early days. But it contributed to the slow emergence in the trends of prices; it also made the assessment of competition during those early days in the pool quite difficult (see Fehr and Harbord, 1993; CC, 2001).

### 2.5.4 Capacity

Generators' control capacity and simultaneously control prices in both the contract and spot markets (Bolle, 1992). Since economic and efficient dispatch is achieved when plants are scheduled in an ascending order of marginal cost, the ownership of marginal plant(s) will determine the effectiveness and the emergence of efficient competition post-deregulation. This means that the underlying objective of policy initiatives that will influence the vertical and horizontal structure in the industry, will be directed at ensuring that no generator has the monopoly over the marginal plants within each load regime.

Another way to create competitive regimes in generation is to give the generators the incentives to lock in significant proportions of their physical deliveries into long-term contracts. The threat of effective entry can also help to restrain high prices. The initial policy to have a complementary bilateral contract market with the pool achieved this. The other way to ensure that prices are as close as possible to the generators costs is if there is an equitable distribution of the marginal plants between the generators (Rudkevitch, et al, 1997). The initial policy for vesting in 1990 did not handle this properly. The Government had the option to split the ownership of the mid-merit plants between a bigger number of participants (see Green, 1999). But they chose to create a duopoly non-baseload sub-market; and expected that entry of IPPs would stimulate the level of competition that was required to sustain marginal cost pricing. On the other hand, when it became apparent in 1991 that a fairly competitive market might not be achieved within the duopoly structure (see OFFER, 1991), the decision to divest NP and PGs mid-merit plants was unduly delayed.

### 2.5.5 Competition

Apart from Sizewell B and the capacity upgrade to the Scottish Interconnector, most of the new capacity that came on-line in the 1990s was CCGTs. They initially operated as baseload plants because they had long term take-or-pay gas contracts that lasted in some cases for more than 15 years, with associated off-take agreements. Their covenants with lenders required them to run at approximately over 80% of their load factors; running as flat loads enabled them to achieve the targets. But the structure meant that NP and PG did not have effective competition for the industry's residual demand within the mid-merit portion of the LDC. Therefore the two incumbents set the marginal price in over 80% of the time in the pool, until mid 1990s, when the contractual agreements for CCGTs changed. The changes paved the way for CCGTs to operate on a two-shift basis and at 50% of their load factors; it was also then possible for them to move up the LDC into mid-merit (see CC, 2001). Nonetheless, since the price rule allowed them to earn the price that a higher cost marginal plant had set, their dominant strategy was aimed at ensuring that they were in merit whilst maximising their marginal private benefit. The main issue here was the inefficient price rule combined with the provisional requirement for declaring the operational availability of a plant.

### *2.5.6 Summary*

The experience in the UK confirms that any economy that wishes to implement any regulatory reform of its electricity sector should have the institutions and professional expertise to deal with deregulation challenges. It is also important that it is based on the economics of the initiatives and not driven merely by the politics of verbal assurances; it is also dangerous to recruit well paid consultants to write up good looking proposals.

## 2.6 *Conclusion*

This section has shown that Britain set out clear expectations, policies and transitory arrangements that helped the participants adjust into the privatised industry, before 1990. These facilitated the smooth transition to full retail competition as well as the success of regulatory regimes on the NTS. All the regulatory and competition policy initiatives that the UK applied into the industry were not one off; they changed as the industry evolved. The regulator also devised ways to deal with some of the issues that inhibit promotion of competition in generation such as cross-subsidisation, economic purchasing, price discrimination and refusal to supply. These are problems that are often more complicated and difficult to monitor, enforce and deal with when competitive regimes develop. But all of those issues were handled effectively because there were supporting institutions as well as the expertise that analysed and monitored the progress of competition and the development in the industry. This implies that an emerging market that plans its regulatory reform well in advance, combines that with the existence and developed institutions, can curtail market failure; it will also be possible for such a system to earn significant efficiency gains by deregulation.

Deregulation does not mean that the new market will be free from regulation. This section shows that regulatory oversight is evolutionary; meaning that the type of regulatory oversight that is required to ensure the reliability of supply over a safe and secure network and which lowers prices to the final consumers, changes as the reformed industry develops. This suggests that the Regulator will continue to act as the *visible hand* that facilitates cohesion between the multiple agents in the system throughout its evolution.

Economic regulation requires a lot of finance, commitment, foresight and professional expertise. Competition policies are more likely to succeed within economies that are self-sustaining and in which the governments are better placed to steer forward its own growth and development. Emerging markets that lack appropriate institutions, combined with the finance and expertise as Britain, should be aware that they may not earn efficiency gains from their electricity industry reform that will be comparable to what the UK has recorded from its reform. The regulatory reform in the Republic of Armenia failed because it was hastily organised; the country did not have the right institutions and its socio-political instability, combined with the high level of endemic corruption, were primary barriers to foreign investment into its network capacity building (see Kaiser, 2001).

It appears that some of the emerging and developing economies that are concerned about the inefficiency of their vertically integrated electricity firms but which lack appropriate institutions and professional expertise can use other forms of competition policies to improve efficiency. For example, they can use yardstick competition techniques such as regression analysis (RA) and / or the data envelopment analysis (DEA), to benchmark regional business units (BU) within the same vertical integrated utility (see Cubbin & Tzanidakis a & b). The determination of the right input measures, criteria and the treatment of differences in regional cost structures affects the credibility of a chosen frontier. After the benchmarking it is always useful to conduct an independent audit exercise as a back up and to understand each BUs location around the frontier.

Using benchmarking to monitor performance in the utility sector in England and Wales was an on-going process before the regulatory reform that was implemented in 1990. The CEGB successfully used it to monitor the performance between the Area Boards. And after 1990, the utility Regulators in water and sewerage: Office of Water Regulation (OFWAT) and (OFFER, latter OFGEM) use both RA and DEA, to as part of their regulatory tool for efficiency standards in the monopoly businesses of the industries. In an unpublished research, Rochino (2001) also finds benchmarking improving the efficiency between business units (BU) in electricity utilities in 19 OECD countries.

The emerging markets can also reduce the subsidies that they give to these public utilities, re-define the managers' objectives to reflect productivity gains. This approach was a success in Poland, where the Government improved managers' productivity before privatisation by curtailing their open-ended subsidies, made them more accountable and placed them on reward based targets (see Pinto & Van Wijnbergen, 1995). There is also a similar result from Hungary, where Ruggers and Leslie (1991) report that managers' productivity improved significantly when the Government gave the managers' greater shareholder accountability. This suggest that any Government that wishes to make its public servants more accountable will do so by stipulating enforceable penalties for poor performance. To achieve that, the Government will cease to see public utilities as an extension of their political parties, in which case, they will cease to use them for granting favours to party supporters. Therefore, it is possible for the threat of enforceable penalties for poor performance to induce public servants to improve their efficiency.

Finally, this paper shows that the entire electricity segments are integrated. That could be one of the reasons why electricity systems were never built initially as market places. The secular rise argument that generation is contestable is at best based on an incomplete argument. It requires huge sunk costs and there is high transaction costs for entry and / or exist into the industry. Moreover, the peculiar features of the product, which limits tradability, are the same issues that make pricing for commodity or transmission access very difficult in deregulated electricity systems. Consequently and consistent with Boreinstein & Bushnell (2000), we should not expect the type of competition policies that appear to have worked in telecommunications, trucking and the airline industries, to deliver comparable efficiency savings when applied into electricity systems (see also Gordon, 2000). It seems that the main issue in some of the emerging markets is the quest for what ought to be the appropriate role of the state; hence policy advisors are divided about whether government should continue to provide a *public good* such as electricity. But some of the emerging markets that lack the right institutions yet choose a hastily organised competition policy regime may have no choice but to re-visit and possibly revise those polices shortly thereafter.

# SECTION 3

**(Please note that all the tables that are included within the text are in the appendix)**

## **Quantitative Analysis of the Components of Pool Selling Price (PSP) in England and Wales' Pool:** January 1994 to December 2000

*Abstract*

*There are five quantitative parts in this section that examine the effect of the prices rule and the design of the market on the pool selling price (PSP). Fehr and Harbord (1993), Green (1994), Wolfram (1998), Wolfram (1999) and Wolak and Patrick (2001) provide empirical evidence that shows the inherent inefficiency in the pool mechanism. The essays in this section revisit the same issue; and my contribution is the examination of the commercial strategies that the Generators might have used to manipulate prices.*

*I apply different quantitative techniques to examine the patterns in the components of the PSP between January 1994 and December 2000. I find that the system marginal price (SMP) is the only variable that reduced after 1998; in relation to that, both the capacity payment (CP) and Uplift increased at a constantly increasing rate. I conjecture that the Regulators close surveillance on market operations placed a downward pressure on Generators; consequently, they offered their capacities at prices closer to their avoidable costs. The higher costs, which the National Grid Company (NGC) might have incurred to keep the system within its balanced tolerances within the day, increased Uplift. But I find that the dramatic increase in the CP was due to the increases in the loss of load probability (LOLP). Over the period, the reserve margin on the network was significantly above the notional twenty-percent that guarantees the reliability of supply; the plants used to meet demand were thermally efficient and operationally flexible. These revelations lead me to conclude that the price mechanism was inefficient; consequently, it enhanced anti-competitive practices in the pool. Therefore, I suggest that the policy advisors should be aware that the rule(s), which they adopt for setting electricity prices, might lead to socially inefficient outcomes some of the time. The implications of the results are highlighted and the potential areas for further research are identified.*

*Key words: Capacity payment (CP), England and Wales, Loss of Load Probability (LOLP), System Marginal Price (SMP), Uplift,*

# 3.1

# Introduction

### *3.1.1 Objective of the section*

The objective of this section is to use the original data that I purchased from the Energy Settlement and Information Systems Limited (ESIS), the pool Settlement Administrator to investigate the systematic patterns in the market-derived prices. The data covers the period between January 1994 and March 2001.

Fehr & Harbord (1993) examine prices setting in the pool plus NP and PGs commercial strategies. Green (1994) incorporates bilateral contracts and shows that locking in significant proportion of the aggregate supply into long-term contracts curtails high spot market prices. Tirole (1998) explains how firms in capacity constrained industries conduct themselves; therefore, are able to keep market prices above the Betrand equilibrium but they can be lower than the monopoly levels. It seems that the intuitive insight that Tirole provides, as well as the strategic behaviour of agents in capacity constrained industries, is the theoretic foundation for the studies by Helm & Powell (1992), Powell (1993) and Green (1992). They find that pool prices are higher than Generators avoidable costs; nonetheless, they conclude that the presence of a contract market and the threat of effective entry restrain high prices. Wolfram (1998) uses auctions theory to investigate NP and PGs commercial strategies and Wolfram (1999) estimates Generators mark-up. Wolak and Patrick (2001) consider the impact of the price rule on capacity and market-derived prices and Bower (2002) investigates the effect of regulatory interventions and announcements on pool prices.

This section uses different quantitative techniques to examine how the strategies that the Generators might have used affected the emergence of the price trends. It focuses on the PSP because it allows an investigation into the Generators' aggregate earnings. Plus by decomposing PSP into system marginal price (SMP), Uplift and capacity payment (CP), it provides an understanding of the impact of the Generators' behaviour not only in the determination of the commodity prices but also on energy balancing on the *national transmission system* (NTS).

*3.1.2 Overview of the developments that occurred in the industry during the 1990s*

The compulsory England and Wales' pool was incorporated in March 1990 as part of its electricity privatisation project. It was a half-hourly day-ahead mechanism through which Generators: *independent power producers* (IPPs), interconnector suppliers—Electricite de France (EdF), Scottish Power (SP) and Scottish and Southern (SS)—sold into; and the demand-side: the *regional electricity companies* (RECs) and bulk suppliers, bought electricity. The pool was not a physically located entity; instead, wholesale trading was carried out through contractual agreements for physical delivery.

At its conception stages, it was designed to facilitate only energy trading; as a result, its prices did not provide locational signals about the input (generation) or exit (demand) points on the network. It had a *non-profit* making *independent system operator* (ISO): the National Grid Company (NGC) that had the statutory right to own and operate the *national transmission system* (NTS) efficiently and in a manner consistent with the way the *Central Electricity Generating Board* (CEGB) did, under the vertically integrated regime. Therefore, it took cognisance of all the plants located across the network as well as the operational flexibility of declared plants to plan the system operations on a day-ahead basis.

There was a complementary bilateral contracts market that offered two products: the electricity futures agreement (EfA) and the contracts for differences (CfD), which the Generators used to hedge the volatile movements in the pool prices. In relation to the EfA, the Generators used mainly the *contracts for differences* (CfDs) to hedge over 80%, of their physical deliveries (MMC, 1996; OFFER, 1994). Thus the bilateral segment provided them with the opportunity to hedge their operational and volumetric risks. The distribution companies and the suppliers entered into varied and tailored agreements—that met customers' specific needs—against deliveries. There were also multiple supplier-customer agreements, which were usually tenured for 12 months that existed in the industry; these made it possible for the rates that customers paid for their demand to be based on stable price structures (Newberry & Green, 1998). Although Generators availability profiling and the influence of inter-system transmission shocks, caused price variations over the forty-eight half-hours within the day, on average, the contract market and average pool prices were closely related.

The pool fulfilled two primary roles: one was to set prices and the other was the day-ahead centralised co-ordination of the plants that were used to meet demand. The *pool purchase price* (PPP) was set on the day-ahead; and the pool-selling price (PSP) calculated after physical delivery occurred because it included the costs, which NGC incurred to resolve constraints and to keep electricity flow across the network within the Grids tolerances limit. As an efficient market maker and the *system operator* (SO), the National Grid Company (NGC) used its *Grid Ordering and Loading* (GOAL) simulator to derive the half-hourly prices at which volumes of power were traded. Between vesting in 1990 and its closure in 2001, system operations were regarded as successful because supply was reliable; the Grid safe and secure and competition led to entry of IPPs. Nonetheless, right from its inception, the pool price mechanism was considered inefficient. The price rule was criticised for failing to reflect the capacity mix of the system, which saw thermally efficient as well as the operationally flexible plants being used to meet demand as the reformed industry evolved. There was also the concern that prices did not reflect the level of competition, which the Government expected to occur from the significant entry of the independent power producers (IPPs) and the divestment of the mid-merit capacity of the duopoly incumbents: National Power (NP) and PowerGen (PG).

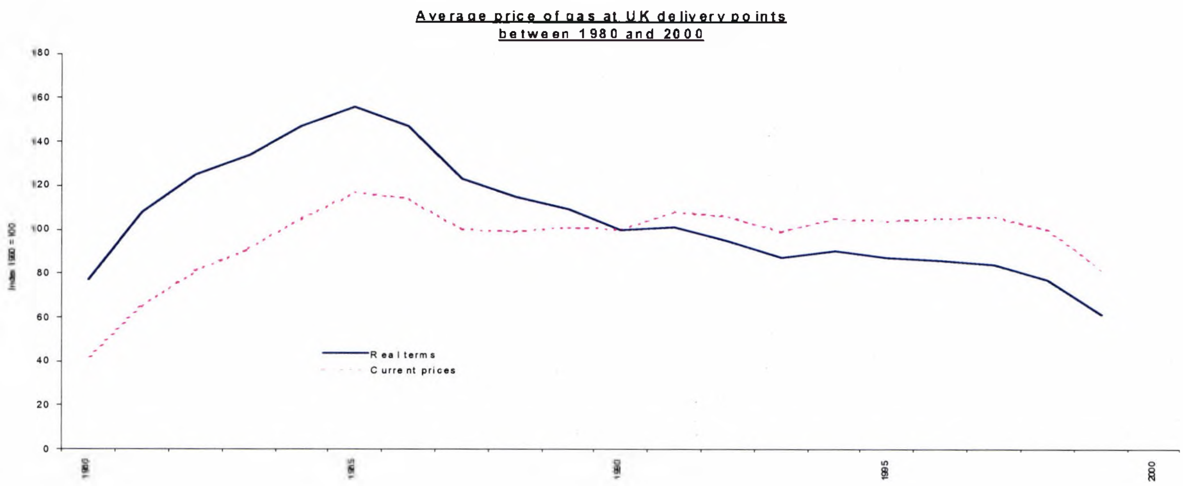
#### *3.1.2.1 Evolution of capacity mix*

The average costs for generating power through coal plants began to increase back in the 1960s (see Joskow and Rose, 1985). In the UK, a number of factors contributed to a further increase in the costs of coal-fired generation during the 1990s. The British Government changed its energy policies presumably in response to the awareness about the negative effect of global warming and climatic change. It is also possible that there was influence from the EU that might have followed its lifting of the ban on gas-fired generation in the early 1990s. In the UK, the Government steered forward initiatives that ended its nuclear expansion programme (Pollitt & Newberry, 1997) and paved the way for the rapid development of renewable sources of power generation. The British Government also initiated policies to restrict nitrogen oxides $NO_x$, sulphur dioxide $SO_2$ and dust emissions. All of these measures meant that coal-fired electricity Generators required retrofitting and procuring *Flue Gas*

*Desulphurisation* (FGD) equipment, which was the way to enhance compliance to the thresholds set for emissions.

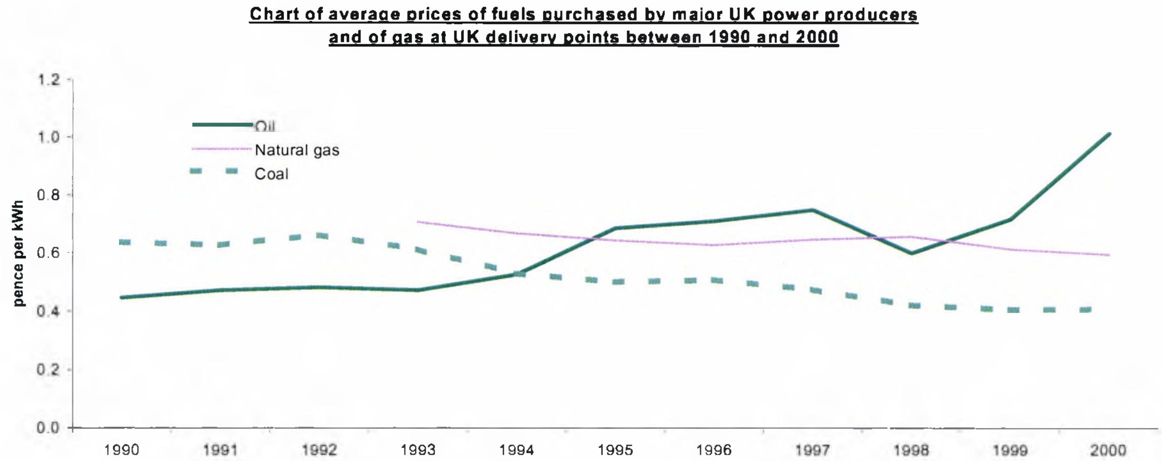
These measures were happening at a time when the cost of generation fuel was declining. Apart from that, the cost of entry for the *Combined-Cycle Gas Turbines* (CCGTs) followed a constantly decreasing rate compared to the pool prices. This lower cost entry signalled a massive entry of IPPs into the baseload portion of the *load duration curve* (LDC). Figure 3.1.1 shows that the average and real costs which the major UK power producers[13] spent to purchase gas from *notional balancing points* (NBPs) started to reduce in 1985 and continued to do so throughout the 1990s. Figure 3.1.2, is the comparative cost of generation fuels in the 1990s.

**Figure 3.1.1**




[13] These were companies that produced electricity from nuclear sources and those that existed primarily to generate electricity such as: AES Electric Ltd., Anglian Power Generation, Barking Power Ltd., BNFL Magnox., British Energy Plc., Coolkeeragh Power Ltd., Corby Power Ltd., Deeside Power., Derwent Cogeneration Ltd., Edison Mission Energy Ltd., Enfield Energy Centre Ltd., Entergy Power Group Ltd., Fellside Heat and Power Ltd., Fibrogen Ltd., Fibropower ltd., Fibrothetford Ltd., Fife Power Ltd., Humber Power Ltd., Innogy plc., International Power Plc., Killingholme power Ltd., Lakeland Power Ltd., Medway Power Ltd., Midlands Power Ltd., NIGEN, Peterborough Power Ltd., PowerGen plc., premier Power Ltd., Regional Power Generators ltd., Rocksavage Power Company Ltd., Sita Tyre Recycling Ltd., Scottish and Southern Energy Plc., Seabank Power Ltd., SELCHP Ltd., South Coast Power Ltd., South Western Electricity., Sutton Bridge Power ltd., Teeside Power Ltd., TXU Europe Ltd (OFGEM, 2001).

**Figure 3.1.2**




Data Source: OFGEM / 2000

Oil was cheaper than coal until 1994, when they converged; thereafter it became more expensive, this increase might be due to the first Gulf War (Kuwait). It then increases until 1998 when it declines but increases again in 1999. The cost of coal was fairly constant between 1990 and 1992; it decreased between 1992 and 2000. It seems that the slight increase in gas prices in 1998 (figure 3.1.2) might be the knock on effect of the capacity problems experienced at the natural gas terminals: St. Fergus and Bacton, during the summer of 1998 (see OFGAS, March 1999).

The lower costs for operating gas plants, fragmentation and the profit potentials in the market induced a massive entry of IPPs who invested mainly in *combined-cycle gas turbines* (CCGTs) (see Larsen & Gary, 1998). As the privatised electricity market evolved, the *Regional Electricity Companies* (RECs) diversified; they entered into the generation market and most of them invested jointly with IPPs in gas-fired plants. By 1995 / 1996, the total quantity of CCGT capacity on the system was approximately 17% (MMC, 1996) and by 2001, 32 gas-fired plants had direct connections to Transco's NTS; and together they consumed approximately 37% of the gas throughput (Transco, 2002).

The excessive entry increased capacity; it caused approximately 0.34% annual increases in Generators declared available between January 1994 and December 2000. In relation to the situation with other markets such as in the USA, the England and Wales' Grid was unconstrained; although transmission constraints occurred, they were transient and disappeared as the system changed within day (CC, 2001). Some Generators such as National Power (NP) and PowerGen (PG) mothballed and / or withdrew their older and inefficient plants because capacity was significantly in excess of demand. Also as the *'initial portfolio'* (IP) contracts, which the Government placed between the RECs and the Generators at vesting in 1990 expired, the RECs that had diversified opted to use gas plants, to meet demand, which were mainly located closer to their loads and within the *local distribution zones* (LDZs). The capacity mix on the system changed as the reformed market evolved; the remarkable thing was that NGC used more thermally and operationally flexible plants to meet demand as the industry matured. This suggests that if the Generators did not manipulate capacity to earn higher rents, the *loss of load probability* (LOLP) would have followed a constantly decreasing trend between 1990 and the closure of the pool in March 2001. Nonetheless, this section finds that the LOLP does not exhibit this expectation.

#### *3.1.2.2 The emergence of a competitive market*

The Electricity Act 1989, the Statute upon which the industry was restructured, gave the *Director General of Electricity* (DGES) the duty *to 'promote competition in generation and supply of electricity'*. This role was consistent with the guidelines and initial policy initiatives stipulated in the White Paper: *'Privatising Electricity'* (1988), which provided amongst other issues that the *Central Electricity Generating Board* (CEGB) be divided into three generation companies: National Power (NP), PowerGen (PG) and Nuclear Electric (NE). A duopoly non-baseload sub-market segment was created; and the thermal plants, which CEGB owned, were divided between NP and PG; whilst the NE inherited the nuclear plants. Competition was expected to evolve as the market matured and new IPPs entered into generation.

The RECs were joint owners of the transmission system that the National Grid Company (NGC) operated. There were 14 *Public Electricity Suppliers* (PES) that had supply monopoly licences called: *'franchises'*. They were responsible for supplying power to the customers that consumed less than 1MW of power. Competition between suppliers for the customers that consumed over 1MW of power was allowed at vesting; but full retail competition was phased over three years between vesting in 1990 and 1998 (Littlechild, 2001). All the suppliers, including the RECs, were allowed to enter the competition market outside of their local franchise zones to supply electricity to customers subject to a different and special licence called *'second tier supply'*, which the Regulator issued.

The Government expected that the pool would be an efficient mechanism in which the effects of the changes in the industry would be incorporated instantaneously into the market prices. That is, prices would reflect the reduction in the costs of production as well as the efficient competition between participants, as the number of IPPs increased and the ownership of marginal plants was diversified (Littlechild, 2001). Nonetheless, the pool prices failed to reflect Generators' one-year avoidable costs. The general consensus was that the way, in which the pool was governed combined with the rule for setting its prices, inhibited the significant reduction in prices. Consequently, the final consumers' could not benefit from reductions in their energy bills that were in any way comparable to what the industry participants earned (OFFER, 1999). Generators earned over 25% mark up (Wolfram 1999); in relation to the retail consumers, the large customers benefited more from the industry's privatisation (Newberry & Pollitt, 1997; Bower, 2002). These types of consumers were usually metered sites; they had access to alternative fuels, could enter into interruptible contracts with their suppliers and benefited from the economies of bulk purchasing.

The Electricity Act 1989 did not give the DGES powers to intervene in the pool; but he had limited authority through the *Pooling and Settlement Agreement* (PSA) on issues that related to the trading arrangement and governance procedures. He could modify aspects of the PSA only if the Competition Commission (CC)—formerly Monopolies and Mergers Commission (MMC)—found that the issue was against the public interest and would inhibit the 'promotion of competition in generation and supply'. He could also propose amendments to participants' licence (see for example CC, 2001).

The DGES carried out his first inquiry into high pool prices in September and published his conclusion in December 1991 (OFFER, 1991). This document highlights constraints, which the industry might face in developing efficient competition within the duopoly market that NP and PG dominated. It points out that the Grid Codes definition of the operating plant availability, combined with the inefficient attributes in the rules for setting prices might exacerbate anti-competitive practises. In April 1994, the DGES implemented the first divestitures of NP and PGs mid-merit plants as a way to reduce their control of the industry's residual demand. He capped prices during the period it took NP and PG to conclude transfer of the divested plants to Eastern in 1996. NP and PG voluntarily divested additional mid-merit plants at different times in 1999.

### *3.1.3 Methodology*

This section consists of five quantitative parts. Part 3.6 uses data exploration technique to provide a better understanding of the commercial strategies that the Generators' might have used to earn high rents. Thus it extends and complements the earlier—April 1, 1991 to March 31, 1995—analysis that Wolak and Patrick (2001) carried out. It confirms that some of the strategies, which they identified, were indeed permanent features in the regime. Part 3.7, applies Franzini & Harvey (1983) three components decomposition approach; and using the variance of the variables, it estimates the stochastic properties of the SMP and CP. Part 3.8 uses the regulator—Office of Electricity Regulator (OFFER) and Office of Gas and Electricity Market (OFGEM)—definition of a spike to investigate if SMP spikes reflect demand, supply and system security situations. Granger et. als (1979) use ordinary least squares (OLS) in structural regression analysis to examine the factors that determine within the day electricity consumption by a sample of households in Connecticut, USA. I also use the same methodology, to estimate the time-of-day (TOD) SMP in part 3.9. Finally, in part 3.10, I use the same approach to investigate the quarterly relationship between reserve margin and SMP, Uplift and declared availability.

The rest of this section is presented as follows: section 3.2 reviews previous empirical studies on the pool data; 3.3 cover the highlights in the industry's evolution as well as what we expect to see in the data results and 3.4 presents the dataset. To avoid repeating the way in which I organised the raw dataset from ESIS, (except the estimation of the unobserved properties of the SMP and CP in part 3.7), in part 3.5, I summarise that; and discuss the software used for the analysis. Part 3.6 uses the exploratory data approach, which Wolak and Patrick (2001) used, to investigate the price determination process in the pool. In part 3.7, I investigate the stochastic properties of SMP and CP. I assume that the SMP is the closest proxy in the dataset for Generators' behaviour and in part 3.8, investigate whether spikes in SMP are attributable to market conditions or merely a reflection of the Generators' opportunistic strategies. In part 3.9, I estimate a competitive TOD SMP; part 3.10, investigates the relationship between reserve margin and the components of PSP and part 3.11 discusses the public policy implications of the results from the whole empirical analysis.

# 3.2

## Literature Review

This sub-section reviews some of the earlier studies on the pool data. Fehr and Harbord (1993) pioneered empirical analysis of production and allocative efficiency in the pool. The studies thereafter appear to follow a pattern; each one fills the gap in the knowledge that an earlier investigation might have identified for further research. As a result, this literature review is presented in a chronological date order. After the review summary, there is a shared criticism on the way that these earlier investigations categorised and treated market dominance.

### *3.2.1.1 Von der Fehr and Harbord (1993)*

Von der Fehr and Harbord (1993) use auction and game theory as the basis for price and capacity setting in electricity markets. They assume that the pool mechanism is a first price, sealed-bid multiunit auction arrangement. Using data from July 1990 to April 1991, they explain NP and PGs bidding strategies; they examine the average weekly offers, which NP and PG made as well as their costs for particular plant size and fuel types. Consistent with Friedman (1971) and Kreps and Schienkman (1983), they argue that electricity trading is a two-stage game, in which the Generators set the capacity that they bring to the market during Stage 1. They compete and set prices when they arrive at the market, which is Stage 2 in the game.

They categorise three load regimes: low, high and variable; with these, they estimate the avoidable costs within each load segment. They find that during the periods of low demand, the system marginal price (SMP) is on average closer to the *short run marginal cost* (SRMC) of the least efficient Generator. The marginal price during peak periods is equal to the highest admissible price; and the mid-merit price lies between the baseload and peak. In relation to the peak and baseload, the mid-merit prices are more volatile.

They also find that NP and PG make offers closer to their estimated costs during the first 8 to 10 months of operation in the pool; but their offer patterns change around 31, March, 1991. The period coincides with the run up to the expiration of the first tranche of the Governments *'initial portfolio'* (IP) *contracts*. They also find the lower prices occurring during the first 8 to 10 months in the regime, which was a period of low demand. In contrast, higher prices emerge as the colder months, which are the periods of high demand, approached. Based on this, they conclude that Generators' commercial strategies reflect weather seasonality and demand.

They acknowledge that the structure of the industry, the repeated nature of the interaction between the agents to procure capacity, which happens with the knowledge that they will meet again in the future, inhibits pure strategy equilibrium and enhances tacit collusion. In addition to these, that the presence of an active bilateral contracts market influences entry, enhances stability of output as well as prices and provide the right signals about the industry's *long run average costs* (LRAC). Therefore, they recommend that contract markets should be an integral part of any electricity industry reform. Although they do not incorporate the effect of long-term contracting in their investigation, it is the area that they identify for further research.

#### *3.2.1.2 Green (1994)*

Green (1994) incorporates *contracts for differences* (CfDs) into his analysis and examines price and capacity patterns in the pool. He assumes that each Generator is an economic unit therefore, are in the business to maximise their aggregate earnings from participating in both the spot and futures markets. His methodology is the calculation of the difference between the Generators' total income from the pool and contracts market, and the cost that they incur to generate the energy, which they sell into the pool.

He finds that Generators do not offer their true costs of production; they also earn high rents by manipulating the capacity that they bring to the market. He warns that the commercial strategies, which the Generators adopt, may have a long-term adverse effect on production and allocative efficiency. Therefore, he concludes that the final consumers' electricity bills might not reduce in proportion to that which might be comparable to the falling costs for generation and transportation:

'... most of the larger generating sets have bid prices that are tolerably close to the estimated level of their marginal costs. Some sets have bid lower prices, but if they ran continuously, they may not have affected prices at the margin. Many of the smaller sets have bid prices that are above their estimated marginal price. A few stations have exploited local monopolies to set prices far above their marginal costs, which move them too far down the merit order to operate when the constraint does not apply. When the constraint binds, neither system operation nor the system marginal price is distorted, but the cost of electricity to the consumer is increased by the additional Uplift payments. In the short term, the price of electricity will be too high, while if plants that appear at the margin are in the wrong place in the merit order, the cost of generation will be raised. In the longer term, this gaming does not inspire confidence that the industry will be sufficiently competitive to ensure that bids are set equal to costs and provide an efficient outcome' (page 91).

#### *3.2.1.3 Wolfram (1998)*

Wolfram (1998) uses the same theoretical basis as Fehr and Harbord (1993). She assumes that the pool is a multi-unit auction mechanism; and uses monthly data for six months between 1992 and 1994 to examine the 'relationship between bid mark ups and infra-marginal capacity' (Page 13). The months are January, February, March, April, July and November. She defines mark-up as the difference between Generators' SRMC and their offer prices. And she estimates Generators cost as well as the NP and PG mark-up in the pool.

She finds NP making higher offers than PG on plants of similar costs. Consistent with Fehr and Harbord (1993), she also finds that the two incumbents offer their residual demand and infra-marginal capacity at higher prices than their estimated costs; they also submit mutually reinforcing offers. These factors lead her to conclude, consistent with Fehr and Harbord (1993) and Green (1994) that Generators collude to keep prices above competitive rates. In addition, she conjectures that the threat of effective entry, combined with the uncertainty that the Generators faces about being in-merit, can restrain high spot prices. This implies that the policy initiatives in markets that desire to have prices closer to Generators average costs should be directed at having a very efficient bilateral contracts market.

One interesting conclusion that she makes is that NP and PG might not have earned her estimated rents if the pool rule was based on discriminatory *pay-as-bid* (PAB). Whilst the rule for setting prices in a market can exacerbate anti-competitive behaviour, it seems that in reaching this conclusion, she paid little attention to the fact that price rule alone will not lower prices if there are pockets of strong monopolies across the LDC. This is worse if there are very few numbers of Generators that own the marginal plants. It also appears that Wolfram (1998) discounts the fact that generation is not contestable and the peculiar features of the industry make it difficult to price commodity and restrain monopoly power. It seems that she subsequently acknowledges these. Therefore, she warns that moving from the uniform SMP under the pool to a discriminatory PAB in the *New Electricity Trading Arrangements* (NETA) might not be socially beneficial because it will inhibit production and allocative efficiency in the residual *balancing mechanism* (BM). This conjecture is consistent with OFFER (1994), Bower and Bunn (2000), Nicolaisen (2002) and Abink et. al. (2002).

#### *3.2.1.4 Wolfram (1999)*

Wolfram (1999) obtains 'detailed' information on plants efficiency levels, uses the same data as Wolfram (1998), assumes that Generators are in the business to maximise profits and measures the difference between pool prices and Generators marginal costs. She also examines and explains the distortionary effects of regulatory interventions on the pool prices.

She finds little evidence that contracts restrain high prices; instead it is threat of entry and regulatory announcements that place downward pressures on price. She also finds that the mark up which Generators make are greater than 'zero'; however, they are closer to their true costs of production. This leads her to argue that the mark-ups are not as large as predicted by some of the earlier theories and supply function equilibrium models[14], regarding the effect of capacity manipulation on outcomes in the England and Wales' pool.

Her empirical framework is based on marginal costs (MC) assumption; but she does not provide a clear definition of the relationship between her MC estimate and the independent variables. On the other hand, the profit maximisation framework that she uses in this paper is suited for a Generator that makes offers from only one plant. In the real word, Generators often have a portfolio of plants, which is why they are able to make offers from infra-marginal generation. If one assumes that a Generator possesses a portfolio of plants, it becomes relevant to separate the private problem of the Generator and that of the industry. The marginal benefit and optimisation problem of the Generator will be directed at maximising the aggregate profits that it makes from participating both in the contract as well as the spot markets. The profit level is subject to the total capacity, which the Generator sells into the

[14] See for example the Cournot expositions by Cardell et al (1997) and Boreinstein and Bushnell, (1999); there is also the Green and Newbery (1992) supply function equilibrium analysis.

physical market. This Generator will face conflicting objectives between maximising the profit and being included in the merit order, which is its route towards ensuring that it earns the highest possible income in the pool. For whatever strategy it adopts, it always trades off its gains from losses and on average settles for the level of profit that enables it to cover its costs plus have some income to sustain its operations. The strategy that a Generator adopts to remain in the business and meet its licence conditions, are quite different from industry equilibrium requirements, which are usually directed at maximising social welfare. Wolframs current work does not quite bring out this distinction. Finally, her sample does not cover a full annual circle; consequently, the results do not provide an insight into the time, such as weather dependency of prices, capacity and mark ups.

#### *3.2.1.5 Wolak and Patrick (2001)*

Wolak and Patrick (2001) use exploratory data approaches similar to Fehr and Harbord (1993), to investigate the impact of market design and price rule on outcomes. They use half-hourly pool data between 1 April, 1991 and 31 March, 1995; categorise four load regimes and analyse trends in the components of the pool selling price (PSP): system marginal price (SMP), capacity payment (CP) and Uplift. They use mean as a measure of central tendency; standard deviation for volatility; ratio to calculate proportional changes between the variables and coefficient of variation to measure relative variability within the same variable over time. They calculate the same statistics for the total system load (TSL) within the load regimes; and estimate the production cost for NP and PG as well as their mark-ups.

Consistent with Fehr and Harbord (1993), they find that SMP is closer to their estimated costs during the periods of low demand; but that it increases in an ascending order of magnitude along the LDC. However, peak SMPs are higher than competitive rates. In relation to the SMP and Uplfit, CP is the most volatile; it is also the variable that is responsible for the tremendous variability in PSP over time. The TSL exhibits less volatility; it does not exhibit significant variation between the years across their full sample. Consistent with Newberry (1995); Bunn et al (1997) and Bunn and Larsen (1992), they report evidence that the Generators withhold capacity; and conclude that it is the cause of the high peak prices. Based on their estimate of NP and PGs avoidable costs, they conclude that the two Generators make offers in excess of their SRMC.

### *3.2.1.6 Bower 2002*

Bower (2002) uses regression analysis to investigate the effect of major regulatory interventions on prices between 1 April, 1990 and 31 March, 2002.

He finds a number of cumulative factors leading to the dramatic reduction in the pool prices after 1998. The RPI-X regulatory tool enhances best practises and efficiency. In the UK, Bower conjectures that efficiency gains were recorded in transmission, distribution and supply. As a result, the post-privatisation prices are lower than those which existed before vesting in 1990. He finds the increased competition in setting the marginal price, resulting from the divestment of NP and PGs mid-merit plants and the significant entry of IPPs. He acknowledges that the substitution of imported and cheaper foreign coal contributed to the reduction in the aggregate cost of generation. In relation to the larger sized coal plants, the improved technical efficiency and operational flexibility of CCGTs was partly the reason for the massive investments that Generators made into them.

The surprising aspect of Bower's conclusion is his conjecture that the Governments withdrawal of the gas moratorium[15] contributed to the reduction in prices after 1998. It seems that he paid little attention to the length of time that it takes for prospective entrants to secure gas licences, source debt finance, conclude network exit agreements with a gas transporter and commission a plant. Section 3.6 finds that the only component of the PPP and PSP that reduced after 1998 was the SMP; Uplift and CP increased at a constantly increasing rate. In part 3.7, I found a break in the trend of SMP occurring in week 14 in 1999, this was the run up to the initial period that the Regulator expected to implement the NETA. Only 5% of the power consumed in the UK was purchased at the PSP because the Generators locked in over 80% of their physical deliveries into long-term contracts. I learned from discussions with some of the industry participants during that time, that there was low liquidity in the market because of the uncertainties regarding regulatory risks, costs of renegotiating contracts as well as opportunistic behaviour post NETA. These meant that prices were quite low before the implementation of the new trading arrangement; the even lower prices immediately after NETA was implemented in 2001, reflected the steep learning curve for all the industry

[15] The government placed an embargo on licensing of further gas plants in October 1998, but lifted it in November 1999, after the implementation of the new gas trading arrangements (NGTA) in October 1999.

participants. That was a time during which they learned the market rules, completed the compliance to changeover to the new systems and developed their commercial strategies.

Finally, and consistent with Pollitt and Newberry (1997), Bower acknowledges that industrial consumers benefited more from the industry's privatisation; and consistent with Newberry (2002) he also expresses the view that NETA was a high transaction cost, which could have been avoided if the price rule was modified with CP removed.

### *3.2.2 Shared criticism about the analysis of dominance*

All the studies reviewed above investigated the commercial strategies that NP and PG used and the resulting mark-up. This might be because they controlled the marginal plants; as a result, it set the SMP over 80% of the time during the regime. The Reviewers simply focused on the mid-merit portion of the LDC; by doing that, they paid very little attention to the baseload, where Nuclear Electric (later British Energy) dominated. What they did was to consider that the market was a duopoly that consisted of two categories of Generators: the large and small. National Power (NP) and PowerGen (PG) were the large Generators and both utilities supplied the industry's residual demand. The Interconnector suppliers: Electricite de France (EdF), Scottish Power (SP) and Scottish and Southern (SS) and other IPPs, made up the fringe suppliers. However, transmission constraints prevented the Scottish players who were predominantly baseload plants from being strong competitors to the on-shore NP and PG (Wolak and Patrick, 2001).

These earlier studies could argue that BEs behaviour could not have had any effect on the spot market prices because it locked in most of its throughput into long-term CfDs and merely used the pool to fine-tune its contracted positions closer to real-time. In addition to that, the whole of England and Wales was one price zone in which irrespective of a Generators offer, it earned the price set by the marginal plant (Wolak & Patrick, 2001). Excluding BE as a dominant Generator based on these arguments was not reasonable because of some of the following reasons.

NGCs scheduling and dispatch of plants was based on a merit order; all the plants that Generators declared available regardless of its technical constraints were included in the unconstrained schedule and in the determination of the SMP. What really mattered in the rule for setting prices is the 'ownership' of the marginal plant. A Generator that controls a small proportion of the market could still set the price within a load regime, just by having the 'right' plant. An example is the case of the pumped hydro (Mission) which owned a very small share of the generation market; nonetheless, it had the right peaking plant and was able to set the marginal price in approximately ten percent of the time in 1995. Table 3.2.1 shows that BEs market share in 1993 / 1994 were approximately 23.2% and its output grew throughout the 1990s (see tables 2.2 and 2.3 in appendix). These figures suggest that it is possible for BE to have adopted strategies that might have maximised its private marginal benefit in the industry.

Table 3.2.1

Market Share in Generation (OFFER, July 1994 TOP)

| *Generator* | *Market share (%)* |
|---|---|
| National Power | 35.0 |
| PowerGen | 26.1 |
| Nuclear Electric | 23.2 |
| Interconnector and pumped storage | 8.4 |
| Others | 7.3 |
| Total | 100 |

Give that BE was the sole owner of the nuclear plants market, it had the power to act independently; therefore, it could have adopted commercial strategies that ensured that prices were above competitive rates. For example, if there were instances when it chose to withhold a significant proportion of its baseload capacity, then the aggregate supply curve would have shifted leftwards and resulted in an increase in the SMP. The other strategy that it could have used, and which would have been a dominant one to ensure inclusion in the merit order, was to persistently offer its capacity at £0.00/MW. This is a sensible option because it guaranteed dispatch, but given that it had the least marginal cost of production; it would earn the higher price that a marginal plant sets.

The other approach to the question of dominance is the legal application of competition law and fair-trading acts in the utility market in the UK. The precedence for defining the target market includes the assessment of the use of a product, its time of use, close substitutes and the suppliers' cost structures, to mention but a few (see Kora, 1998). Electricity variables are time, weather and volume dependent; and different plant types are used along the LDC. The prices over the 48 half-hours within the day are also different. These suggest and support the treatment of each load regime along the LDC as a distinct sub-market. This is a familiar approach and the basis upon which target markets have been defined in competition cases. The EC made it in Hoffman-LaRoche and Co. AG. Commission case 85/76, United Brands Co and United Brands Continental BV. v. Commission case 27/76 and in NV Nederlandsche Baden-Industrie-Michelin v. Commission case 322/81. In electricity, the Competition Commission used this same approach to assess the economic impact of the then proposed mergers between NP and Southern Electric Plc; as well as in that of PG and the Midlands Electric Plc (MMC, 1996a & b).

Based on the above, the earlier studies could have considered BE a dominant baseload player and assessed its commercial strategies and the resulting mark-up.

### *3.2.3 Summary*

There are a number of empirical studies on the pool data and all investigated the strategies, which NP and PG used in the pool as well as their mark-up. They find that pool prices are time, volume and weather dependent. They also report that the dominant incumbents used capacity manipulation to earn higher rents.

# 3.3

# Implications of earlier studies and expectations for the present investigations

This part uses the implications from the earlier studies that are reviewed in part 3.2 and some of the main developments in the industry and knowledge acquired by interacting with the industry participants, to formulate generalised propositions about the patterns that I expect to see in the data. The Office of Gas and Electricity Market (OFGEM), the Department of Trade and Industry (DTI) and the National Grid Transco (see http://www.OFGEM.gov.uk; http://www.dti.gov.uk; http://www.nationalGrid.com/uk/), document some of these developments in communications such as consultation documents, reports and press releases.

Electricity prices reflect usage; therefore, they are time, volume and weather dependent. The implication from the earlier studies by Fehr and Harbord (1993) as well as in Wolak and Patrick, (2001) is that each load regime is a separate sub-market.

*Expectation 1*

*Prices will increase in an ascending order along the LDC; it will be higher during the day than at night time; similarly it will be higher during the table A than B indicated half-hour. Over weekdays, and weekends, there will be higher prices occurring during the week. There will also be higher prices during the winter months because this is when the highest demands in the year occur. Consumers need more power to operate heating facilities in the winter, than during the summer months (see Granger et al, 1979). Comparing the situation during the summer and winter, prices are lower during the autumn and spring months; this is because the weather is mild during these two shoulder seasons and consumers do not need to run cooling and / or heating facilities as such.*

In April 1994, the DGES introduced the Uplift Management Incentive Scheme (UMIS), which was later modified and it became the Transmission Services Project (TSP) on 1 October, 1995. This initiative was aimed at reducing the costs that NGC incurred to balance the system. Also several Working Groups were set up; they met regularly to review the progress, problems and ways to improve capacity and energy balancing of the NTS. For example, two groups that considered issues to do with transmissions services included the Reactive Power Market Working Group; Reactive Power Market and Transmission Users Group. The DGES also gave NGC incentives to procure reactive power in an economic and efficient manner (see OFFER, December 1998).

*Expectation 2*

*As a result of some of the initiatives summarised in the last paragraph, this study expects that the within the day balancing costs would follow a constantly decreasing trend between 1994 and the closure of the pool on 26 March, 2001. If this holds, then the cost that NGC incurred to resolve constraints within the day, for start-ups and availability payments would also reflect an evolutionary trend. This implies that the Uplift in the dataset will exhibit the same pattern.*

National Power and PowerGen had significant excess capacity in 1993 / 1994; however, coming into 1994 / 1995, they withdrew, and in some cases mothballed most of that excess. There was also a report of a sudden outage of nuclear plant that occurred in January and in December 1995 (MMC, 1996; CC, 2001).

*Expectation 3*

*Relative to the other years in the sample, reserve margin will be lowest in 1995.*

The DGES investigated the causes of high pool prices several times between 1991 and 2000. He concluded at each inquiry that prices failed to reflect comparable decreases in the costs of input, the demand and supply as well as the security situation on the network (see OFFER, 1991; 1992; 1993; 1994; 1996; 1998; and OFGEM 1999).

*Expectation 4*

*The dataset will reflect highly volatile price variables. This suggests that the price series will not be Gaussian; instead, they will be more predictable and skewed.*

*Expectation 5*

*The spikes will not reflect demand and supply situations. In relation to the other year's in the full sample, prices will be highest in 1995 because of the lowest levels of reserve margin in that year. Also within 1995, prices will be highest during the months of January and December; this can be attributed to the sudden outage of the nuclear plants.*

Newberry (1995; 1998 & 1999); Powell (1993) and Bolle (1992) show that it was difficult to obtain reliable assessment of the industry's performance during the early years of the regime. This suggests that the underlying level average of prices as well as the capacity patterns will emerge slowly as the industry evolved.

*Expectation 6*

*There will be an evolutionary emergence of the underlying level average of the pool prices.*

Evans and Green (2003), Bower (2002) and OFGEM (2002a & b), show and argue that cumulative factors contributed to the dramatic reduction in pool prices after 1998. The broad influential issues that they raise include the increased number of owners of the marginal plants, which resulted from the massive entry of IPPs and divestment of NP and PGs mid-merit plants. The Regulator particularly considers that the announcement of the model to trade outside the pool signalled an imminent change to the trading arrangements; and the prices reductions were simply a reflection of the effect of the threats of a regime change. This suggests that the announcement restored to some extent, the inefficiency in the price mechanism.

*Expectation 7*

*There will be a break in the path of the underlying pool prices after 1998.*

Agents in short-run capacity constrained industries use capacity (supply function) to manipulate prices (see Tirole, 1998; Powell, 1993). In electricity, OFFER (1991) highlights some of the capacity manipulation strategies that the Generators might use to earn higher revenues from the pool. It confirms that they might rely on capacity manipulations outside the SMP determination processes, to increase pool prices. This is consistent with the later prediction that Bunn and Larsen (1992) made about the relationship that might exist in the market between the LOLP and prices.

*Expectation 8*

*In relation to the SMP, Uplift and Capacity payment will exhibit very high variability with more incidences of irregular values and high relative variability. In addition to these, the multiplier effect of the loss of load probability (LOLP) on the difference between the values of loss load (VOLL) in the calculation of CP will make the latter the most volatile component of the PSP.*

*Expectation 9*

*Consistent with Wolak and Patrick (2001), this study expects to uphold CP as the most volatile component of the PSP.*

# 3.4

# The Data

I purchased the original half-hourly pool data 'bundles' in 2001, from the *Energy Settlement and Information Systems Limited* (ESIS), the pools settlement administrator. The data contained aggregated monthly 'bundles' that covers the period from 1 January, 1994 to 27 March, 2001; there are a total of ninety months within seven and a quarter years. Table 3.4.1 describes and 3.4.2 summarises the variables in the original dataset.

Table 3.4.1 Listing of variables in the full dataset: 1 January 1994 to 27 March 2001

Total number of Observations: 126,814
Total number of variables: 13

| Variable name | Description |
|---|---|
| SDD | Settlement Day Date |
| PERIOD | Settlement Period |
| SMP | System Marginal Price (£/MWh) |
| LOLP | Loss of Load Probability |
| PPP | Pool Purchase Price (£/MWh) |
| PSP | Pool Selling Price (£/MWh) |
| AB | A/B indicator |
| TL | Transmission Losses (MWh) |
| DEMAND | Gross Demand (MWh) |
| DECLARED | Declared Availability (MWh) |
| REDECLARED | Redeclared Availability (MWh) |
| ACTUAL | Actual Availability (MWh) |
| YEAR | Year (1994, 1995, ... 2001) |

Source/year: ESIS/2001

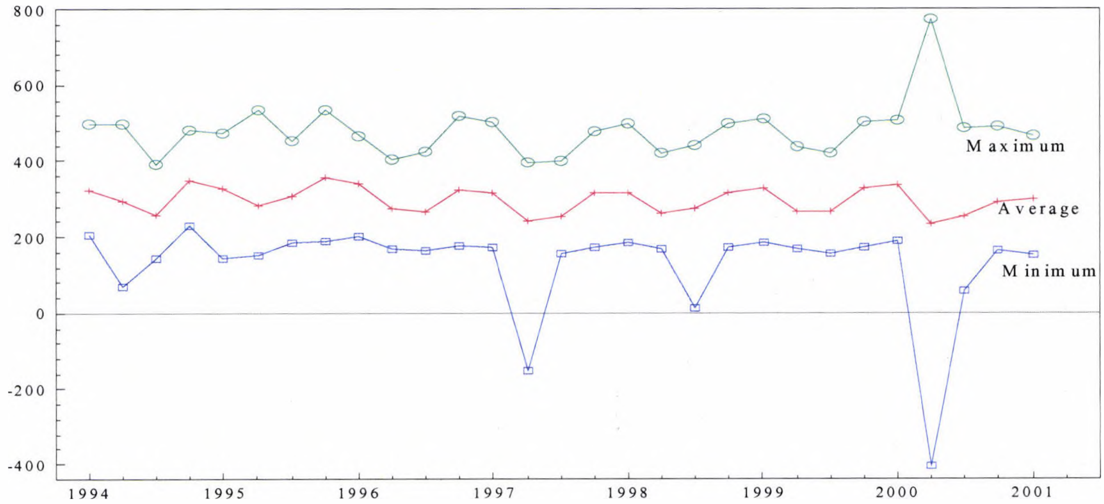
Table 3.4.2 Summary of variables

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| smp | 126814 | 20.95564 | 12.6308 | 0 | 836.1632 |
| lolp | 126814 | .0010963 | .007755 | 0 | .4309897 |
| ppp | 126814 | 23.74751 | 25.24796 | 0 | 1108.118 |
| psp | 126814 | 25.01574 | 28.17157 | 0 | 1180.517 |
| tl | 126814 | 296.793 | 65.35718 | -405.926 | 771.398 |
| demand | 126814 | 16590.99 | 3083.837 | 8798.115 | 25923.2 |
| declared | 126814 | 23737.88 | 3094.484 | 15198.66 | 32661.79 |
| redeclared | 126814 | 23145.57 | 2929.025 | 15212.28 | 31239.62 |
| actual | 126814 | 22869.04 | 2959.603 | 15007.52 | 31215.05 |

The dataset contains some negative transmission losses (TLs) as shown in table 3.4.2, which figure 3.4.1 shows to have occurred in 1997 and 2000. All but 1 of these negative TLs, occur on 14-15 May, with longer runs recorded on 25 and 26 May, 2000.

**Figure 3.4.1**

**Quarterly Transmission Losses**



This is ESIS' explanation for the negative TLs. Although NGC simulated an unconstrained schedule on the *day ahead* (D-1) and SMP calculated ex-ante; Generators could profile their capacity within the day (D). Within the day changes in prices are partly due to availability profiling. The final pool prices which are published 28 days after the day is based on a revised unconstrained schedule, which incorporates changes in Generators availability within the day (PSA, 1990; OFFER, July 1994).

The process is as follows, ESIS published a provisional *pool purchase price* (PPP) by 4 p.m. on the day-ahead: $(D-1)$. The 28 days lag between the D and the publication of a final settlement price gave ESIS the time to receive any revised unconstrained schedules and to collate all data handling errors including any disputes that Generators' might have raised. Once they have this information, they carried out a *settlement re-run. As a result, the final* PSP and PPP, which was published on $(D+28)$ might be significantly different from the provisional (D-1) figure (OFFER, 1991:24).

In theory and practice, negative TLs are not strictly speaking accurate. But ESIS confirmed that the case in those half-hours as they occur in this data is attributable to technical constraints and faulty meter readings. Such errors were corrected if the producers' affected raised what were called *'dispute queries'*. They explained to me that since the affected producers' for those half-hours did not raise any disputes, the associated settlements were based on these negative values.

### *3.4.2 Summary*

The data 'bundles' consists of 13 variables; and a preliminary summary revealed that it contains some negative TLs that occur in 1997 and 2000. Although negative TLs are theoretically incorrect, they often occurred in the pool due to technical reasons such as faulty meter readings. The settlements were done based on the negative TLs because the affected Generators failed to raise any dispute queries about them.

# 3.5

# Organisation of Data

This part explains how the original dataset was organised for the essays in this section. Since it is a heteroskedastic dataset, in which variables are weather dependent, it is possible that the error term in the estimations will be serially correlated; in addition to that, the models will violate the constant variance assumption. Therefore, the estimated parameters may be biased. Part of the objective in this part, is to explain how I intend to curtail the effect of heteroskedasticity and multi-collinearity in the models. Finally, it summarises the software that is used for organising the data, data exploration and estimations.

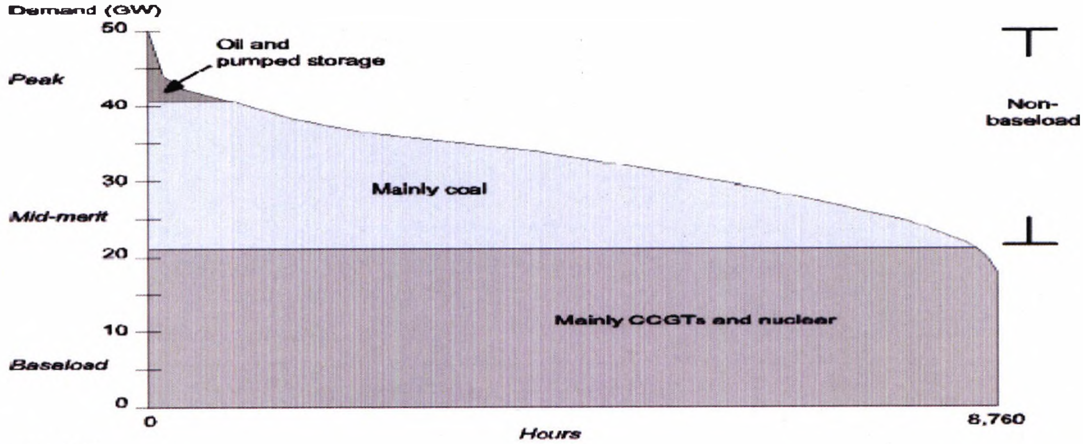
The monthly bundles were first collated into annual sub-samples that run from January to December each year. The basis for this classification is to enhance an easier investigation of the weather seasonality of the variables over the four seasons in a year. It also seems a sensible base upon which to investigate the systematic or consistent patterns in the variables over the years.

### *3.5.2 Load regimes*

This section adopts MMCs (1996) categorisation of a one-year LDC in England and Wales, which figure 3.5.1 shows. The vertical axis calibrates the load regime whilst hours are represented on the horizontal axis. There are three load regimes, classified as follows: baseload, 0MW to 22MW; mid-merit, 23MW to 42 MW and peak, above 43MW. Using three categories for load regimes in this current study is consistent with the earlier work that Fehr and Harbord (1993) carried out. It gives me the opportunity to examine patterns and systematic changes in prices and capacity within and between the regimes over time.

**Figure 3.5.1**

**Categorisation of Regimes on the Load Duration Curve**




Source: MMC (1996)

It is important to note that the LDC presented above is specific to the capacity mix in England and Wales over the period covered by the data. The energy balance and capacity mix on any network will determine the plant types that will be used to meet demand within each portion of the LDC.

Nuclear plants and combined cycle gas turbines (CCGTs) were the plants that were used to meet baseload demand. In relation to the other plants along the LDC, Nuclear had the highest capital costs, but it was the most flexible and had the least marginal cost (MC). Since dispatch was based on the least cost, Nuclear was scheduled first. As demand increased along the LDC, the more expensive plants were called on; so long as it was still within the baseload, CCGTs were next in line after the Nuclear plants. Once the load was between 23MW and 42MW, mid-merit plants would be required; and at this time, the Coal plants will be called on. It does not mean that CCGTs and Nuclear would not run any more; instead, Coal plants will also be part of the number of plants running. The most expensive plants such as the Oil and Pumped-hydro were used when demand was at its peak; which was above 43MW. Again during the peak periods, a combination of all the plants that were used along the LDC were running.

One thing to note was that CCGTs could have operated as mid-merit plants during the 1990s, but during the early years of the regime, their debt covenants meant that they had to operate above eighty-percent of their load factors. They also had long term take-or-pay gas contracts that lasted up to fifteen-years ahead; these requirements meant that the Generators could not construct their CCGT plants to run on a two-shift basis, but just as flat loads.

### *3.5.3 Quarters*

The price, capacity and demand variables in the dataset reflect weather seasonality. Categorising the data into the four seasons in the year: winter, summer, autumn and spring, is a reasonable basis; this is because the usage pattern for electricity differs remarkably between these periods (Granger et al, 1979). And since prices are volume driven, implying a positive correlation between gross demand and prices one should expect to find a significant difference between prices during the four seasons in the year. These support the approach in this study to conduct quarterly and monthly analysis. Therefore, this section relates prices and capacity to weather seasonality within each year, through quarterly and monthly analysis.

It defines quarters as Q1 = January to March, Q2 = April to June, Q3 = July to September and Quarter 4 = October to November. Each quarter and month in the analysis consists of all the half-hourly observations within that sub-sample

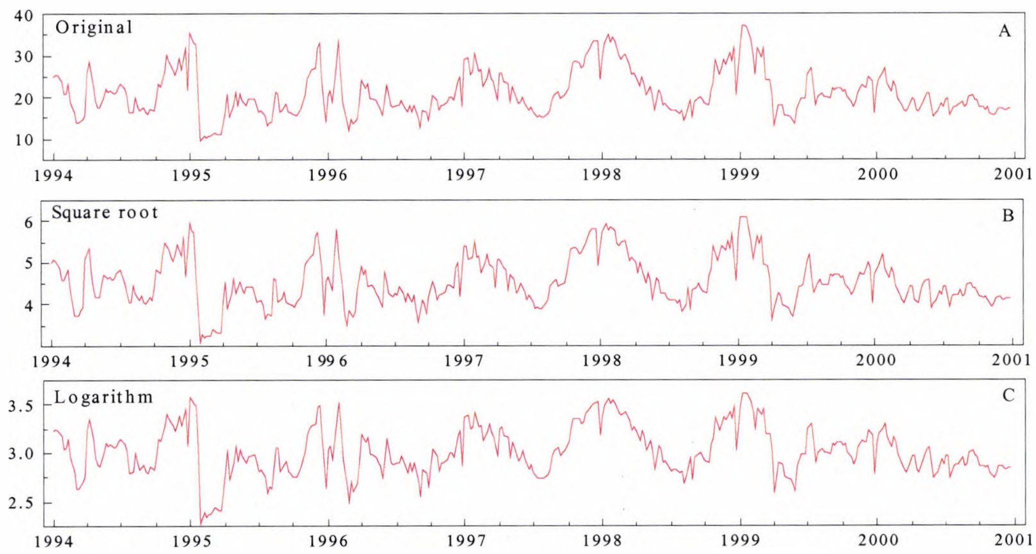
### *3.5.4 Transformation of variables*

My preliminary analysis and survey of the data involved the testing of the variables for normality, checking the strengths of the relationship between the variables, running simple OLS estimations and an examination of the residuals. Tables 3.5.1 reports the annual skewness and kurtosis test (*sktest*), which on average rejects the null hypothesis of normality. The histogram of transformation of the variables (see Hamilton, 2004) showed also that on average, all the variables except Uplift are normal in Square Roots (*sqrt*); Uplift is nearly normal as identity. Section 3.7 assumes that the original series can be obtained as a summative of the relationship between the components of the variance of the series (Makridakis et. al, 1998) and uses a *maximum likelihood* (ML) technique. Therefore, the estimations in that section are based on the logarithm of the series.

Figure 3.5.2 plots the series for the original and the transformed square root plus the logarithm of the SMP. The series shows a trend-cycle pattern in the series; there is also evidence of seasonality and irregular values. But the series: graph A, appears more stable in the transformed square root: graph B and logarithm: graph C.

**Figure 3.5.2**

**Weekly average time series of SMP: Original, square root and logarithms**




For the structural OLS regression estimations, the study uses the transformed square root of all the variables except Uplift, hence defining the normal functions as: $f(x) = sqrt(x)$. It will use identity for Uplift; therefore, defining its normal function as $f(x) = (x)$. The ML estimation is based on the logarithms of the variables.

### *3.5.5 Treatment for multicollinearity and heteroskedasticity*

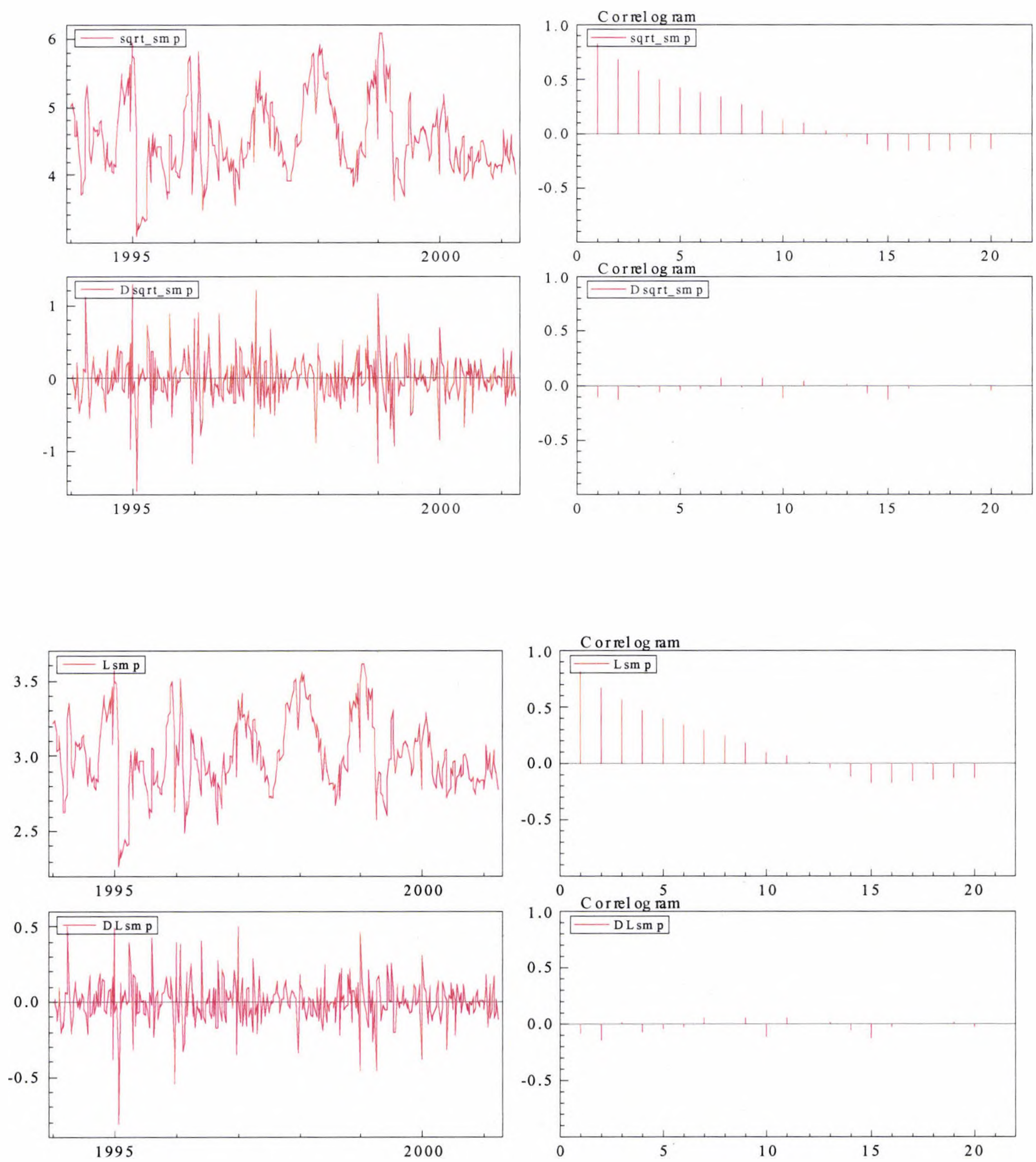
The variables used in setting pool prices were functionally dependent on each other. This immediately suggests the possibility that these variables will move together; in addition to that, it is likely that the same factors will cause them to change together and in some cases, it may be in the same direction.

The time dependency of electricity prices and demand within the day also suggest that the variance of the observations may not be constant; and the error term in our models may be serially correlated. Given the inherent features of this dataset, which includes being heteroskedastic, it appears that just carrying out a non-robust OLS estimation will give biased estimates because there will be a violation of the classical assumptions for regression analysis. But then it is important to point out that even after modifications are made so as to enhance the reliability of the results and the inferences that may be made from them, the errors in the specified model might still be serially correlated. Also there might be evidence of heteroskedasticity. In such a situation, where these violations are evident following modifications, it will not reflect violation of the classical assumptions or the mis-specification of a model, anymore. Instead, it will be a reflection of the nature of the dataset. This is a common problem that is encountered in economic time series; and it is because the variables that go into the market equilibrium state, usually depend on each other (Neter et. al, 1996).

Dropping serially correlated and redundant variables is usually the way to remedy multicollinearity. However if the strength of correlation between independent variables in a regression analysis lies between the range of $-0.70$ and $+0.70$, it does not cause difficulties in models (see Lind et. al, 2001; Ott, et al, 1992). Therefore, I will not use independent variables in any of the models if the strength of relationship exceeds $\pm 0.70$. If one assumes that the value, which a variable takes at a point in time depends on some factors that occur in a preceding period, using lagged values which for example is obtained by differencing, minimises the effect of serial correlation. This is evident in the plot of the square root and logarithm of the weekly average SMP series that figure 3.5.3 shows. In relation to the 'no difference', the serial correlations reduce dramatically in the transformed logarithms of the first differences; this is the same thing with the square roots of SMP.

## Figure 3.5.3

## Graphs and Correlograms of logarithm and square root of weekly average SMP in levels and first differences




An electricity price during each half-hour depends on the type of plant, the volume of electricity dispatched and the avoidable costs of production. A past event such as long-term constraints influences the price at which a Generator will be willing to offer its capacity in the market. But on average, the price during each period within the day does not depend on the one during a preceding period. Given this circumstance, using the differences (first or second) of the series is appropriate if it is only an academic exercise that is intended; it may lead to biased inferences if it is a policy prescription that is required. Since part of the objective for this study is the prescription of policies, it uses a combination of approaches to minimise the effect of heteroskedasticity on our OLS estimates.

Each half-hour in the dataset, consists of energy, capacity and Uplift prices; there are also the LOLP, VOLL, demand and Generators': declared, redeclared and actual, availabilities. Since these variables take on different values during each half-hour within the day, it is possible to assume that each half-hour consists of a cluster of 10 variables. This also means that one can assume that each electricity day consists of 48 distinct and separable half-hourly clusters. This is consistent with the fair-trading approach for defining a target sub-segment in electricity markets.

The assumption of clusters that is made in the last paragraph allows for the consideration of independence of the observations between clusters. It also provides the foundation upon which one can use a structural modelling approach for the models. Combining that with the White (1980)-corrected standard errors; that is, a robust variance estimator, it will enhance the reliability of our estimates, which come from samples that are not independently distributed (see Huber, 1967; Wooldridge, 2002; Carroll et. al, 1998).

When robust standard error estimator is used in a model, the point estimates are the same as that which will be obtained from a conventional method of calculation; and the ANOVA values do not change. What changes though, are the standard errors and confidence intervals of the estimates, which are adjusted. The *F* test in robust specified models 'becomes a Wald test based on robustly estimated variance matrix'; 'and the variance of the residual varies by observation' (STATA; page 337) (STATA, V. 8—R: 331-341 & U23: 270-276; also see appendix). Citing Hansen et al (1953), Williams (2000) notes that the use of robust variance in estimations has been in sample survey literature since 1953. But its use in applied statistical models gained prominence in the 2000s (Gutierrez, 2003).

### *3.5.6 Software*

The study uses STATA v.8 and Structural Time Series Analyser, Modeller and Predictor (STAMP). STATA is used to organise the original data 'bundles' into sub-samples such as annual, table indicators, load regimes, weekly, monthly and quarterly. It is also used for running the OLS estimations. The properties of the unobserved components of the SMP and capacity payment (CP) are modelled in STAMP.

### *3.5.7 Summary*

The study will carry out sub-sample investigations on prices, demand and capacity patterns in the pool. The OLS regressions are run on STATA v.8; the standard errors are White (1980)-corrected and period, which consists of 48 half-hours within the day, will be the cluster variable. The structural univariate models are done in STAMP.

The next sub-section starts with the exploratory analysis of the dataset.

**Table 3.5.1A**

Annual Skewness and Kurtosis tests for normality

| | | | | —Joint | — |
|---|---|---|---|---|---|
| *Year* | *Variable* | *Pr(Skewness)* | *Pr(kurtosis)* | *Adj chi2(2)* | *Prob>chi2* |
| 1994 | ESM | 0.000 | 0.000 | 37.67 | 0.0000 |
| | SMP | 0.000 | 0.000 | . | . |
| | CP | 0.000 | 0.000 | . | . |
| | Uplift | 0.000 | 0.000 | . | . |
| | declared | 0.000 | 0.000 | . | . |
| 1995 | ESM | 0.000 | 0.000 | . | 0.0000 |
| | SMP | 0.000 | 0.000 | . | . |
| | CP | 0.000 | 0.000 | . | . |
| | Uplift | 0.000 | 0.000 | . | . |
| | declared | 0.000 | 0.000 | . | 0.0000 |
| 1996 | ESM | **0.718** | 0.000 | . | 0.0000 |
| | SMP | 0.000 | 0.000 | . | . |
| | CP | 0.000 | 0.000 | . | . |
| | Uplift | 0.000 | 0.000 | . | . |
| | declared | 0.000 | 0.000 | . | 0.0000 |
| 1997 | ESM | 0.038 | 0.000 | 54.67 | 0.0000 |
| | SMP | 0.000 | 0.000 | . | . |
| | CP | 0.000 | 0.000 | . | . |
| | Uplift | 0.000 | 0.000 | . | . |
| | declared | **0.085** | 0.000 | . | . |
| 1998 | ESM | 0.000 | 0.000 | 41.48 | 0.0000 |
| | SMP | 0.000 | 0.000 | . | . |
| | CP | 0.000 | 0.000 | . | . |
| | Uplift | 0.000 | 0.000 | . | . |
| | declared | **0.704** | 0.000 | . | . |

| Table 3.5.1A (cont.)<br>Annual Skewness and Kurtosis tests for normality | | | | | |
|---|---|---|---|---|---|
| | | | | —Joint — | |
| *Year* | *Variable* | *Pr(Skewness)* | *Pr(kurtosis)* | *Adj chi2(2)* | *Prob>chi2* |
| 1999 | ESM | 0.054 | 0.000 | . | 0.000 |
| | SMP | 0.000 | 0.000 | . | . |
| | CP | 0.000 | 0.000 | . | . |
| | Uplift | 0.000 | 0.000 | . | . |
| | declared | 0.000 | 0.000 | . | . |
| 2000 | ESM | **0.310** | 0.000 | . | 0.0000 |
| | SMP | 0.000 | 0.000 | . | . |
| | CP | 0.000 | 0.000 | . | . |
| | Uplift | 0.000 | 0.000 | . | . |
| | declared | 0.000 | 0.000 | . | 0.0000 |

# 3.6

# Investigating the price determination process in England and Wales' pool

*Abstract*

*This part examines the impact of the strategies that Generators' might have used to offer capacities in the pool, on the components of the pool-selling price (PSP): system marginal price (SMP), Uplift and capacity payment (CP). The results show that the mean values of SMP reduced significantly after 1998, in contrast the Uplift and capacity payment increased dramatically over the same period. The increases in CP and Uplift suggest the evidence of capacity manipulation because the results mean that the costs, which the National Grid Company (NGC) spent to balance the system increased after 1998. Since the security position, demand and supply and reserve margin on the system during that period do not justify the increases in balancing costs, this paper conjectures that the Regulators close surveillance of the daily market operation, might have placed a downward pressure on the Generators' opportunistic behaviour. Consequently, it might be that the reduction in SMP confirms that after 1998, the Generators began to offer capacities at prices that were closer to their true costs. Since the capacity manipulations outside the SMP process is very difficult to detect, one possibility is that after 1998, they re-defined their commercial strategies and began to use the CP and Uplift more, to earn higher rents.*

*Keywords: Capacity Payment, Electricity Market, England and Wales, System Marginal Price, Uplift*

## Introduction

The England and Wales' pool started in March 1990 as a supply-only arrangement; and in 1993, the *Director General of Electricity Supply* (DGES) introduced the demand-side participation. The mechanism was based on a uniform payment, compulsory membership; and no trades were contracted outside it.

Part of the initial policy for the industry's de-integration and deregulation was the creation of three generation companies and an *independent system operator* (ISO). The non-baseload plants were divided between two Generators: *National Power* (NP) and *PowerGen* (PG) whilst the third company, which was Nuclear Electric (later British Energy), inherited all the nuclear plants. As a result, right from vesting in 1990, the market was very concentrated. NP and PG became the dominant portfolio Generators; they owned the non-baseload coal and oil plants, supplied the industry's residual demand and set the marginal price 80% of the time during the regime (OFGEM, 1999; Littlechild, 2001; MMC, 2001). NP and PG competed with the small and fringe suppliers, such as the other *independent power producers* (IPPs) and the interconnector suppliers—*Electricite de France* (EdF), Scottish Power (SP) and Scottish and Southern (SS). However, transmission constraints limited the possibility of fair competition between the Scottish baseload suppliers and their on-shore counterparts (Wolak and Patrick, 2001). The industry transited from a duopoly to an oligopoly that had many IPPs; the combination of mandatory divestment of NP and PGs mid-merit plants in 1994, their voluntary divestment in 1999 and entry of IPPs contributed to this.

NGC was the fourth company that was created at the time the industry was vested in 1990. It was charged with the responsibility to carry out the role as the ISO; consequently, it performed the dual roles as a *transmission owner* (TO) and a *system operator* (SO). It had the statutory monopoly to transport power over the *national transmission system* (NTS), which consisted of 400KV and 275KV lines. NGC also co-ordinated the centralised scheduling and dispatch of the varied capacity mix across the system in an ascending order of marginal costs. This process is technically referred to as *merit-order*. As the energy and capacity balancer on the transmission system, it organised and procured transmission services[16] directly from the

[16] These were reactive power, frequency responses, reserves—spinning and non-spinning—and black start capabilities.

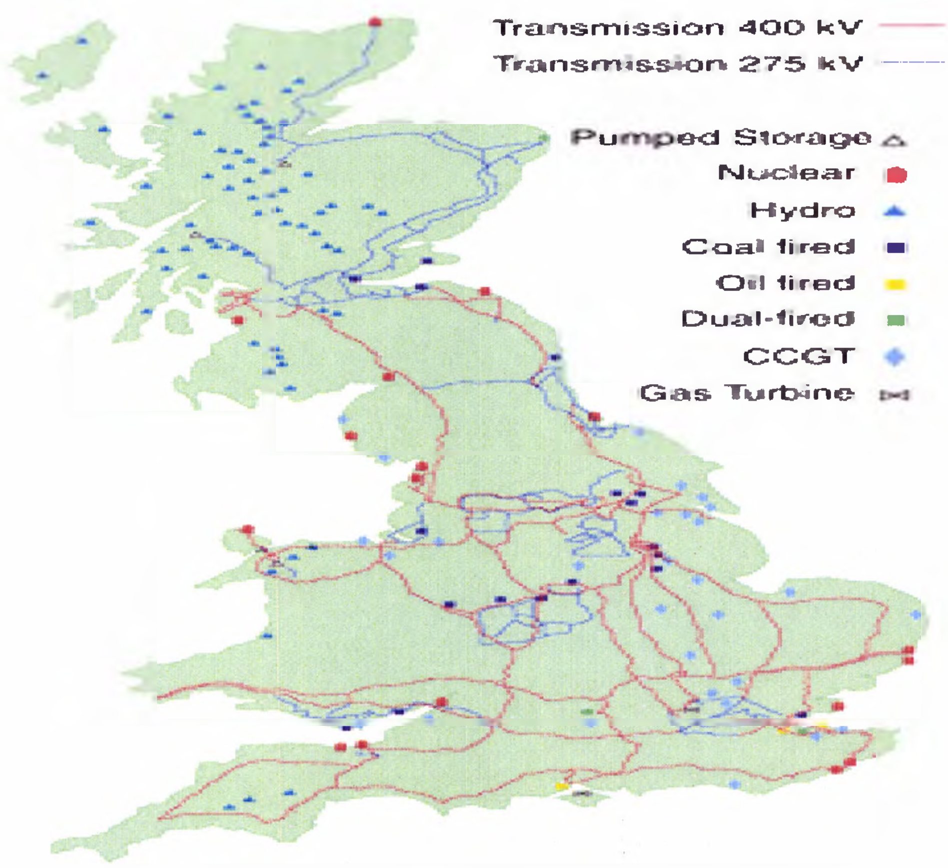
Generators that had the facility to offer those services. If within the day, that the security and safety of the system demanded a re-direction of the flow of power across the network, it *constrained 'on' or 'off'* plants, as appropriate. These were some of the ways by which the NGC facilitated within the day reliability of supply, over a safe and secure NTS. In summary, the pool served two related purposes: (1) it was a mechanism for setting the *48 half-hourly* commodity prices and (2) it facilitated physical trading.

The pool buy price, the *pool purchase price* (PPP) and its sell price, the *pool selling price* (PSP), were based on a half-hourly determined *system marginal prices* (SMP). This paper methodologically based on the earlier work by Wolak and Patrick (2001) focuses on the PSP. This approach gives us the opportunity to investigate the systematic and consistent patterns in Uplift. As a result, and consistent with Wolak and Patrick (2001), we gain an insight into the Generators aggregate earnings; thereby enabling this study to highlight the policy areas that may facilitate efficient competition in commodity trading. It will also enhance the regulation of capacity and energy balancing on the transmission system, in electricity markets.

PSP consisted of three elements: a commodity, capacity and system security costs; it was not location specific; therefore, it did not provide any signal for short-term remedial management initiatives or even for appropriate *long-term investments* (LTI). However, there was a location specific *transmission network use of system* (TNUoS) charge, which was applied to capacity on the transmission system. TNUoS was based on the *long run incremental cost* (LRIC) of meeting a marginal increase in demand. Figure 3.6.1 shows that generation was predominately located in the North and consumption in the south of England. Using the *LRIC* methodology meant that exit charges reflected the cost of infrastructure required to meet demand; consequently, and in relation to the North, TNUoS charges were higher in the South of England.

**Figure 3.6.1**

**Location of Generation Stations across England and Wales' Grid**




Source: http://www.british-energy.co.uk/pet/images/industry_map3.gif

At the design stages of the pool, Littlechild (2001) notes that the Government was concerned about the participants' potential exposure to the operating and volumetric risks that characterises electricity markets. Consequently, they placed *an initial portfolio (IP) contract for differences (CfDs),* which they backed against the British coal and structured as take-or-pay coal contracts, between the Generators and the *Regional Electricity Companies* (RECs). The additional costs of these contracts were passed on to the final consumers through the 'Y' factor in the distribution price control, which under the licensed-based regulatory reform, was calculated as the retail price index minus an efficiency rate depicted as 'X' that gave the formula: $RPI - X + Y$ . The objective of the IP CfDs was to protect the British coal industry (see Armstrong et al; 1998). Nonetheless, it seems that the wider aim of the Government was to give the Grid Users the opportunity to curtail the effects of hard landing, which usually characterises industry reforms. It seemed also that it was intended to allow the Generators to hedge against volatile pool prices whilst the financial segment in the privatised industry developed.

The financial segment of the pool developed and offered two products; one was the *contracts for differences* (CfDs), which the industry participants used mainly for long-term contracts that lasted up to several years ahead of real-time. The other was the *electricity futures agreement* (EFA) that they used predominately for short-term transactions. It was possible to tailor EFAs to suit the preferences of the contracting parties; hence they were usually standardised. Nonetheless, there were downsides to the EFA sub-market: it lacked liquidity and the instrument was often difficult to trade (Bower and Bunn, 2000). In contrast to EFA, Generators locked in over 80% of their physical deliveries into CfDs (Green and Newbery, 1998[17]; CC, 2001).

Only 5% of electricity demand in the UK was purchased at the pool price: PSP. Suppliers entered into various forms of long-term contracts with their customers, which were usually tenured for one year. Consequently, the price that the customers paid during the period of the contracts did not change with the volatile movements in the pool prices. This implies that the gross demand in the dataset will not be responsive to within the day variations in the PSP. On the upside, given the significant volume of physical capacities that the Generators locked into long term contracts, the industry output was stable and contract prices reflected *long run average costs* (LRAC) (Newberry, 1995).

At the design stages of the pool, it seemed that the processes for setting commodity price, and LRIC capacity charging on the transmission system combined with an active futures market, would deliver production and allocative efficiency (Littlechild, 2001). That is, the Government thought that the pool design was very robust to the extent that it might be difficult for Generators to systematically manipulate capacity and prices. But Fehr and Harbord (1993) reveal that the Generators started manipulating capacity and prices as early as the first 10 months into the operation of the pool. There are other empirical studies on the pool data including the DGES' investigations into the causes of high pool prices that uphold

[17] 'The Electricity Industry in England and Wales' in Competition and Regulated Industries, edited by Dieter Helm & Tim Jenkinson.

that prices remained above competitive levels throughout most of the period in the pool. This was presumably because the Generators sustained their anti-competitive practices (see for example OFFER, 1991; OFFER, 1994; OFFER, 1999; Green, 1994; Wolfram, 1998; Wolak and Patrick, 2001). The contribution of this current study is that it uses the Grid Codes definition of the operating availability of a plant, and the process for setting the system marginal price (SMP), to show the loopholes that the Generators relied upon to manipulate capacity and prices.

The rest of this paper is structured as follows: 3.6.2 reviews the earlier literature on the pool data and formulates expectations of what I expect to see in the data. 3.6.3 summarises the methodology for the study, 3.6.4, the loopholes in the Grid codes definition of the operating plant output and prices setting and 3.6.5 the results of the data exploration. Discussion of the results is in 3.6.6 and in 3.6.7 I conclude this sub-section.

### *3.6.2 Literature review & expectations*

Games and industrial organisation (IO) theories provide insights into capacity and price setting in electricity markets; it also provides expositions about the potential behaviour of the multiple agents in the system. The general rule in capacity constrained oligopoly markets is that prices will always be above the Bertrand equilibrium but may be below monopoly levels (see Tirole, 1998). Prices' setting in electricity markets is similar to the two-stage price and capacity setting in *supergames*, which Friedman (1971) proposes. Kreps and Scheinkman (1983) show that tacit collusion makes the non co-operative equilibrium in such markets inefficient some of the time; consequently, one should expect that the aggregate cost of generation may be above competitive levels (see Fehr and Harbord, 1993). Tirole (1998) and Laffont and Tirole (1993) prove the theoretical basis for collusive practises in the oligopoly markets. Borrowing from their exposition and applying it to the electricity market, Generators' will determine capacity first; thereafter, they will set prices and the actual physical delivery of the commodity will be the last event that occurs. They will use capacity, which is their supply function (see Green and Newberry, 1992) to manipulate prices. Bolle (1992) uses this argument to analyse price and contracts in the England and Wales' electricity market. Also Wolfram (1998) finds that the incumbent duopoly Generators: National Power (NP) and PowerGen (PG) submit mutually reinforcing offers into the pool.

In commodity trading, Generators engage in repeated interactions to procure capacity with which they inject power into the Grid; but they do so with the knowledge that they will meet again. Information about some aspects of their operations such as their registered capacity (which was called *Generator Registered Capacity* (GRC) in the pool), and the variables, which the system operator (SO) uses to estimate gross demand are usually domiciled in the public domain. Over these repeated interactions to procure capacity, the skilled traders' learn the market rules and how to estimate the industry's load curve with some accuracy. Once they are able to do so, they can forecast the residual demand that their firms will supply to the industry; and offer such capacities, including their infra-marginal capacity at prices that are above competitive rates. Mobility of labour between firms facilitates knowledge transfer about competitors' cost structures as well as their potential strategies. NP and PG were formerly under one management; therefore, even after the industry's reform, the staff in both companies had near perfect knowledge about each other's possible strategies.

The inefficient outcome that the last paragraph highlights is not peculiar to electricity markets. Indeed, it is also reported in some of the other capacity constrained industries that use a repeated interactive mechanism such as auctions, to allocate capacity and to determine price (see for example, Aron, 1998; Porter & Zona, 1993; Cramton & Schwartz, 2000). But the peculiar feature of electricity, which limits it tradability, worsens efficient pricing in electricity markets.

Apart from *economy seven: storage heaters*, electricity cannot be stored in appreciable quantities. The generation, transmission and distribution processes happen as though the entire system is one single vast machine. (2) The Grid *injections* and *off-takes* must be balanced on a real-time basis with the Grid kept within its acceptable energy, frequency responses and voltages tolerances. (3) Demand is inelastic and predictable over time. (4) Electricity prices are time, quantity and location (geographic) dependent. They are also very volatile, and mean reverting. In addition, within the day, inter system transmission and capacity constraints, worsens the incidences and the duration of price spikes. (5) Once a plant is up and synchronised, there is no dramatic and continuous variability in its avoidable cost. In practice, agents are more likely to offer the same price for quantities across the whole period within day. This is partly why spikes occur in blocks of identical prices. (6) The variables that go into price setting depend on each other and marginal cost of production varies throughout the 48 half-hours in the day. How did these features affect price behaviour in the pool?

Green and Newberry (1992) pioneered theoretical analysis on price setting in the pool. They assume that the Generators: NP and PG offer step supply functions; and use supply function equilibrium analysis to investigate competition between the two incumbents. They find that the duopoly structure of the industry exacerbates monopoly power; therefore, they recommend that 5 equally sized firms may be required to bring prices down to competitive levels. In a similar analysis, Bolle (1992) argues that the increase in the number of players will lead to competitive prices. These imply that pool prices ought to decrease after the divestment of the NP and PGs mid-merit plants in 1996 and 1999. In contrast, to Bolle (1992) and Green and Newberry (1992), Rudkevitch et al (1997) finds that prices can be above competitive rates even with an equally sized number of firms. This suggests that the Generators' in England and Wales might have earned abnormal profits even as the number of owners of the marginal plants increased.

Fehr and Harbord (1993) were the first to investigate NP and PGs profits from the pool. They used information on costs of fuel and the thermal efficiency of plants to estimate the firms' production costs. They found that the monopoly power contributes to the high rents that NP and PG earn. They also find evidence of a change in NP and PGs commercial strategies towards the expiration of the first tranche of the IP CfDs.

They categorise three load regimes: baseload, mid-merit and peak; and find prices increasing in magnitude from base load through to peak. They explain that pure-strategy equilibrium leads to the least efficient Generator setting prices during the base load; the full capacity of the Generator with the least cost is dispatched first and the most expensive plant supplies the residual demand. I borrow the expositions from Fehr and Harbord (1993) and Wolak and Patrick (2001), to explain this.

I assume that there are two Generators in the market: A and B, if Generator A makes the lowest offer, and B the highest, B will set the marginal price. But A's full capacity will be dispatched first and B will be left to supply the residual demand. So both Generators will receive the same price per MW of power dispatched. This suggests that baseload price setting is analogous to Bertrand competition. During peak periods, no one Generator can supply the industry's aggregate demand alone. They will then work on the expectation that both of them will be called; as a result, they will offer their estimated residual demand at the highest possible price. What will happen in this case is that the highest admissible offer will set the marginal price. The two Generators' face conflicting objectives, which derives from them wanting to maximise their marginal private benefit; as well as being called to supply. Here again, the level of profit is volume dependent, thus the higher the quantity sold, the more the profit. Since no single Generator will want to price itself out of merit, each player is likely to offer its residual capacity at a price that is slightly lower than its perceived competitor. In relation to the mid-merit and peak, baseload prices will be lowest; those at the peak, highest and the mid-merit will lie between the baseload and peak and it will exhibit the highest relative variability.

*Expectation 1*

*This study expects to find prices increasing in order of magnitude from the baseload to peak; and in relation to the bas load and peak regimes, the highest variability will occur during the mid-merit.*

Wolfram (1999) uses three different methodologies to estimate NP and PGs mark up. Consistent with Von der Fehr and Harbord (1993), she estimates marginal costs as consisting of the costs of fuel and the thermal efficiencies of plants. She finds that NP and PG earn over 25% mark up; in addition, to that, the regulatory oversight restrains high price. This suggests that the pool prices would have been higher than the observed values. There were three key regulatory interventions in the pool: (1) the price caps placed between 1 April 1994 and 1996 (2) the divestment of plants in 1996 and 1999 and (3) the announcement of the model for the new trading arrangement in 1998.

*Expectation 2*

*Compared to the other years in the full sample, prices will be stable during the cap period.*

*Expectation 3*

*Prices will fall after 1998, and the lowest prices and volatility will occur in year 2000; however, technical problems from the test runs meant that it was re-scheduled and eventually implemented in the first quarter in 2001.*

Wolak and Patrick (2001) analyse the impact of price rule on production and allocative efficiency between 1 April 1991 and 31 March 1995. In contrast to Von der Fehr and Harbord, they categorise four load regimes; but consistent with Fehr and Harbord (1993), they find prices increasing in magnitude from the lowest to the highest load regime. They also find that capacity and Uplift cost increases PSP; CP is the most volatile component of the price setting variables. They report relative stability in the *total system load* (TSL) across all years in their sample, which leads them to conclude that it is possible to carry out reliable forecasts of the TSL than the other price setting variables. They also find that NP and PGs earn excess profits.

*Expectations 4*

*Relative to the other variables, TSL will be stable across all years.*

*Expectation 5*

*Capacity payment will be the most volatile component of the PSP. This is because of the multiplier effect of the LOLP on the difference between VOLL and the SMP, in the formula for calculating CP.*

### *3.6.3 Methodology*

Wolak and Patrick's (2001) exploratory data analysis is the methodological basis for this present study. It focuses on using the loopholes in the Grid Codes definition of operating plant availability and the process for setting the SMP, to explain the commercial strategies that the Generators might have used to manipulate prices in the pool. Therefore, it provides an insight into the ingenuity, which is the rule for setting prices induced on Generators.

Consistent with MMC (1996), we define three load regimes as *full total system load* (FTSL) between 0MW and 22MW; 23MW to 42MW and above 43MW. Categorising load regimes into three is consistent with the earlier work by Von der Fehr and Harbord (1993). We use quarterly analysis to investigate price response to weather seasonality and table indicators to compare peak and off-peak patterns.

### 3.6.4 Operating Plant Output and Prices Setting

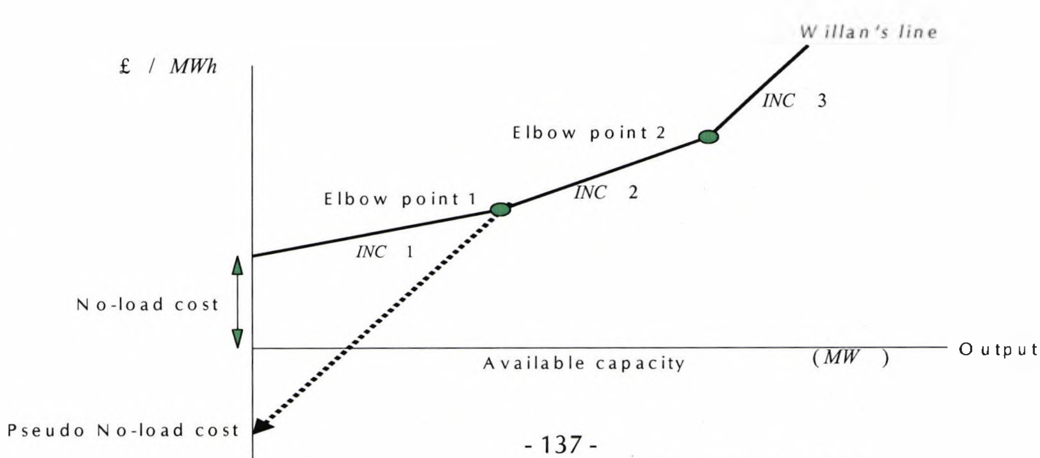
This section discusses the Grid Codes definition of the operating plant availability and determination of SMP. It highlights the loopholes that Generators might have used to earn higher revenues.

#### *3.6.4.1 Operating plant availability*

The offers that Generators made into the pool consisted of five elements: a no-load prices; a start-up price and up to a maximum of three incremental offers. 'The no-load and incremental prices define the price of the operating plant at different levels of output once the plant has been synchronised. These parameters were chosen as being representative of the way that the engineers characterise the cost curves of thermal stations. The resulting cost curve for a genset is known as a *'Willan's Line'* and might look like that shown in [figure 3.6.2]. The points where the first incremental switches to the second and the second to the third are known as elbow points' (OFFER 1999:7).

**Figure 3.6.2**

**Incremental Bids and Prices of Operating Plant per Output**




(Offer, 1991:27; Offer, 1999:7)

'If a genset has more than one incremental price, a *'pseudo no-load price'* is calculated for each extra incremental price. These pseudo no-load prices can be, and often are, negative. During the calculation of SMP all no-load and start-up costs are allocated pro rata to scheduled generation in Table A periods. If, for any Genset, the sum of the start-up cost plus the total no-load cost is negative, then the set's Table A price will be lower than its Tables B price' (OFFER, 1991:28).

Generators' could submit £0/MWh for start-up costs and incremental offers. There was also the *Greater than or Equal (GE) inflexibility marker*, which allowed them to specify the operational flexibility as well as the minimum output that kept a plant stable. GE marked plants did not set the SMP. The structure of the offer prices and the GE marker system presented some flaws.

A Generator that offered capacity at £0/MWh was more likely to be in-merit; also it earned the higher price, which the marginal and most expensive plant used to meet demand during that half-hour sets. This allowance to make 'zero' offers enhanced lack of costs bidding; given that it was a legal provision, it was not possible to claim predatory pricing against a Generator that persistently offered its capacity at such a price.

A plant is warm and ready to dispatch if it runs at its minimum stable generation; and operating at plants' stable generation helped to curtail the high costs associated with cold start-ups. Moreover, intermittently switching on or off a generation plant may damage its components. GE marking was included in the Grid Code to enable the Generators to sustain the viability of plants that might be called on intermittently within the day. But even in the early days of the regime, the Regulator identified that it was a potential instrument that the Generators might use to manipulate prices (see OFFER, 1991). The type of capacity manipulation that they used the GE marker to do was to mark the cheapest plants as inflexible; once that was done, SMP would be based on the more expensive plants.

### *3.6.4.2 Table A and B indicators*

One of the issues that the Government addressed at the design stages of the pool was how the RECs might sustain their off-peak tariffs; in particular, their viability in relation to consumer facilities such as *economy seven* storage heaters (see OFFER, 1991). Two types of half-hours were included in the price setting: *table 'A'*, depicted periods of high-demand when the system had low levels of excess capacity. T*able' B'*, were the half-hours of low-demand; in relation to A, there was more excess capacity during table B. On average, table A occurs mainly during the daytime and table B at nights. Another distinguishing feature between the two table indicators was in the way that the SMP was calculated plus the assignment of Uplift costs. SMP during table A, included the fixed element of a Genset, such as the no-load price and its start-up price; but the price during table B was set at the 'marginal operating price of the marginal plant' (OFFER, 1991:26). In addition to these, Uplift was included in the PSP during table A but it was excluded during table B.

*3.6.4.3 Price setting*

NGC used historical *seasonal normal demand* (SND) from pumped storage; they used large customers that consumed 250MWh and the non-daily metered (NDM) sites, to forecast gross demand. Although the DGES introduced demand-side participation into the pool in 1993, NGC did not include their bids when determining the SMP, because they were usually over £50/MWh and therefore, higher than the peak SMPs.

NGC derived the industry supply function by stacking the Generators' pairs of offer and quantity bids in an ascending order of marginal costs; thereafter, it used its *linear optimisation Generator ordering and loading* (GOAL) model to derive an unconstrained least-cost schedule, which it used to dispatch plants on the day. However, within day technical, transmission and capacity constraints might require NGC to re-direct the flow of power across the network; in such situations, NGC used out of merit plants, which for the particular half-hour might not be the least cost, to meet demand.

*Pool Purchase Price (PPP)*

All the Generators that supplied power into the pool received a PPP, whilst the demand-side, paid the PSP to take off power from the pool.

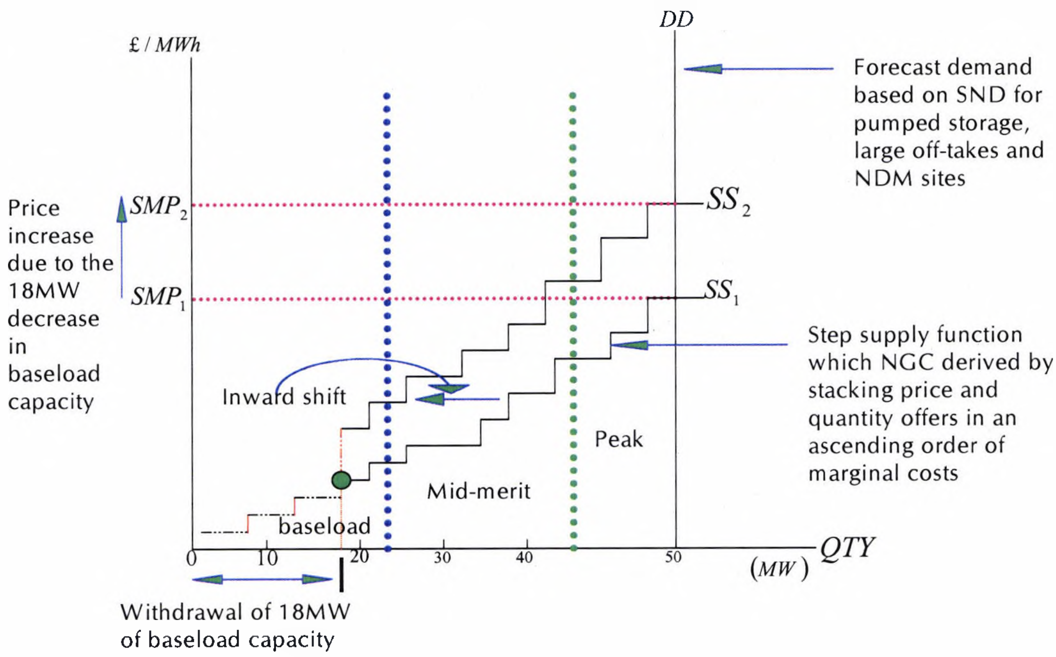
The PPP was calculated as:

$$PPP = SMP + CP \qquad (1)$$

Where the system marginal price (SMP), was derived at the intersection of the aggregated industry's step supply function and NGCs forecasted gross demand as shown in figure 3.6.3.

**Figure 3.6.3**

**Determination of SMP and the Impact of Capacity Withdrawals**




Source: authors drawing

The SMP determination process is analogous to price determination in free markets (see Sloman, 2002). Therefore, economics theory can provide a useful insight into the way in which the multiple agents in the system might have behaved in the pool. It can also help in explaining the consistent systematic patterns in the price movements and the factors that might have caused the shifts in the demand and supply curves over time.

Capacity Payment (CP), determined at the day-ahead stage, was calculated as the product of the *loss of load probability* (LOLP) and the difference between the *value of lost load* (VOLL) and SMP:

$$CP = LOLP\,(VOLL - SMP) \qquad (2)$$

$LOLP$, was an index that changed through several orders of magnitude across the 48 half-hours within the day. It ranged between 'zero' and 'one' and reflected the reserve margin on the network. The latter was calculated as a ratio of the difference between declared availability minus the highest demand, based on the *average cold spell* (ACS) conditions (NGC, 2000) and divided by the declared availability.

*The Value of Loss Load (VOLL),* was an imputed value, which Kwoka (1997) noted, was 'derived from utility planning models'. At the design stages in the regulatory reform, capacity payment (CP) was designed to reflect scarcity; to create the right incentives for the Generators to manage their load efficiently including re-location where feasible; and to make appropriate investments in remedial management infrastructure.

*Pool Selling Price (PSP)*

The demand-side paid PSP to buy electricity from the pool. It was calculated as the sum of the PPP and Uplift, with the latter being the necessary costs that NGC incurred to maintain safety and security on the transmission system:

$$PSP = PPP + Uplift \qquad (3)$$

*Uplift,* which was determined *ex-post,* was an unhedged component in the pool price. It included costs such as those for availability payments; start-up, constraints and for procuring transmission services such as the reactive power, reserves (spinning and non-spinning), frequency responses and black start.

*3.6.4.4 Commercial strategies that the Generators' might have used to manipulate prices*

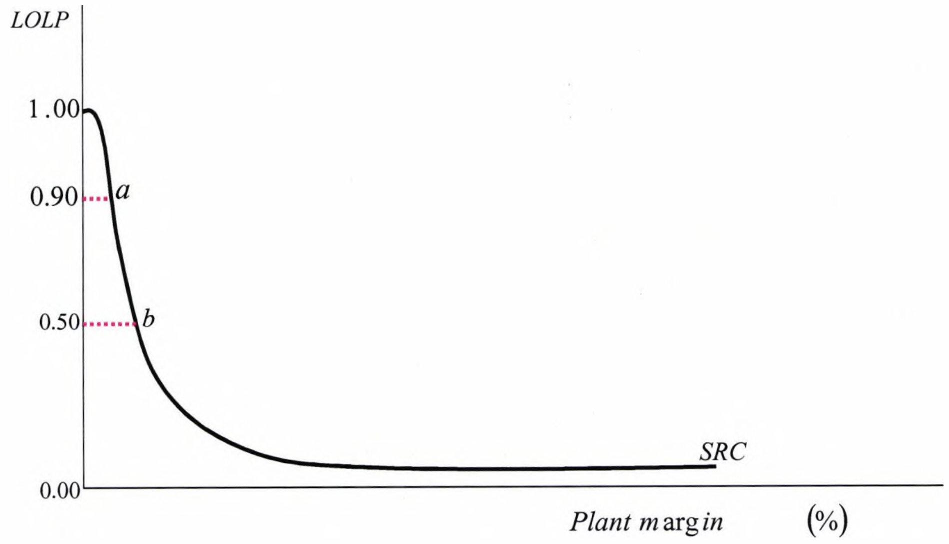
Next, I examine the strategies, which the Generators might have used to manipulate capacity to earn higher values of the prices setting variables.

*Loss of Load Probability (LOLP) / Capacity Payment (CP)*

Figure 3.6.4 is a simplified illustration of the downward sloping relationship between LOLP and reserve margin. I introduce a *security reliability curve* (SRC), which is the line that connects co-ordinate reserve margin and LOLP points. The figure shows that LOLP decreases with increases in reserve margin, along the SRC.

**Figure 3.6.4**

**Relationship between LOLP and Reserve Margin**




Reserve margin was 'zero' whenever the LOLP was equal to 'one'. At such time, the system does not have any reserve capacity to meet increments in demand; therefore it will rely solely on the imports possibly from neighbouring markets. Any system that uses the same calibration for its LOLP, when reserve margin is 'zero' commodity price can take astronomical values. This was the case in California in 2000 (see Sweeney, 2002). If LOLP is 'zero' reserve margin can also take abnormal values. Although the pools upper limit was set at 'one', in practice, it was not observed because, as the diagram shows, the SRC never touches the reserve margin axis. Instead, it is asymptotic to the horizontal axis.

If we take a point such as $a$, the reserve margin is very low; hence, the LOLP is high at 0.90. This indicates that the system does not have sufficient capacity to cope with increased levels of consumption such as a violent surge in demand within the day. Suppose on the short run, Generators increase declared availability, it will increase reserve margin. If the system now moves along the SRC to a point such as $b$ , there will be both a direct and an indirect effect of the change. LOLP will decrease by 0.45 and signify that there is now an increased level in the reserve margin.

This illustration shows that the Generators relied on using capacity to manipulate LOLP; and consistent with this, the data set for this study reveals in table 1' that declared and LOLP are positively correlated across all the years as well as the quarters between January 1994 and December 2000. The strength of the relationship varies across all the quarters and years. Focusing on the quarterly statistics and in relation to the first and last quarters which contains the peak winter months, there are high correlations seen in the second and third quarters in 1995, 1996, 1998, 1999 and 2000. There are two possible explanations for these. 1. Fewer plants are usually scheduled to meet demand during the off-peak periods; in some cases, many of these plants will be inflexible base load generators. Therefore, LOLP will reflect the greater probability of not meeting demand in the event of say a technical problem leading to a plant outage. 2. The other sensible conjecture is capacity manipulation. Installed and the GRC grew year on year and most of the plants that came on line were smaller and more flexible, meaning that they were easier to start. Given this development in the capacity mix across the system, it seemed that the Generators manipulated their declared availability; thereby reducing reserve margin and increasing LOLP.

**Table 3.6.1'**

Annual and Quarterly Correlation between LOLP and Declared Availability

| | | | *Full sample* | | *Quarter 1* | | *Quarter 2* | | *Quarter 3* | | *Quarter 4* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Nos. | Variable | Lolp | declared | Lolp | Declared | Lolp | Declared | Lolp | Declared | Lolp | Declared |
| 1994 | 17520 | Lolp | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | |
| | | Declared | 0.1427 | 1.0000 | 0.1190 | 1.0000 | 0.2176 | 1.0000 | 0.2083 | 1.0000 | 0.2598 | 1.0000 |
| 1995 | 17520 | Lolp | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | |
| | | Declared | 0.2140 | 1.0000 | 0.2559 | 1.0000 | 0.2220 | 1.0000 | 0.3812 | 1.0000 | 0.2260 | 1.0000 |
| 1996 | 17568 | Lolp | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | |
| | | Declared | 0.2071 | 1.0000 | 0.2568 | 1.0000 | 0.0739 | 1.0000 | 0.4613 | 1.0000 | 0.2114 | 1.0000 |
| 1997 | 17520 | Lolp | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | |
| | | Declared | 0.1197 | 1.0000 | 0.0830 | 1.0000 | 0.1498 | 1.0000 | 0.2027 | 1.0000 | 0.1619 | 1.0000 |
| 1998 | 17520 | Lolp | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | |
| | | Declared | 0.1312 | 1.0000 | 0.1784 | 1.0000 | 0.2385 | 1.0000 | 0.1657 | 1.0000 | 0.1802 | 1.0000 |
| 1999 | 17520 | Lolp | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | |
| | | Declared | 0.0454 | 1.0000 | 0.1424 | 1.0000 | 0.2738 | 1.0000 | 0.4413 | 1.0000 | 0.1992 | 1.0000 |
| 2000 | 17568 | Lolp | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | | 1.0000 | |
| | | Declared | 0.1133 | 1.0000 | 0.1634 | 1.0000 | 0.3447 | 1.0000 | 0.4237 | 1.0000 | 0.1627 | 1.0000 |

Notes: The statistics is based on half-hourly observations of the variables. The difference between declared availability and gross demand is the proxy for reserve margin. The quarters are based on annual quarters, thus Q1 = January, February and March; Q2 = April, May and June; Q3 = July, August and September; and Q4 = October November and December. LOLP ranges between '0' and '1'.

The indirect effect of manipulating declared availability in price setting is seen in the capacity payment. Remember that CP is derived as $CP = LOLP\,(VOLL - SMP)$. If SMP and VOLL remain constant, a decrease (increase) in LOLP will cause CP to reduce (increase). This formulation gave LOLP a multiplier effect on the difference between VOLL and SMP; therefore, made LOLP the most important variable in setting CP. The Generators used capacity to manipulate LOLP; as a result, earned higher values for CP. Increases in CP increased the values of SMP, PPP and PSP. This is consistent with the non-linear relationship between LOLP and prices, which Bunn and Larsen (1992) conjectured. Since higher CPs increased PSP and PPP, the data set for this section would show that these three variables are highly correlated.

*System Marginal Price (SMP)*

The equilibrium price derived at the intersection of the demand and supply curves in competition is the short run marginal cost (SRMC) of production. In electricity, the SRMC is equal to the avoidable cost of generating power into the pool. In theory, the SMP was deigned to increase (decrease) with decreases (increases) in Generators declared availability. So how might Generators have manipulated SMP?

Suppose there is a withdrawal of 18MW of base load capacity from the baseload in figure 3.6.3. Since demand is inelastic, the demand curve will remain constant, but the supply will shift to the left. For instance, imagine that it causes a movement from $SS_1$ to $SS_2$. The new supply curve will intersect the constant demand at a position above the original equilibrium, and result in a higher SMP of $SMP_2$.

A marginal change in capacity causes SMP, LOLP, CP, PPP and PSP, to vary. If we go back to figure 3.6.3, once the 18MW of capacity was withdrawn, reserve margin decreased; and in figure 3.6.4, the system would move leftwards along the SRC, increasing the probability that the system might not meet increases in demand within the day. Increases in LOLP will increase CP; the knock-on effect would be an increase in the PPP and PSP respectively. The magnitude with which LOLP increases when capacity is withdrawn, say due to a sudden plant outage will depend on the location of the system along the SRC. That is, whether the LOLP on the system is close to 'one' or 'zero'. If it is closer to one, depicting very low reserves, a change in capacity will lead to a higher change in LOLP. But if the system is closer to 'zero' meaning there are no threats to the system operations, then a change in capacity will change reserve margin, but it may not really lead to a significant change in prices.

One other thing to note is that Generators are more likely to carry out the capacity withdrawals discussed in figure 3.6.2 within the base load portion of the load duration curve. The threat of incurring high cold and warm start-up costs, combined with the incremental costs of damage to the components of the plant, deterred Generators from carrying out such capacity manipulations within the mid-merit as well as in the peak segments of the LDC.

*Other capacity manipulations outside SMP determination*

The other capacity manipulation strategies that Generators might have used within the day were those for which the associated costs were charged to Uplift. There were a few temporary and long-term constraint boundaries that existed across England and Wales. Generators that had plants behind such boundaries maximised their private marginal benefits by offering their residual demand from such plants at abnormal prices. These were excluded from the SMP determination process, but they ran out of merit because NGC needed them to keep the transmission system safe and secure.

Half-hourly variations in price results from changes in Generators declared availability termed *availability profiling*. If the deviation between the declared availability and gross demand is above 1000MW, NGC decreased injections by constraining off some generation until demand peaked. The in-merit plants that were constrained 'off' earned their offer prices. A Generator with monopoly power within a geographic area could also offer its transmission service capacity above competitive rates.

### *3.6.4.5 Summary*

The definition of the operating plant availability was directed at enhancing viability of Generators' production costs and the price rule designed to signal appropriate investments into short and long-term infrastructure. Yet, Generators manipulated capacity and earned higher income.

### *3.6.5 Results*

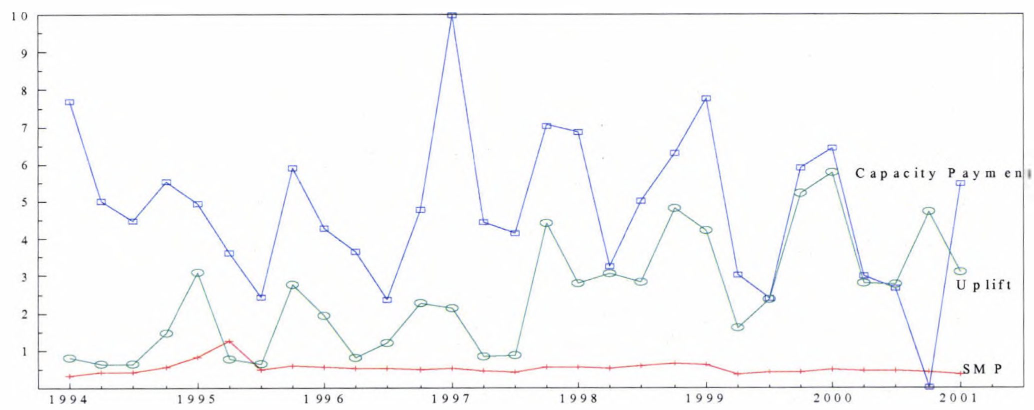
Table 3.6.1 is the tabulated descriptive statistics for the raw variables in the dataset. The highest component of PSP is SMP. There is significant variation in the PSP across all the sub-samples; the source of this variation is the combined volatility in the capacity payment and Uplift. As expected, PSP decreases after 1998; but this decrease is only a result of the large decrease in the SMP because the CP and Uplift increased over the same period. Since SMP is a proxy for the price at which Generators were willing to supply electricity, and therefore, a reflection of their bidding behaviour, it appears reasonable to conjecture that the decrease in SMP reflected an increase in competition. The divestment of NP and PGs mid-merit plants might have contributed to this. This conjecture is consistent with Evans and Green (2003) and Bower (2002).

The mean values and volatility for CP and Uplift increased dramatically after 1998. Since VOLL was constant and SMP decreased, the only plausible source for the higher values in CP was the corresponding increases in the LOLP. Given that the GRC increased a situation that reflects the increased, capacity on the network, the only reason for the increases in LOLP would have been that the Generators withheld capacity. The increase suggests that the NGC spent more money in resolving constraints, availability payment and start-up costs. These conjectures imply that after 1998, the Generators refrained from manipulating capacity in the SMP setting process; instead, increased capacity manipulation through the LOLP and Uplift. This is consistent with the earlier prediction by Bunn and Larsen (1992) who find that despite the projected excessive capacity that might come on-line in the UK, there was a non-linear relationship between LOLP and prices. This meant that the Generators would use capacity manipulations through the LOLP to earn higher prices.

As expected and in relation to the mean values of CP and Uplift, SMP exhibits lower volatility; it is also the component with the lowest relative variability as figure 3.6.5 shows. This result is consistent with Wolak and Patrick (2001); and since SMP is the most stable component, it implies that it may be easier to forecast it than Uplift and CP. The variability in CP and Uplift suggests that SMP may not always be the highest component of PSP during some half-hours.

**Figure 3.6.5**

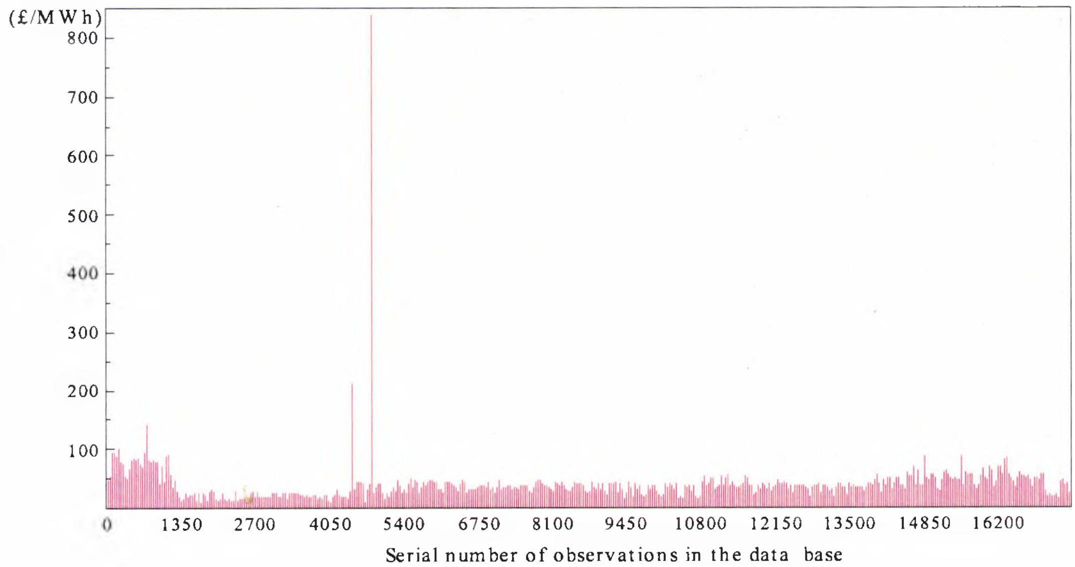
**Quarterly Coefficient of Variation: SMP, Uplift and Capacity Payment**




The Kurtosis for both the CP and Uplift confirms, as expected, that the two variables had more irregular values and possibly over longer runs than the SMP. The SMP Kurtosis in 1995 is abnormal and suggests that it might be due to abnormal observations, which figure 3.6.6 shows to be the combined effect of approximately £836/MWh, which occurred on 11 April, 1995 and £211/MWh that occurred on 4 April, 1995.

**Figure 3.6.6**

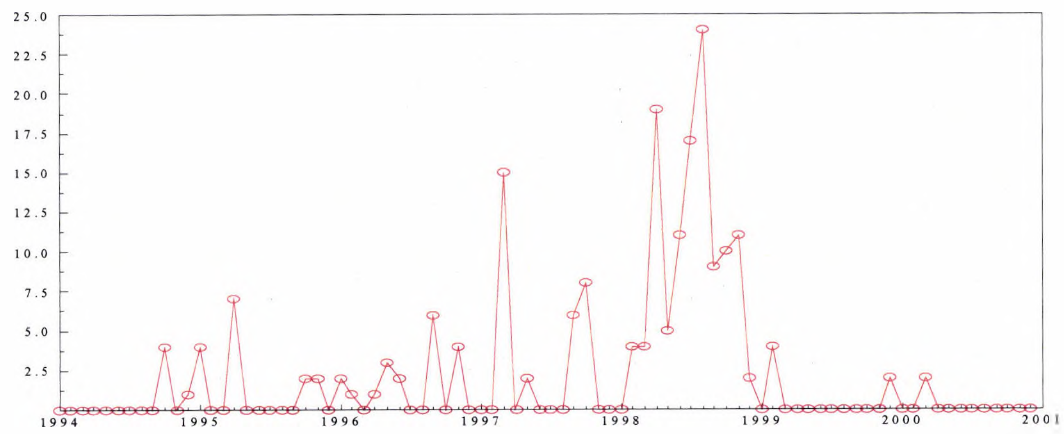
**Half-Hourly SMPs in 1995**



One of the expectations was that in relation to the other years in the full sample the low reserve margin in 1995 would cause higher prices during that year. Sub-section 8 uses three definitions of spikes to examine the factors that might have led to the spikes in SMP in the data. Using a definition of a spike as SMPs greater than three and half times the monthly average, figure 3.6.7 shows that in relation to the other years in the sample, there were lower numbers of spikes, which occurred in 1995.

**Figure 3.6.7**

**SMP Spikes: SMPs Three and Half Times the Monthly Average**

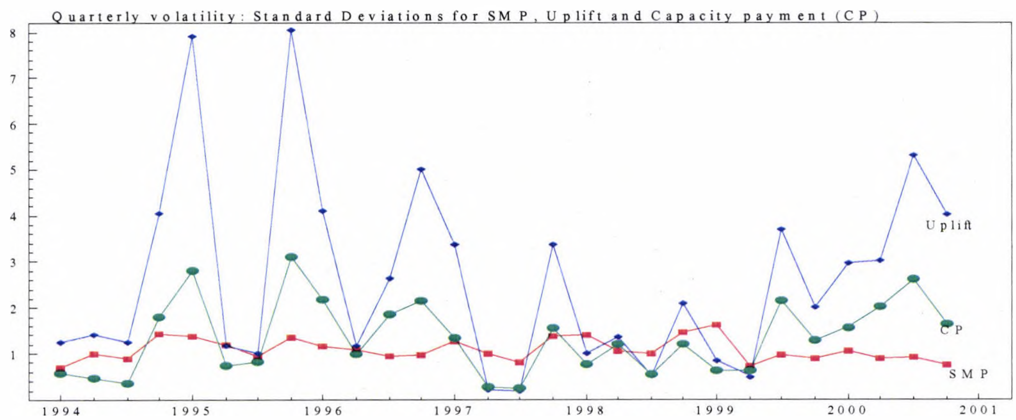



Statistical data exploration analyses assume normality of the variables. My preliminary skewness and kurtosis test: *sktest*, of these variables reported non-normality; therefore, I transformed all the variables except Uplift into square roots but worked with Uplift as an identity because it was nearly normal in its raw form.

I re-calculated the summary statistics across all the years, quarters and load regimes. Uplift and Capacity Payment still exhibit very high volatility; but as table 3.6.2 shows, Uplift becomes the more relatively volatile component of PSP; figure 3.6.8 also shows that Uplift has a higher standard deviations than CP after 1998. This result is only possible because Uplift is in its raw form, whilst capacity payment is transformed into square roots.

**Figure 3.6.8**

**Quarterly Standard Deviations for Square Root of: SMP, Uplift and Capacity Payment**




This does not change the overall implication of the earlier result. On the other hand, capacity is a factor that causes increases in availability payment, constraint costs and the costs for transmission services; that is, changes in the Uplift. It is also the causes of changes in CP; therefore, decreases (increases) in capacity will cause both Uplift and CP to increase (decrease). This positive correlation between the two variables means that they may cause problems when used together in models; but they can be used as proxies for each other.

SMP, Uplift, CP and PSP vary tremendously across the days in the week as shown in tables 3.6.3 to 3.6.6. On average, and as expected, prices are higher and with more volatility and increased incidences of the number of spikes, during the peak than off-peak periods. Consequently, prices are lower during the weekends, table B, baseload and the third quarter in the year; and it is higher during table A, weekdays, peak load regimes and the first and last quarters in each year. This confirms that prices exhibit weather seasonality. It is also volume dependent; therefore, prices are highest during the winter seasons, which is when more power is needed to operate heating facilities. Similarly, prices are higher during the peak-day period when there are more customers that use all types of electric gadgets, than at off-peak nighttime, when a significant number of consumers are asleep. This conjecture is consistent with Granger et al, (1979).

Tables 3.6.7 to 3.6.9 shows that price increases in magnitude from the baseload to the peak period along the LDC. As expected, the highest relative variability between the load regimes occurs during the mid-merit. This is consistent with the earlier findings by Fehr and Harbord

(1993). The high number of mid-merit spikes accounts for the wide range in the observed values of Uplift and Capacity Payment. For example, the lowest range in Uplift is £50.93/MWh, which occurs in 1994 and the highest of £95.60/MWh, which occurs in the year 2000. In contrast, the lowest range for capacity payment is £175.03/MWh, which occurs in 1998 whilst the highest of £1022.45/MWh occurs in 1995. This result is also a reflection of the conflicting objectives that the players have to face about maximising their marginal private benefit, yet make the right offer for capacity that guarantees inclusion in the unconstrained merit-order.

There is an aspect of the SMP statistics that is quite surprising; it has high Kurtosis during some weekends, which suggests that there might have been some spikes even during periods of low demand. Sub-section 9 shows there were other factors other than the market forces that determined the values of the SMPs. Based on this, the study conjectures that the weekend spikes might reflect the traders' opportunistic strategies; and is evidence that spikes might not reflect any intersystem transmission shocks, which may lead to *demand stress*. This is also consistent with Hogan (1998) who finds high prices occurring when there were no capacity constraints or demand stress in Pennsylvania, New Jersey and Maryland (PJM). On the other hand, it is also possible that the weekend spikes reflected genuine cases of temporary transmission constraints, which CC (2001) reports to occur across the network, were transient and disappeared as the system changed within the day. This is reasonable since most network maintenance programmes on the network utilities are usually scheduled for the off-peak weekend periods. If that was the case, then the higher prices, hence weekend spikes might be justified; and it will be consistent with Hale et al (2000). They find that higher prices occur in New England (NEPOOL) and New York (NYPP), when transmission constraints prevented power exports from regions of lower costs and excess demand in Michigan, Indiana, Kentucky, Ohio, West Virginia and Northern Virginia (ECAR) and most of Pennsylvania, New Jersey and Maryland (PJM).

There is a slight increase in the annual total system load (TSL) as shown in table 3.6.10. TSL is also less variable than prices; this implies that it may be easier to forecast TSL than any of the other variables in the pools price determination process. These are consistent with Wolak and Patrick (2001). Demand is highest during the first but lowest during the third quarter in the year; demand was lowest on Sundays and on average Thursday was the day in the week with the highest demand, as table 3.6.11 shows. Since the Generators matched demand, table 3.6.1 shows that the highest declared availability occurred during the period when demand was highest and which was over the weekdays.

The lack of responsiveness of aggregate load to prices is due to the volume of power that was routed through the pool. RECs and most of the other bulk electricity suppliers used *contracts for differences* (CfDs) to hedge against the volatile movements in the pool prices. CC (2001) reports that over 80% of the physical deliveries across England and Wales were locked into long-term CfDs; leaving only approximately 5% of TSL to be purchased at the PSP. Most of the suppliers had fixed prices and usually one-year contracts with their customers. The contract prices did not change with fluctuations in the pool prices.

### *3.6.6 Discussion*

The Generators did not rely on the SMP to manipulate prices after 1998. I conjecture that it might have been because the DGES had direct access to information about specific plant offers. Also he could use the powers vested on him through the *Electricity Act* 1989 (as amended), to refer any Generator that he perceived was inhibiting the development of efficient competition to the Competition Commission. It might also be a result of the Regulators persuasive communications with the Generators towards the run-up to the implementation of the NETA regime, to reduce market prices. This is consistent with Green (1999) that reports the Regulators use of strongly worded communications with the Generators as a way to ensure that prices were low in the pool. Once this assumption is made, it then makes sense that the Generators might have done all in their power to keep their offer prices low. It was very difficult for the Regulator to identify capacity manipulations that occurred outside the SMP process; therefore, it appears that after 1998, Generators began to use CP and Uplift more to earn higher rents. Moreover, even though the Regulator could identify the high offers, which the Generators made for their plants that were located behind constraint boundaries, he could attribute that to scarcity. The only issue might have been the magnitude of the increase to consider reasonable.

Table 3.6.13 shows enormous variation in the ratio of table A to B indicated half-hours between 1994 and 2000. The increased rates in these ratios over time reflect the increments in the costs for unscheduled availability (Littlechild, 1998). This ratio ought to have decreased as the market evolved if the Generators did not manipulate capacity. This is also a sensible assumption, because IPP entry led to excess capacity; and as the reformed industry matured, the Generators learned the market rules and acquired the skills with which they accurately forecast load. The SO also used more thermally efficient and flexible plants to meet demand. Also table 3.6.14 shows enormous variability in the proportion of Uplift to SMP; it increased across the sub-sample and as the market evolved. The cause of this increase is seen in the difference between the minimum and maximum values from 1994 through to 2000. The pattern of the Generators declared and redeclared availability also changed between 1994 and 2000. Table 3.8.10 shows on average, that declared was lower than redeclared between 1994 and 1995; but this changed between 1996 and 2000 when declared became significantly higher than the redeclared. This suggests that Generators earned income from inefficient and idle capacity.

**Table 3.6.13**
**Ratio of table A to B half-hours: *'Stretch'***

| *Year* | *N* | *Table Indicator* | *Frequency* | *%* | *Cum. %* | *Ratio of A:B* |
|---|---|---|---|---|---|---|
| 1994 | 17520 | A | 13,059 | 74.54 | 74.54 | 2.93 |
| | | B | 4,461 | 25.46 | 100.00 | |
| 1995 | 17520 | A | 13,019 | 74.31 | 74.31 | 2.89 |
| | | B | 4,501 | 25.69 | 100.00 | |
| 1996 | 17568 | A | 13,632 | 77.60 | 77.60 | 3.46 |
| | | B | 3,936 | 22.40 | 100.00 | |
| 1997 | 17520 | A | 15,026 | 85.76 | 85.76 | 6.02 |
| | | B | 2,494 | 14.24 | 100.00 | |
| 1998 | 17520 | A | 14,377 | 82.06 | 82.06 | 4.57 |
| | | B | 3.143 | 17.94 | 100.00 | |
| 1999 | 17520 | A | 14,294 | 81.59 | 81.59 | 4.43 |
| | | B | 3,226 | 18.41 | 100.00 | |
| 2000 | 17568 | A | 12,594 | 71.69 | 71.69 | 2.53 |
| | | B | 4,974 | 28.31 | 100.00 | |

### *3.6.7 Conclusion*

This section has shown the inefficiencies in the rules for setting prices in the pool; and provides an insight into the possible strategy that the Generators might have used to manipulate the PSP. These important revelations conclude that the SMP was the only component of the PSP that reduced after 1998, whilst the mean values and spikes in Uplift and CP increased significantly over the same period. CP remained the most volatile component of PSP throughout the regime; also, its higher values after 1998 were due to the increases in LOLP. The higher values of Uplift after 1998 suggest that the NGC might have incurred more costs to resolve constraints, procured transmission services within the day and made more availability payments. This confirms that Generators used capacity manipulation outside the SMP determination process to earn higher income in the pool.

# 3.7

# **On the Basic Structural Time-Series Modelling of Electricity Prices**: System Marginal Price (SMP) and Capacity Payment (CP)

*Abstract*

*This paper uses a structural modelling approach to investigate the weekly average SMP and CP between January 1994.week 1 and December 2000.week 52. It assesses the cumulative effects of competition and the knock-on effect of the changes in the natural gas trading arrangements, as well as the new electricity trading arrangements (NETA) on the trend of the two variables after 1998. It finds a structural break in SMP after 1998; thereafter, it decreases at an increasing rate. In contrast, increases in the loss of load probability (LOLP) caused CP to increase at an increasing rate. There was a significant entry of independent power producers (IPP) in the 1990s and which led to excess gas-fired capacity on a relatively unconstrained Grid; the incumbents also mothballed or withdrew inefficient fossil plants, meaning that NGC used thermally efficient plants to meet demand during the same period. The increases in LOLP suggest that the Generators were using capacity to manipulate CP; therefore, it concludes that the pool was an inefficient mechanism.*

*Keywords: Capacity Payment, Electricity Prices, England and Wales, Pool, System Marginal Price, Uplift*

## Introduction

The last section used exploratory data analysis to examine how the rules and processes for setting prices and capacity in the pool exacerbated gaming, leading Generators to earn excessive rents for each MW of power that they injected onto the Grid. The methodology reveals important issues for public policy. Firstly, it shows that relative to *Uplift* and *capacity payment* (CP), the *system marginal price* (SMP) is the most stable component of the pool-selling price (PSP). This may be because through the market surveillance team in his office, the *Director General of Electricity Supply* (DGES) had direct access to information about specific plant bids and offers. Consequently, it was possible for him to '*name and shame*' or issue strongly worded communications (see Green, 1999) to the Generators that persistently made abnormal offers into the pool. In extreme cases, and relying on the powers vested in him through the Electricity Act 1989, he could refer such Generators to the *Competition Commission* (CC), on the grounds of, for example, *pre-emptive*[18] behaviour.

The salient powers, which the Regulator used to maintain daily and close surveillance of market operations, combined with the series of inquiries into the causes of abnormal prices, restrained higher prices throughout the life of the regime. The experience is consistent with what happened in Hong Kong where Lam (1999) finds that regulatory oversight curtailed high spot prices. In an earlier investigation of the mark-up that National Power and PowerGen made from participating in the pool, Wolfram (1999) conjectures also that the DGES' close monitoring of the pool placed downward pressure on prices.

Secondly and consistent with Wolak and Patrick (2001), capacity payment (CP) is the most volatile component of the PSP. Wolak and Patrick's work covered the period between 1 April 1991 and 31 March 1995. The last section (3.6) confirms that CP continued to exhibit the highest relative volatility throughout the life of the regime; the evidence supports that capacity was the primary tool that the Generators relied upon to manipulate price setting. The industrial economics (IO) theory of strategic behaviour of firms in capacity constrained and oligopolistic industry, supports that the firms in such a market will use their supply functions

[18] 'pre-emptive behaviour' describes abuses by dominant incumbents in markets that are being opened to competition or are newly opened to competition, which are designed to adversely affect the development of competition' (OFT, March 2001:10).

to manipulate the market-derived prices (see Tirole, 1998). Capacity manipulation was a permanent feature in the pool; and capacity payment became a distinct and separate sub-market (Newberry, 1995). The question for public policy consideration is whether the excess reserve margin on the relatively unconstrained British Grid justified the levels of CP recorded in the pool.

Data exploration provides useful univariate information, but it does not provide insights into some of the unobserved properties in the variables. Consequently, this section uses a *basic structural modelling* (BSM) in the Kalman Filter[19], to estimate, quantify and obtain real values for the changes in the stochastic components of the system marginal price (SMP) and capacity payment (CP). A BSM consists of 3 components: a trend, a seasonal and an irregular (Koopman, 2000). Therefore, this paper investigates some of the important behavioural attributes that these two variables exhibited between January 1994 and December 2000, which are not easy to quantify in data exploration analysis. It is also not possible to gain an insight into these properties through *ordinary least squares* (OLS) regression analysis.

It answers the following questions:

- Is there a long-term average pattern in the variables?
- Did price trends change after 1998; if they did, what might be the plausible reasons for that?
- The original data bundles are organised in weekly basis just for this paper; therefore, it quantifies the deviation between the actual weekly average pattern and the long-term trend. Based on the literature on the developments in the industry during the period, it explores possible reasons why the deviations occurred.

Therefore, this part complements and reinforces the findings in 3.6. Parts 3.9 and 3.10 conduct multivariate analysis on this research to ascertain whether some of the variables in the pool prices setting move together; and provides the opportunity to investigate if a change in one variable affects another, and the direction as well as the magnitude of such changes.

[19] Control engineers developed the Kalman filter in the 1960s for 'applications concerning space craft navigation and rocket tracking' (Bentz, 2003:224; see also Kalman, 1960).

Finally, by identifying the dynamic and predictable path in the variables, the study will lay the foundation on which to forecast the series.

The rest of this paper is organised as follows: the literature review is in 3.7.2 and a discussion of the properties of the series, the organisation of the original dataset and the estimation software is in 3.7.3. The propositions of what the study expects to see in the estimation results are in 3.7.4; and the formulation of the model is presented in 3.7.5. The result of the generalised models is in 3.7.6 and 3.7.7 investigate the changes in the trend of CP and SMP after 1998. 3.7.8 discusses the implication of the investigations and 3.7.9 concludes this part of this empirical analysis section.

### *3.7.2 Literature Review*

The methodological basis for this paper is a common application in econometrics and statistics studies; therefore the relevant literature is different from the studies reviewed earlier in part 3.2; nonetheless, it complements the overall empirical investigation.

Modelling variance as an unobserved stochastic process in time series estimations is an age long methodological approach that economists use in quantitative macro economic studies, where the focus has been on the investigation of the non-stationary trend-cycle in the components of for example the *aggregate demand* (AD) curve. These economists have been particularly interested in predicting and providing appropriate polices that might enhance stable inflation rates, minimise the levels of unemployment and maximise the national output.

The underlying belief in some of the studies that the last paragraph refers to, is that it is possible to separate economic activities into those that are business cycle related from those that result from periodic seasonality. Businesses are characterised by events that are not regular, better still, one-off, which in statistics literature are referred to as random. When separating economic activities, these random related events can be identified as the difference between the sum of the trend plus seasons and the actual observations in the series. The random observations form what is collectively captured and known as 'others'; they can be thought of as the *'abnormal'* and / or *'irregular'* observations. These are symbolised as the *'error term'* in statistical models (Makridakis et. al, 1998).

The principle of decomposition analysis and smoothing of a series is based on obtaining the original series by addition or multiplication of the decomposed components of the variance of the series. The trends, seasonal and irregular, make up the original variable. And the methodological approach is to add together or multiply these components, to derive the original series. Transformed logarithms of series can be modelled as additive relationships (see Makridakis et. al, 1998:85). For example, if the assumption that the original SMP series can be derived by multiplying the trend, seasonal and irregular / error of its variance; the linear relationship can be defined in the multiplicative form as:

$$smp_t = \mu_t \times \gamma_t \times \varepsilon_t \qquad (1)$$

Where

SMP is system marginal price; $\mu_t$, the trend component; $\gamma_t$, the seasonal fluctuations are $\varepsilon_t$, the irregular or error term in the model. The multiplicative relationship becomes summative by taking the logarithmic transformation of SMP. That is,

$$\log smp_t = \log \mu_t + \log \gamma_t + \log \varepsilon_t \qquad (2)$$

Franzim & Harvey (1983) show that reliable and useful information about the stochastic properties of a variable can be obtained by defining a functional linear form that consists of the *trend, seasonal and irregular* components; having transformed a series into its logarithm, adding them to obtain the original.

There are two ways through which the best estimates of the unobserved components can be carried out: (1) by using seasonal *autoregressive integrated moving average* (ARIMA) models, thereafter decomposing and maximising the irregular component. (2) By estimating an explicit model, which consists of the trend, seasonal and irregular components in a time domain in the Kalman Filter or a frequency domain (Harvey & Scott, 1994: 1327). In the second option, the relevant model may be 'formulated in terms of the distribution of the one-step ahead predictor error' (Harvey et al 1994:247) and carried out by a *maximum likelihood* (ML) estimation.

With respect to the first option cited above, Harvey and Todd (1983) note that the three components decomposition approach is relatively better than Box and Jenkins (1976) ARIMA models. Bentz, (2003) compares three methods of estimating time varying factor sensitivities and concludes that the Kalman Filter based stochastic decomposition approach gives 'optimal estimates for non-stationary factor exposures' (page 213). Marvall (1985) '... compares the *structural model components* (SMC) with those obtained from the X-11 and the ARIMA-based signal extraction methods' (page 350), and finds that the SMC avoids data mining. It also provides useful insights into the behavioural pattern of the three components over the range of the sample as well as the mean square error of the estimated components. Harvey et al (1986) and Harvey & Scott (1994) uphold similar arguments.

The general assumption in stochastic variance models is that the long-term trend in the series, which includes a cycle, thus a trend-cycle, usually evolves slowly over time. The level average values are non-stationary, but are stationary in the first and second differences; therefore it is misleading to treat them as deterministic, particularly in simple OLS estimations. In contrast, the seasonal, which are the periodic fluctuations, can be constant over a length of time. Depending on the organisation of a dataset, this length of time can be months, weeks or quarters. In electricity, weather dependent elements, such as sunshine and rainfall levels, wind speed, humidity and cold spells, will be common feature within each season. Finally, the permanent trend and observed values will differ by the value of the error term, which will often be random, and hence unpredictable but they can be identified.

Variables with non-stationary trend-seasonal but relatively small error components are common real world models; and deterministic seasonals will be special cases (Franzini & Harvey, 1983). Harvey & Durbin (1986) find this special case of deterministic seasonal component in the data for seriously injured or killed in their three component decomposition modelling of the effect of the compulsory legislation to wear seat belts (implemented in February 1983 in Britain) on the number of traffic accidents.

A consistent argument in the literature about analysis of time varying sensitivity models is that regression analysis should include a linear relationship of the unobserved components of the series (see for example Harvey and Scott, 1994; Harvey et. al, 1986). By omission, where the stochastic properties are not included in a regression model the assumption is simply that the error term captures all the unexplained things that might affect the dependent variable but which are excluded in the model. An example is the effect of changes in the seasonal pattern of a series. It is important to know that the information from the unobserved properties might be very relevant in explaining the regression output; it might also form a reliable foundation for policy inferences (see Davidson et. al, 1978). By including a stochastic component in their regression analysis, Harvey et. al (1986) investigate the relationship between employment and output. They find that the long-term trend picks up technological advancement and on 'average, if output were constant, employment would fall, and productivity increase, by 2.6%p.a' (page 984). Similarly, Davidson et al (1978) investigates the relationship between aggregate income and consumption in the UK. They also discover that the seasonal component accounts for the behavioural hypothesis of consumption in each of the seasons.

Decomposing unobserved variance has also been applied in univariate models. Harvey et al (1986) cites that Harrison and Stevens (1976) as well as Harvey and Todd (1983) fit univariate time series models by treating trend components in this way. Harvey et al (1994) applies this technique to investigate the evolution of the seasonal patterns in consumption. It can also be used in bivariate and multivariate estimations. King et al (1978) used data on 'consumption, investment and output from the post war USA' (page 819) to investigate business cycles. Bell and Hillmer (1994) used the same approach to examine the benefits of seasonally adjusting the variables in econometric studies.

The studies mentioned in the paragraphs above, focus on the components of the *aggregate demand* (AD) curve; and include studies on employment, investment and consumption. Since the studies reviewed above relate to macro statistics, one may be tempted to think that applied stochastic decompositions are the exclusive reserve of macroeconomics, thus cannot be applied easily in microeconomic analysis. The England and Wales' pool was a national market in which the entire country was a single price zone with a uniform price paid to Generators and received from the demand-side. Consequently and in the sense that macroeconomics statistics are national, its prices are proxies' macro data.

None of the earlier studies reviewed in section 3.2 investigates the patterns in the unobserved properties in the pool data. Therefore, this current research is the first to use the stochastic decomposition of the variance of series in a BSM, to investigate the behavioural properties of the weekly average of the half-hourly wholesale prices in electricity markets. This is the case of the England and Wales' electricity pool, and in which the study examines the components in the pool selling price (PSP) in £/MWh, their growth rates and how they changed over time.

Electricity prices are volatile, mean reverting and time and weather dependent. The seasonal dependency of electricity variables is one of the primary factors that account for the serial correlations in the error term in models. The evolutionary emergence of electricity markets suggests that the underlying average price will be non-stationary; however they may be stationary in the first and second differences. The dataset for the study confirms this in figures 3.7.1 and 3.7.2, which provides a visual illustration of the logarithms and correlograms of SMP and CP. The logarithm transformations for both variables are stationary in the first differences of the series. The corresponding correlograms show significantly reduced correlations. When this is combined with the inherent features of the variables, which figures 3.7.3 and 3.7. 4 show, it seems that electricity prices will behave properly if they are modelled stochastically. Moreover, from the literature review on the macroeconomic studies that have used the approach, it seems that modelling the unobserved variance of SMP and CP based on the most recent observations in the series and which is carried out in time using the Kalman Filter, will provide unbiased estimates. These can be a basis for reliable policy prescriptions for the design and price rules in electricity markets.

**Figure 3.7.1**

**Graph of Correlogram of Log of SMP and First Differences**

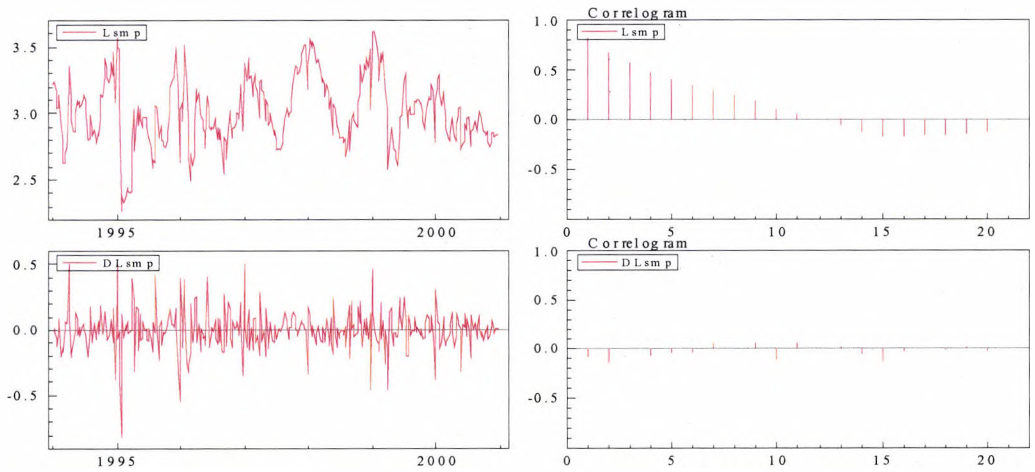
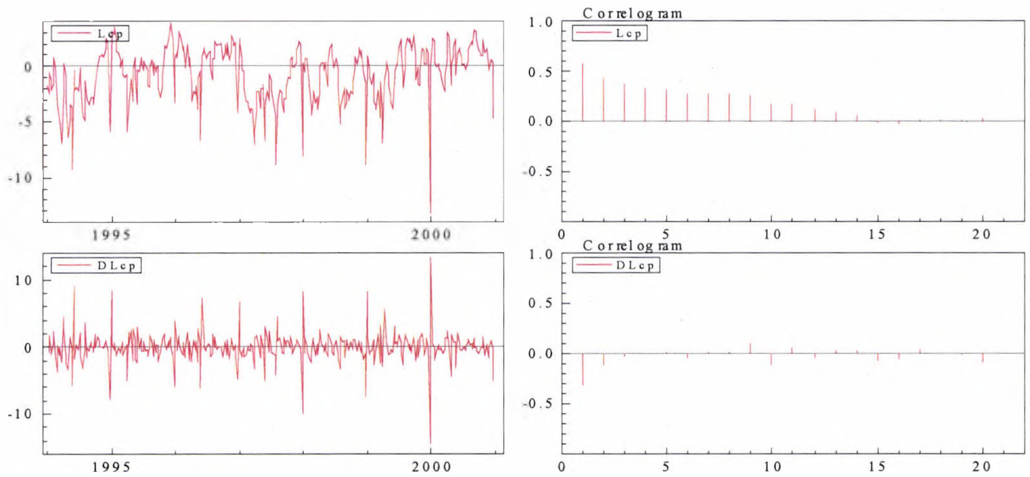


**Figure 3.7.2**

Graph of Correlogram of the Log of Capacity Payment and First Differences




Franzim and Harvey's (1983) additive relationship formulation is the methodological basis for this investigation. Therefore the underlying assumption is that the logarithmic transformation of the original SMP and CP series is obtained by summing up the trend, seasonal and irregular components of the variance. Based on the developments in the literature regarding the evolution of the regulatory reform in the UK electricity industry since vesting in 1990, this study expects that the trend of the pool variables will change after 1998. Therefore, it estimates three separate models: (1) the range in the full sample, from January 1994.week1 to December 2000.week52, (2) and sub-samples: January 1994.week1 to December 1997.week52 and (3) January 1998.week1 to December 2000.week52.

*Summary*

The three components stochastic decomposition analysis is a common technique that economists use in applied macroeconomics analysis. Pool prices, which were national level data, exhibit features that are similar to the macroeconomic time series. Consequently, this research expects that the variances of SMP and CP will behave well in univariate BSM.

### 3.7.3 *The data*

This study uses the weekly average of the half-hourly SMP and CP between January 1994.week1 and December 2000.week52.

*System Marginal Price (SMP)* is the market clearing commodity price. It is the price at the intersection of aggregate *gross demand* (GD) that NGC forecasted and an industry step supply function (see Green and Newberry, 1992) derived by stacking the Generators declared availability (DA) in an ascending order of marginal costs. GD is the sum of the consumption expected by pumped storage, large consumers that took off over 250MW of power from the *national transmission system* (NTS), none daily metered (NDM) sites and reserve requirement.

*Capacity Payment (CP)* is the payment made to Generators for making their plants available, and calculated as $LOLP * (VOLL - SMP)$. Where LOLP, the *loss of load probability*, ranges between 0 and 1; it is the probability that the system may not have enough capacity to meet increments in demand on the day. Whilst the *Value of Loss Load* (VOLL), is imputed from utility planning models (Kwoka, 1997); it indicates the social cost of electricity, which is the price that the consumers are willing to pay to have power rather than having power outages.

Economic theory of scarcity and price underlies the design of the processes and rules for setting prices in the pool (see Sloman, 2002; Scherer & Ross, 1990). Therefore, CP and SMP reflect scarcity and they increase (decrease) with decreases (increases) in capacity. Tables 3.5.1A and 3.5.1B in part 3.5 present the descriptive statistics for the two variables. The variables present very different statistics; but there are some generic features. The mean of the sub-samples whether annual, quarterly, weekly and load regime, are not constant. The standard deviations are very high both in the original as well as in the transformed square root (*sqrt*) of each series; this is a confirmation that they are quite volatile. The Kurtosis for each variable exceeds 3, with that of CP ranging in hundreds; and confirming the presence of very high spikes. These spikes usually increase the values of the mean. The Skewness and Kurtosis tests (*sktest*[20]) carried out in STATA version 8, rejects the null hypothesis of normality.

[20] This is a combined skewness and kurtosis test of normality carried out in STATA.R8 for each variable. STATA is the primary software used to collate, organise, and calculate some of the variables from the initial dataset. It was also used for the initial examination of the data as well as the data exploration. The original dataset came in 87 monthly 'bundles' for 7.25years.

Figures 3.7.3 and 3.7.4 shows the original seasonally unadjusted series and the logarithm transformations of the weekly averages of half-hourly observations of SMP and CP in £/MWh between January 1994.week1 and December 2000.week52. Both the original and the logarithm of the series exhibit some salient features that are common to economic time series.

**Figure 3.7.3**

**Analysis of weekly average system marginal price (SMP): original series and logarithms**

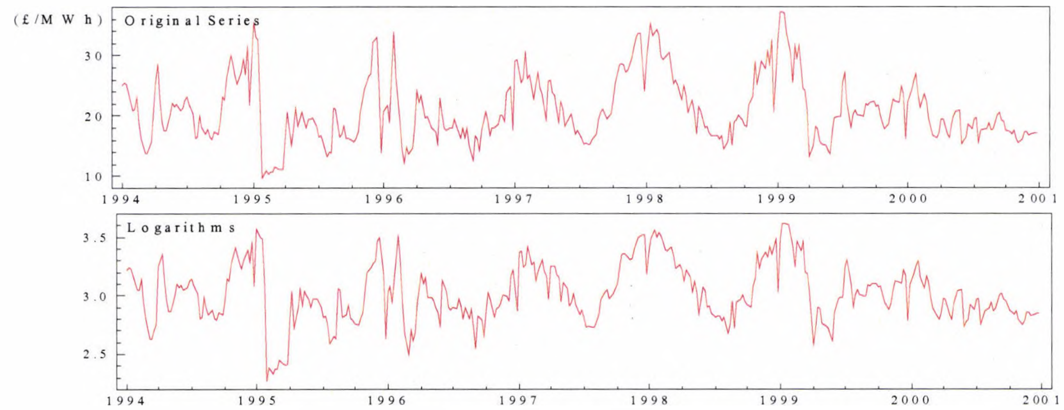
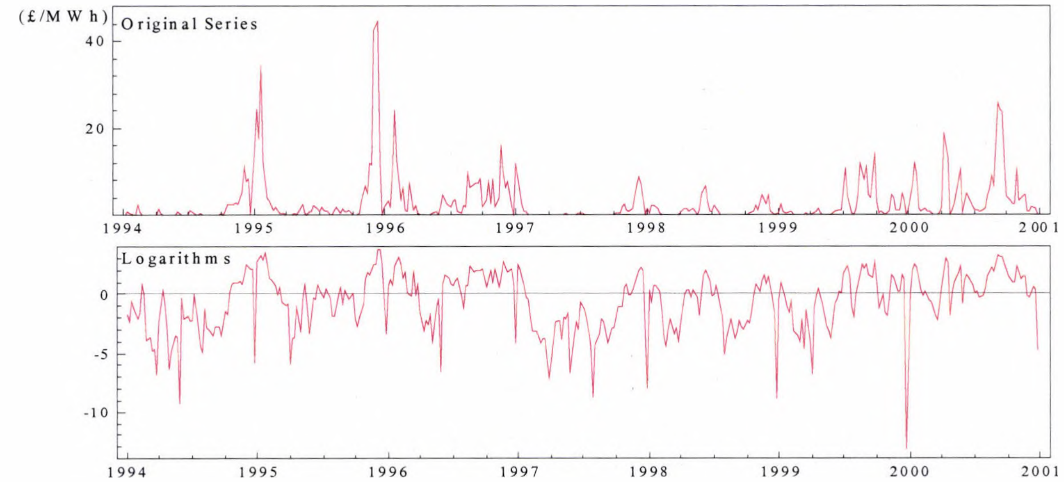



**Figure 3.7.4**

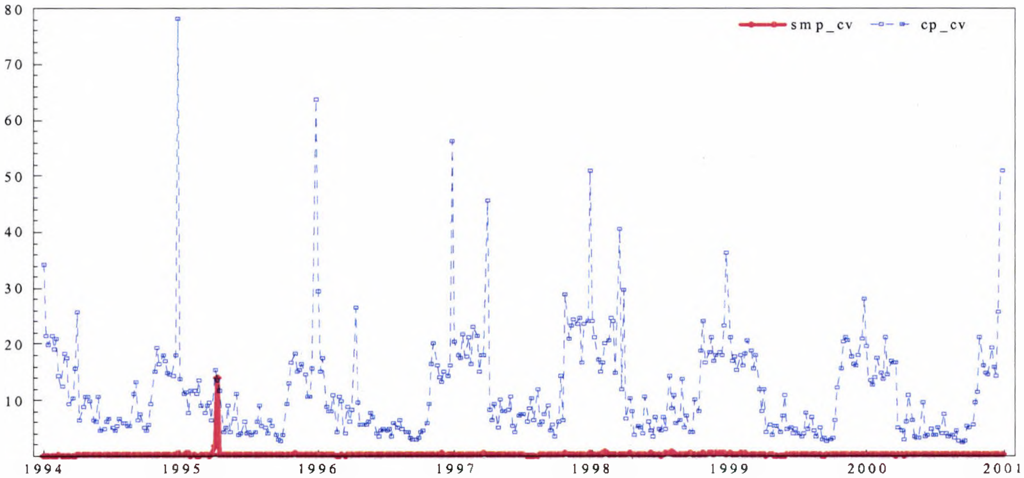
**Analysis of weekly average capacity payment (CP): original series and logarithms**




There is a clear underlying value for each variable, confirming that they are trended. This underlying value shows a long run / permanent pattern in its movement throughout the sample; thus, suggesting that it grows or reduces over time. SMP exhibits a downward sloping pattern, whilst CP increases over the long-term. This suggests that there might be some long-run factors that will affect the growth rate of the level average values. Therefore, it is reasonable to assume that the underlying level CP and SMP are non-stationary. The graphs also show some spikes that suggest that there will be a deviation between a stochastic trend and the observed weekly averages. CP is considerably more volatile than SMP, and the relative volatility between the two variables is clearer in figure 3.7.5, which is a graph of the coefficient of relative variability of the original weekly half-hourly averages.

**Figure 3.7.5**

**Weekly Average Coefficient of Variation for Original Series of SMP and CP**




Finally, the graphs of the logs and original series show that they exhibit seasonal fluctuations; and when compared to figure 3.7.5, the seasonality in CP is much more pronounced. The weather and time dependency of electricity demand is partly the explanation for the seasonal pattern in the series. The seasonality is more stable in the transformed logarithm than in the original series, presumably because the log of the series provides a 'shrinkage' effect; it also creates a near normality for highly skewed data. This research uses the transformed logarithm of the CP and SMP.

*3.7.3.2 Software / Organisation of original data*

The models are run on STAMP, the *Structural Time Series Analyser Modeller and Predictor.* STAMP is specifically designed for conducting structural time series analysis. The idea is to assume independence between each time structure, for example by weeks for this study. Structural time series analysis provides the opportunity to assume that the variables within each period are clusters; therefore, we can consider the independence of observations between periods. Since the variables that go into prices and equilibrium setting depend on each other, it is possible to consider independence between periods (see STATA.8—R: 331 – 341). Heteroskedastic time series such as the dataset for this study are vulnerable to mis-specification; therefore, conducting structural estimations helps to curtail bias in the estimates (see Koopman et al, 2000).

STAMP does not have half-hourly calibrated input options; instead, it has annual, quarterly, weekly and 'other'. It assigns identification numbers that start from 'one', to a database that is built on the 'other' option. It is difficult to carry out seasonal effect investigations on any variable within the 'other' category. If the data for this study is entered as annual, it will have 7 points, 28 if quarterly and 364 in the weekly specification. But the estimation technique, which is the *maximum likelihood* (ML), maximises the probability of obtaining the *best and unbiased* estimates in a larger, than in smaller samples. Consequently, the weekly option, which gives the highest number of observations, is preferred since it will enhance the validity of the estimates. It also gives sufficient numbers of annual observations with which to carry out a reliable investigation of the seasonal effects.

#### *3.7.3.3 Summary*

The weekly average half-hourly data series of SMP and CP exhibits features that support the use of variance to model their unobserved properties. This research is based on the additive linear relationship principle; therefore, the original data will be transformed into the logarithms of the series. I use STAMP to run the estimations.

### 3.7.4 Propositions

This sub-part formulates some of the propositions that the estimation results might reveal. It draws insights from the exploratory data summary presented in tables 3.5.1A and 3.5.1B in section 3.6. It is also based on the review of literature on the earlier studies on the pool data (see section 3.2); the information gathered from the *Office of Gas and Electricity Markets*

(OFGEM) and the *Department of Trade and Industry* (DTI) communications, which the consultation documents, reports and press releases, to mention but a few, contain.

The time-weighted and demand-weighted prices tabulated in tables 2.5.1 and 2.5.2 suggests that the pool variables evolved slowly after the reformed industry was vested in 1990. There is usually a hard landing period after an industry reform occurs; it is during that period that firms' learn the rules, complete compliance to changeover to systems, retrain staff and develop their commercial strategies. Firms also use the period to acquire the skills with which to forecast demand profiles. Errors of commission and omission, which are often due to the lack of knowledge of market rules and system operations, are features of the steep-learning curve. It is also common to find that participants may over-contract during these early days, presumably to curtail volumetric risks and high system top-up charges. These imply that the level average SMP and CP will evolve slowly as the industry developes. Even when their pattern emerges, weather changes plus environmental conditions combined with other forms of inter-system transmission shocks, can cause the observed values to fluctuate around an underlying level average.

Part of the regulatory reform in the UK was the *initial portfolio* (IP) contracts, a transitory arrangement, which the Government put in place between the Generators and the *Regional Electricity Companies* (RECs). This *IPs* was based on take-or-pay agreements; and they were backed against British Coal. These transitory arrangements affected the emergence of efficient competition in the industry (Powell, 1992; Green, 1994); and the contract and the average pool prices were both much lower than expected during the early years of the regime (OFFER, 1991; OFFER, 1994, CC, 2001). The knock-on effect was also that the trend of SMP and CP emerged rather slowly too.

*Proposition 1*

*There will be a linear trend for CP and SMP. This value will consist of an underlying average that will have a growth rate.*

*Proposition 2*

*Inter-system transmission shocks, will cause the actual half-hourly observations, hence the weekly averages in this study, to oscillate around this underlying average level.*

The *Director General of Gas and Electricity* (DGGE) expressed the view at the first anniversary of the New Electricity Trading Arrangements (NETA) that pool prices started to reduce as soon as the model for trading outside the pool was announced in 1998. By the end of the first year of the NETA, wholesale prices had decreased by over 40% (see OFGEM 2002a, b & c). Evans and Green (2003) as well as Bower (2002) uphold that cumulative factors and not just the DGES' announcement of the imminent plan to trade outside the pool, contributed to the dramatic decreases in the pool prices after 1998. There is this generalised notation of the significant decreases in pool prices after 1998. It appears that the focus is on the PSP and PPP, both of which reduced. But as this research has already shown in part 3.6, their reduction was only due to the decreases in SMP; Uplift and CP increased at an increasing rate after 1998.

*Proposition 3*

*This investigation assumes that the Governments policy to change the electricity trading arrangements affected price trends after 1998. Depending on the severity of the responsiveness of prices to policy changes, it could cause a break in the structural path of the respective level average SMP and CP.*

The National Grid Company's (NGCs) forecast of gross demand in England and Wales was based on the historical *seasonal normal demand* (SND) profiles for consumption over fifty years. In the mid-1990s, it seems that the effect of the climate change and global warming led to the emergence of warmer winter and hotter summer months than the SND records. This meant that consumers needed more power and gas to run heating facilities during the winter months or for cooling during the summer. Also more gas was exported through the interconnector to the Continent. The increased demand by Shippers for capacity in relation to the maximum physical that Transco provided resulted in the shortages and higher prices at the St. Fergus gas terminal in the summer of 1998 (see OFGAS, 1999). In electricity, the dataset shows a slight change in the pattern of the gross demand. This means that after 1998 demand became more unpredictable and it might have been difficult to rely solely on the SND profiles for an accurate load estimate.

*Proposition 4*

*After 1998, the stochastic seasonal component in this study may be non-stationary.*

*3.7.4.2 Summary*

This study expects that on average, these four propositions may be sufficient to explain the patterns, which the estimation results will reveal. The formulation of the model follows.

### *3.7.5 Formulation of the Model*

This sub-section explains why the variables will have a trend, a seasonal and an irregular, before presenting the model.

*3.7.5.1 Trend*

The trended-cycle feature of SMP and CP were already established in figures 3.7.3 and 3.7.4 (see also tables 2.5.1A and 2.5.1B). One of the propositions for this investigation is that the underlying average SMP and CP will emerge as the reformed industry matured and traders learned the rules, understood demand profiles and developed their commercial strategies. This study symbolises this trend during each week as $\mu_t$; it will have a path, which it will follow in the full sample. Assuming a point in time in the dataset, the weekly level average of CP and / or SMP, will grow at a rate, which is equal to the slope of the trend. This study identifies this slope as $\beta_t$; and one thing to note about this growth rate, is that it may or may not be constant along its path.

Electricity prices are affected by neighbouring markets (Bunn, 2004). In a system that has active gas and electricity markets, the policies in natural gas will affect electricity prices. For example, the owners of dual fuel Generators will depend on policies in gas to take-off gas and guidelines from electricity for Grid injections.

By the mid-1990s, the natural gas and electricity markets had converged (see Larsen and Gary, 1998). A lot of the energy industry participants held both gas shipping[21] and electricity generation licences; this was in a system in which consumers used electric or gas facilities for heating or cooling purposes at different times in the year. Transco (2002) reports that by 2001, the number of gas-fired generation plants that were direct connects to the natural gas transmission-line had grown from 1 in 1990 to 32. These gas-fired plants consumed a significant proportion of the throughput from the offshore fields and storage sites. This suggests that policy initiatives in natural gas affects capacity and commodity regimes in electricity; however, the differences in balancing periods between the two markets makes the influence in their interaction felt more in electricity than in gas.

The UK belongs to a common energy market that the European Commission (EC) directs. Member states tend to harmonise energy initiatives in line with the EC energy policy directives. Thus, the EC indirectly influences energy policies in the integrated regional market. There is a natural gas interconnector that runs from the Bacton terminal to Zeebrugge in Belgium, which facilitated inter- continental gas trading. In an unpublished research, Zhong (2002) finds that the price differentials between the UK and the Continent created profitable arbitrage; and forward trading occurred through the interconnector. The volume of trades with the Continent meant that the activities in the neighbouring markets affected the on-shore capacity and commodity regimes in the UK.

The slope component of the trend in the models will capture the effects of the changes in public policy, which would include all aspects of regulatory interventions such as price controls and announcements. It will also reflect policy changes that relate to changes in all aspects of fair-trading acts and agreements between undertakings. The severity of the effect of a change in policy could manifest as a permanent upward or downward movement in the level average values of either the CP or the SMP. A permanent change of this nature in the path of the level average of a series is called a *structural break* (see Enders, 2004; Koopman, et. al, 2000).

[21] '[A] Shipper – a company with a Shipper Licence buys gas from producers, sells it to the suppliers and employs the [gas transporter a] GT to transport the gas to consumers. It may also store gas with a Storage Operator to help it manage the balance between its suppliers and the consumer's demand. Its licence requires it to be reasonable and prudent in the way it uses the GTs pipeline Network' (Transco, undated: 6).

#### *3.7.5.2 Seasonal*

Pool prices exhibit weather seasonality (see the data exploration in section 3.6). Figures 3.7.1 and 3.7.2 show the permanent periodic fluctuation in the SMP and CP. The peaks in the series signify the periods of the highest observed values and these are associated with the weeks in which the highest demand occurred. These are usually during the colder winter weeks. The troughs relate to the weeks in which prices were lowest; these were the periods of the lowest demand and on average are likely to be during the spring.

Seasons are categorical variables; and dummy variables are mainly used to capture or to isolate seasonal effects in statistics estimations. Where data is organised in quarters and is represented over a full one-year cycle that runs from January to December, the four distinct seasons will be the winter, autumn, spring and summer. It is also common to assign constant numbers to different seasons; that is a specific number will be used to identify a particular season in the year. For example, winter can be represented as 'one'; autumn, 'two'; spring, 'three' and summer, 'four'. Models with such seasonal-identities provide only a basic intuition about the effect of weather on the variables in the estimation.

Where variables exhibit weather dependency, which is the case in the dataset for this study, using the basic dummy categorisation may lead to model mis-specifications; the knock-on effect of which may be biased inferences and wrong policy prescriptions. Harvey and Scott (1994) recommend the use of trigonometric seasonal formulations in stochastic modelling. The 'trigo' formulation has an inherent feature that allows for an easier adjustment of observations to the changing seasonal patterns. It also decomposes the irregular or random elements in the series whilst bringing out a clearer and smoother pattern in the trend component.

#### *3.7.5.3 Irregular*

The irregular component in the models will capture all the abnormal values in the series. It is the same thing as the residual or the 'remainder' in the model because it is the difference between the stochastic trend and the actual weekly observations. In the electricity market literature, these irregular observations are called 'spikes'.

In some data, irregular entries may result from data input errors. The dataset for this study is characterised by high incidences of irregular and non-erroneous observations; they do not have a pattern. For example, a very high observation can be followed immediately by a very low one, which might be over a long or short run.

Figures 3.7.3 and 3.7.4 shows spikes inherent in the SMP and CP series. The point to note here is that price spike is a permanent and generic feature of electricity markets. Inter-system transmission shocks that make it impossible to transport power easily from places of lower cost and excess supply to those of higher demand and higher cost, exercebates price spikes. Hale et al (2000) finds higher prices occurring in New England (NEPOOL) and New York (NYPP) when transmission constraints prevented power exports from regions of lower costs and excess demand in Michigan, Indiana, Kentucky, Ohio, West Virginia and Northern Virginia (ECAR) and most of Pennsylvania, New Jersey and Maryland (PJM).

It is also possible for spikes to occur when the system is not under any 'demand stress'. That is, their occurrence might not have any relationship to the demand and supply situation on the network; instead, the traders' opportunistic strategies will be the primary driver. In this case, the oligopoly combined with the short-run capacity constraint nature of the industry worsens the Generators exercise of monopoly power. Hogan (1998) shows that without transmission and capacity constraints prices were very volatile and spiky in Pennsylvania, New Jersey and Maryland (PJM). Similarly sub-section 3.8 of this research finds that the spikes in SMP between January 1994 and December 2000 did not reflect the underlying demand, supply and system security situations.

This study symbolised this irregular component as $\varepsilon_t$; it will capture the effect of the spikes on the level average values in the models.

*3.7.5.4 The model*

Franzini & Harvey (1983), Harvey & Scott (1994) and Harvey et al (1986) provide the theoretic statistics basis for the formulation of the linear relationship; what this study does, is to apply the estimation technique.

Franzini and Harvey (1983) define the global model as:'

$$y_t = \alpha + \beta t + \sum_j \delta_j z_{jt} + \varepsilon_t \qquad (t = 1, \ldots, T) \qquad (3)$$

Where $y_1, \ldots y_T$ are the observations, $\alpha$ and $\beta$ are the trend parameters, $\varepsilon_t$ is a normally distributed white noise disturbance term with mean zero and variance $\sigma^2$, the $z_{jt}$ 's are seasonal dummies and the $\delta_j$ 's are their coefficients. If there are $s$ seasons in the year there will normally be $s-1$ seasonal dummy variables. This makes a total of $s+1$ regression parameters and these parameters can be estimated efficiently by ordinary least squares. The corresponding stochastic model is

$$y_t = \mu_t + \gamma_t + \varepsilon_t \qquad (t = 1, \ldots, T) \qquad (4)$$

Where $\mu_t$ and $\gamma_t$ are the trend and seasonal components respectively. The trend is defined as

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \qquad \beta_t = \beta_{t-1} + \zeta_t, \qquad (t = 1, \ldots, T), \qquad (5)$$

Where $\eta_t$ and $\zeta_t$ are normally and independently distributed white noise processes with zero means and variances $\sigma_\eta^2$ and $\sigma_\zeta^2$ respectively. The seasonal component is

$$\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t \qquad (t = 1, \ldots, T) \qquad (6)$$

Where $\omega_t$ is the normally distributed white noise with variance $\sigma_\omega^2$. The irregular component $\varepsilon_t$ is normally distributed white noise with variance $\sigma^2$ (page 674). The four variances: $\sigma^2, \sigma_\eta^2, \sigma_\zeta^2, \sigma_\omega^2$, are mutually uncorrelated, independent and are the only unknown parameters in the model (Maravall, 1985).

Equation 6 is a basic seasonal formulation; but a trigonometric seasonal is the most appropriate for this analysis. Assuming even seasons, Harvey and Scott (1994) define the trigonometric seasonal as:

$$\gamma_t = \sum_{j=1}^{s/2} \gamma_{jt} \tag{7}$$

Where

$$\left.\begin{aligned} \gamma_{j,t} &= \gamma_{j,t-1}\cos\lambda_j + \gamma^*_{j,t-1}\sin\lambda_j + \omega_{jt} \\ \gamma^*_{j,t} &= \gamma_{j,t-1}\sin\lambda_j + \gamma^*_{j,t-1}\cos\lambda_j + \omega^*_{j,t}, \qquad j = 1,\ldots,s/2-1, \quad \lambda_j = 2\pi j/. \end{aligned}\right\} \tag{8}$$

And

$$\gamma_{j,t} = -\gamma_{j,t-1} + \omega_{j,t}, \qquad j = s/2, \tag{9}$$

Where $\omega_{jt}$ and $\omega^*_{jt}$ are normally distributed, zero mean, white noise processes. In estimating [our (5)] we assume these disturbances to have equal variance, $\sigma^2_\omega$, although in principle we could have a different variance for each frequency; see Maravall (1989) for the link with seasonality in ARIMA models' (page 1328) (see also Koopman et. al, 2000).

*3.7.5.5 Summary*

Equation 4 is the is the stochastic model for the logarithm of SMP and CP; and the respective formulations are:

$$lsmp_t = \mu_t + \gamma_t + \varepsilon_t \qquad (t = 1,\ldots,T) \tag{10}$$

$$lcp_t = \mu_t + \gamma_t + \varepsilon_t \qquad (t = 1,\ldots,T) \tag{11}$$

The trend has two components, the level average value: $\mu_t$ and a slope: $\beta_t$; equation 5 provides their extended definitions. The estimation uses a 'trigo' seasonal formulation that equation 7 represents; therefore, the component graphics will show an enhanced optimal smoothing of the dynamic seasonal effect.

The estimation is carried out in STAMP, a menu driven software. It uses ML technique to estimate the variances of the trend: $\sigma_\eta^2$, and the irregular: $\sigma_\zeta^2$; thereafter, uses a Kalman Filter to estimate the underlying average of SMP and CP: $\mu_t$ (see Koopman et. al, 2000).

### *3.7.6 Results*

This sub-section presents the results and discusses the estimation outputs.

The estimation reports a successful and satisfactory ML optimisation, which suggests that the models are reasonable (see Koopman, et al 2000:33). The estimates for the log of CP in all the sub-samples and that for the log of SMP for the sub-sample January 1998.week1 to December 2000.week52, have **very strong** convergence's. There is **strong** convergence for the estimated log of SMP for the period between January 1994.week1 to December 2000.week52. But there was no estimation done for the log of SMP for the sub-sample of January 1994.week1 to December 1997.week52.

#### *3.7.6.1 Summary statistics*

Table 3.7.1 is the tabulated summary statistics. The normality values are quite high, suggesting the dominating presence of spikes in the level average. This is inevitable because the original series had these spikes; and they did not disappear following the logarithmic transformations. Therefore, the normality values simply reflect the more predictable and limited space, which the observed values, occupy; that is, the series are not Gaussian.

The Box-Ljung statistics confirm that the residuals are serially correlated. These serial correlations are attributable to the changing seasonals inherent in the series.

The higher 1.41 in the estimated log of CP between January 1998.week1 and December 2000.week52 indicates increases in the variance during that sample period. This result is also not surprising given that in relation to the SMP, CP is more volatile; figure 3.7.2 also shows an apparent constant increasing trend in the CP after 1998, which means that the variance over that same period will follow an increasing pattern.

It is not reasonable to reject the models based on the shortcomings revealed in the summary statistics and which have been discussed in the last few paragraphs. One thing that is worth mentioning is that Harvey et al (1986) expresses the view that the Box-Ljung statistics is not an ideal test for residual serial correlation (page 981). In some other dataset, it may be possible that a change from a univariate to a bivariate or multivariate estimation and which might give significantly different results, may be a way to deal with the residual serial correlation. In this case, given the seasonal dependency of electricity prices, it does not seem that residual serial correlation and heteroskedasticity will disappear in bivariate or multivariate estimations based on the variables from the pool data 'bundles'.

On the other hand, the functional dependence of the price setting variables may cause a different type of specification concern. Variables that go into the determination of equilibrium prices and capacity in most markets (including the pool) usually depend on each other. The main pool prices, the pool-selling price (PSP), capacity payment (CP) and the pool purchases price (PPP), all hinged on the day-ahead determined SMP. In addition, changes in capacity cause both SMP and CP to change over time. SMP and CP are positively correlated. The strength of the correlation between the two variables exceeds .90 most of the time; and a regression analysis that includes the two variables may simply be estimating their high correlations. The goodness of fit in such estimations may be quite high; and the inferences that are based on such results may be unreliable because the estimates will be biased. The rule of thumb in applied statistics estimations is that there will be no difficulties in estimations where the strength of correlation between the independent variables lie between $-0.70$ and $+0.70$ (see Lind et al, 2001; Ott, et al, 1992; Neter et al, 1996). In theory, CP and SMP cannot be modelled together because of the high correlations between them. This is actually the reason why this study embarked on the univariate estimation analysis as a way to gain a useful insight into the stochastic properties of the two variables.

Finally, the violations which the summary statistics highlight confirm the limitations of the standard statistical techniques, which usually work well in other datasets that often come from normally distributed populations or are nearly normal when transformed, to model electricity prices. One lesson though, is the need to carry out a further investigation into the cause(s) of the spikes in CP and SMP. Section 3.8 investigates this for SMP.

**Table 3.7.1**

**Summary statistics**

Lsmp = Trend + Trigo seasonal + Irregular

$$lsmp_t = \mu_t + \gamma_t + \varepsilon_t$$

| Period | *Nos.* | *Std. Error* | *Normality* | *H(j)* | *r(j)* | *r(j)* | *DW* | *Q* | $Rs\hat{}2$ |
|---|---|---|---|---|---|---|---|---|---|
| 1994 - 2000 | 364 | 0.127 | 78.23 | 0.464 (103) | 0.0065 (1) | -0.0117 (18) | 1.96 | 31.54 (18,15) | -0.073 |
| 1994 - 1997 | 208 | 0.134 | 33.82 | 0.349 (51) | 0.025 (1) | -0.087 (13) | 1.95 | 17.92 (13,10) | -0.213 |
| 1998 - 2000 | 156 | 0.105 | 6.88 | 0.712 (34) | 0.035 (1) | 0.080 (11) | 1.84 | 12.14 (11,8) | -0.265 |

Lcp = Trend + Trigo seasonal + Irregular

$$lcp_t = \mu_t + \gamma_t + \varepsilon_t$$

| Period | *Nos.* | *Std. Error* | *Normality* | *H(j)* | *r(j)* | *r(j)* | *DW* | *Q* | $Rs\hat{}2$ |
|---|---|---|---|---|---|---|---|---|---|
| 1994 – 2000 | 364 | 1.59 | 54.99 | 0.819 (103) | 0.077 (1) | 0.037 (18) | 1.80 | 32.37 (18,15) | -0.356 |
| 1994 – 1997 | 208 | 1.55 | 21.26 | 0.435 (51) | 0.057 (1) | 0.096 (13) | 1.85 | 26.79 (13,10) | -0.548 |
| 1998 – 2000 | 156 | 1.302 | 12.33 | 1.136 (34) | 0.087 (1) | 0.077 (11) | 1.70 | 15.44 (11,8) | -0.229 |

Note: Nos. is the number of observations within the sample period. The standard error is the square root of the prediction error variance, which is *the* goodness of fit of the model: $Rs\hat{}2$ . The normality statistics is the Bowman – Shenton test, 'with adjustment of Doornik and Hansen (1994) based on third and fourth moments of the residuals and having a $\chi^2$ distribution with 2 degrees of freedom when the model is correctly specified' (Koopman, 2000:119). H (j) is the heteroskedasticity; the serial correlations are given by r (j)'s with the number of lags placed in parenthesis right below the values. DW is the Durbin Watson statistics and Q, the Box – Ljung statistics, a test for residual serial correlation.

### *3.7.6.2 Parameter estimates / Components table*

Table 3.7.2 is the tabulated estimates of the variances. It provides the estimates for the standard deviations of the irregular disturbances that drive the level average as well as its growth rate. The non-zero estimates confirm the components that have stochastic trends. For example, the estimated values for the level average are greater than 'zero'. Therefore, it confirms that the underlying averages in all the samples are non-stationary. Consistent with the earlier results obtained in the data exploration in sub-section 3.6, the level average and the irregular for the CP is significantly higher than that for the SMP. This suggests that in relation to the SMP, the CP exhibited higher variability and possibly was not very reliable. This measure of volatility and relative variability in the data exploration results also supports this conjecture.

The 'zero' estimated components are those that are fixed within the sub-samples where they occur. The slope is constant along the estimated growth path in all the estimated samples. The avoidable cost of generation is a short-term factor that affects the slope and the costs of fuel accounts for a significant proportion of this. But a change in capacity is what affects the long run pattern in the slope. The construction of new plants, withdrawal or mothballing of inefficient plants, are some of the factors that cause changes in the level of capacity in an electricity network; and in turn changes in the SMP and CP growth rates. The indirect effect of the long run changes in network capacity is on the CP, because if technological change leads to the more thermally and operationally efficient plants, it will reduce the loss of load probability (LOLP) for each MW of power produced. Of course if the Generators dominant strategies for offering capacity onto the system are not anti-competitive, increases in capacity will decrease SMP. The combination of the reduced SMP and LOLP will lead to reductions in the CP. This implies that the advancement in generation technology has a positive effect on the average costs of producing electricity (see Joskow and Rose, 1985). This is consistent with the 'Real Business Cycles' argument that technological advancement has a permanent effect on macroeconomic trends (see Enders, 2004:157). Harvey et al (1986) also provides an empirical result that supports this intuition.

Since the Regulator used one-year as the benchmark for the short run in the pool, another factor that can affect CP is the changes in value of loss load (VOLL), which was designed to change subject to the retail price index (RPI) and which is an annual index.

| Table 3.7.2<br>**Estimated variances of disturbances** | | | | |
|---|---|---|---|---|
| $lsmp_t = \mu_t + \gamma_t + \varepsilon_t$ | | | | |
| | *Irr* | *Lvl* | *Slp* | *Sea* |
| January 1994 to December 2000 | 0.00050721 (0.0439) | 0.011553 (1.0000) | 0.00000 (0.0000) | 0.00000 (0.0000) |
| January 1998 to December 2000 | 0.0021749 (0.2819) | 0.0077164 (1.0000) | 0.00000 (0.0000) | 7.5356e-007 (0.0001) |
| $lcp_t = \mu_t + \gamma_t + \varepsilon_t$ | | | | |
| | *Irr* | *Lvl* | *Slp* | *Sea* |
| January 1994 to December 2000 | 0.61133 (1.0000) | 0.51216 (0.8318) | 0.00000 (0.0000) | 0.00000 (0.0000) |
| January 1994 to December 1997 | 0.64660 (1.0000) | 0.58290 (0.9015) | 0.00000 (0.0000) | 0.00000 (0.0000) |
| January 1998 to December 2000 | 0.61956 (0.8890) | 0.69691 (1.0000) | 0.00000 (0.0000) | 0.00000 (0.0000) |

Notes: (1) 'Irr' = Irregular, 'Lvl'=level average value 'Slp'= the slope and 'sea'= seasonal. (2) The q-ratios are in parenthesis. It is the relative variance, which is the same thing as the *signal to noise ratio*: $\sigma_\eta^2 / \sigma_\varepsilon^2$ but in terms of standard deviation, it is $\sigma_\eta / \sigma_\varepsilon$ (see Koopman, 2000:121).

The seasonals accounts for the potential behaviour of the variable within each season; it is fixed in all the estimations except for the estimated sub-sample for the log of SMP between January 1998.week1 to December 2000.week52 where it is slightly greater than 0. This suggests a slight seasonal variation after 1998. Weather and environmental factors are two known causes of changes in consumption and supply patterns in electricity. This study conjectures the climate change, which marked the beginning of warmer winters and hotter summer months, is a contributory factor to the seasonal variation. It clearly marked the emergence of unpredictable weather patterns, which might have been significantly different from the historical SND profiles that the NGC could have used to estimate the aggregate gross demand.

The 'zero' seasonal result in this study is consistent with Harvey and Durbin (1986) who find a deterministic seasonal component in the data for the seriously injured or dead in their analysis of the impact of the seat belt legislation in Britain in 1983.

On a full sample basis, the irregular parameter, which usually reflects market shocks and / or the traders' commercial strategies, suggest that CP had more incidences of spikes than the SMP. This still supports a conjecture that it will be easier to forecast the SMP than the CP. These result leads to the same conclusion that I reached in the earlier data exploration analysis reported in part 3.6. This is consistent with the pattern of the prices and the conclusion that Wolak and Patrick (2001) as well as Fehr and Harbord (1993) made in their earlier investigations on the pool data. The parameter estimate for the irregulars after 1998 shows even worse values for the CP in relation to the SMP. Based on the Grid Codes definition of operating plant availability, as well as the procedure for determining the SMP, this study conjectures that Generators began to use CP more after 1998 to manipulate prices. That is, they resorted to the use of capacity outside the SMP determination process to manipulate prices. This is consistent with the expectation that agents in short-run capacity constrained industries will use their supply functions to earn higher rents (see Tirole, 1998).

*3.7.6.3 Component graphics*

The component graphics in figures 3.7.6 to 3.7.10 are the visuals of the smoothened estimates of the parameters. They reveal how the variances govern the changes within the components over the whole of each sub-sample (Koopman et. al, 2000). Graph A plots the trend in the weekly averages of SMP and /or CP. The trend, which is plotted, as a 'broken' line is the underlying weekly average of CP and / or SMP in £/MWh; and a downward (upward) sloping trend implies a reduction (increase) in the long-run pattern in the weekly average.

Graph B plots the seasonal pattern and shows the regularised cyclical patterns with the peaks coinciding with the periods of highest demand, which are in the winter; and the troughs are the weeks with the lowest demand. The seasonal component in SMP after 1998 is small compared to the level average but it is strong enough to influence the movement in the smoothened estimate plotted in figure 3.7.9.

Graphs C plots the slopes, which are constant throughout its part in all the estimations.

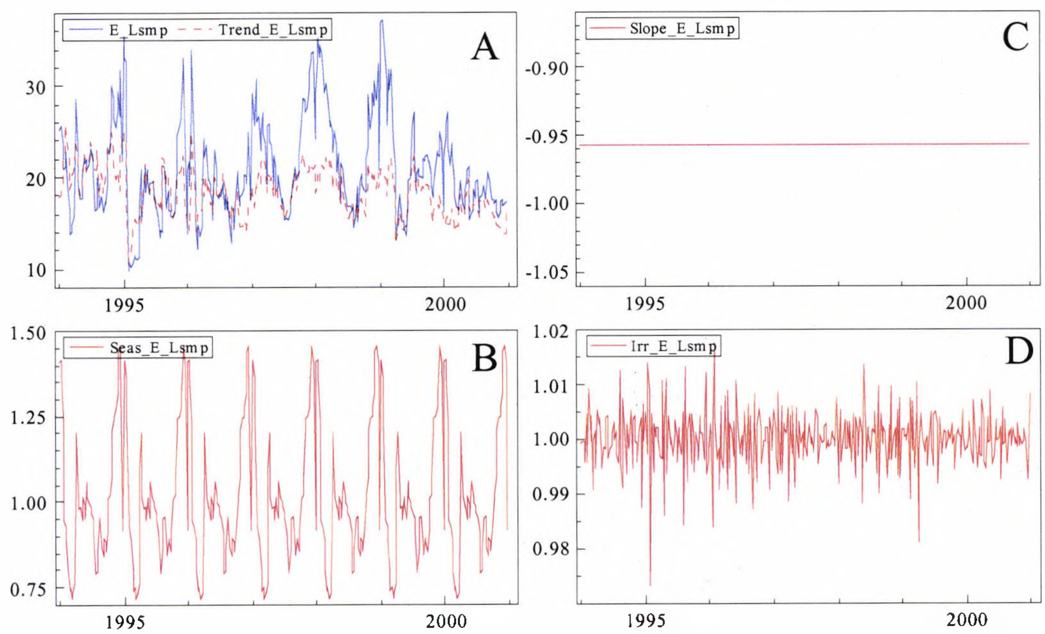
Graph D presents a clearer visual representation of the residuals, which are the relevant values when accounting for the breaks in the trends of the SMP and / or the CP. There is clearly no defined pattern in the spikes; there are runs of very high positive and negative values and in some cases, very high (low) is followed by a very low (high) value.

These types of spikes cannot be removed from electricity prices dataset both in the short or long term; they are indeed, predictable aspects of the trend, which accounts for within the day inter-system transmission shocks and / or the traders' opportunistic behaviour. If it is the case of a transient transmission constraint, it is likely to disappear as soon as the associated technical constraints are resolved. Apart from these, where there are defined temporary or permanent constraint boundaries across a Grid, the Generators that are located behind them are likely to strive to maximise their marginal private benefits by making abnormal offers for their residual and infra-marginal capacities. This is because they know that the SO will require the commodity to keep the system stable and within its tolerance limits. This was the case in England and Wales' pool where National Power persistently offered its Fawley plants, which was located behind constraint boundaries, at abnormal prices.

## Component Graphics of models Anti-Log Analysis

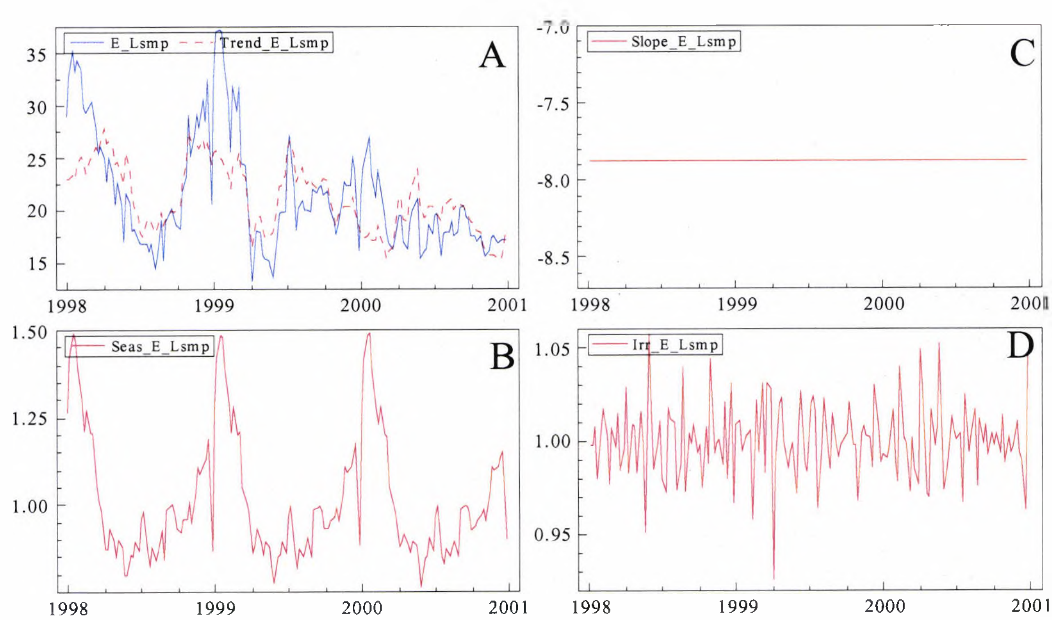
### Figure 3.7.6

### SMP: January 1994wk.1 to December 2000wk.52



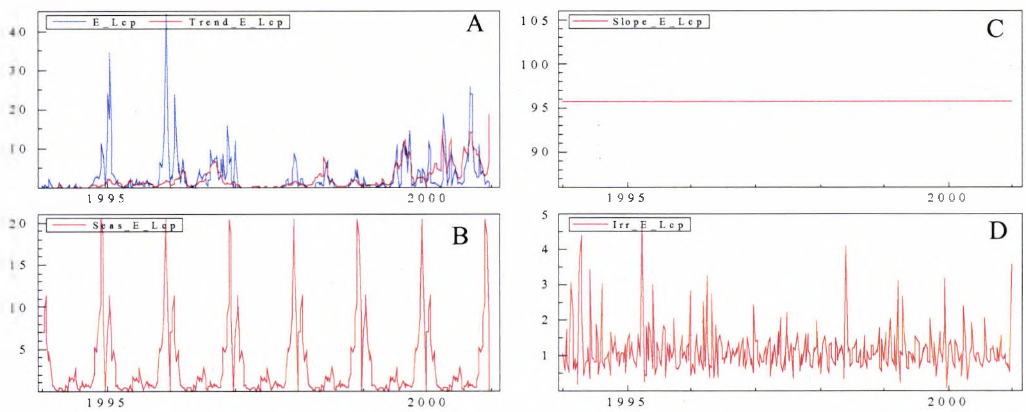

### Figure 3.7.7

### SMP: January 1998wk.1 to December 2000wk.52



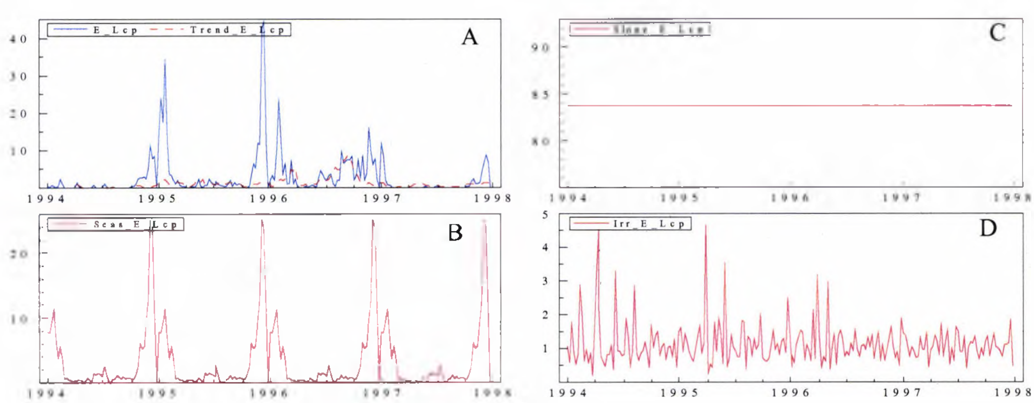
## Figure 3.7.8

## Capacity Payment: January 1994wk.1 to December 2000wk.52



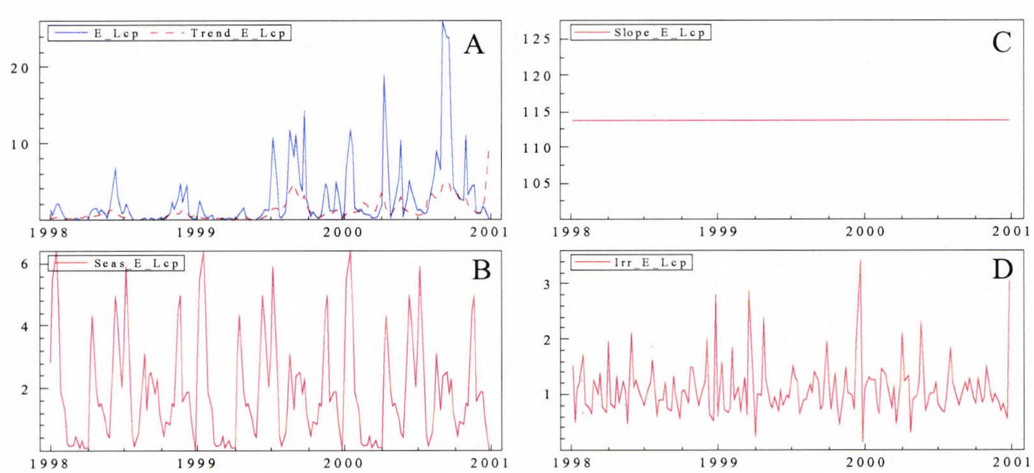

## Figure 3.7.9

## Capacity Payment: January 1994 wk.1 to December 1997wk.52




## Figure 3.7.10

## Capacity Payment: January 1998wk.1 to December 2000wk.52



### 3.7.6.4 *Anti-log analysis*

Table 3.7.3 summarises the anti-log analysis for the models. The full results are attached as table appendix: 3.7.9 to 3.7.14.

**Table 3.7.3**

**Anti-log analysis for the trend and annual growth rates**

| Variable | *January 1994 to December 2000* | *January 1994 to December 1997* | *January 1998 to December 2000* |
|---|---|---|---|
| SMP | £19.08 (-3.13%) | NED* | £19.14 (-16.20%) |
| CP | £9.10 (162.37%) | £0.04 (-106.95%) | £17.62 (195.13%) |

Notes: The slope, which is in parenthesis, is the annual percentage (%) at which the trend grows within each sample period.
*No estimation done, for this sample.

On average, the highest growth rates each year occur during the first four and the last weeks in the year. These are the peak winter weeks when more power is demanded for heating purposes. During these periods, price setting is done in a Betrand fashion (see Tirole, 1998), with the most inefficient plant setting the SMP. The full capacity of the least efficient is fully dispatched and the residual capacity is also met in ascending order of cost for the other Generators (see Fehr and Harbord, 1993). The other feature of this peak winter season is that the full range of the capacity mix across the load duration curve (LDC) will be used to meet demand.

### 3.7.6.5 *Summary*

The underlying average SMP declined between January 1994.week1 and December 2000.week52; but CP grew. They maintained these growth patterns after 1998. There were significant numbers of spikes, which affected the slope of the trend across all the estimated samples. The seasonal pattern was constant in all the samples except for a very slight change in SMP after 1998. The next section investigates whether there were structural breaks in the two series after 1998; and if there were, seeks plausible conjectures for them.

### 3.7.7 Changes in the level average of Capacity Payment (CP) and System Marginal Price (CP) after 1998

This sub-section uses a level disturbance in a smooth trend to estimate the log of CP and SMP between January 1998.week1 and December 2000.week52. A smooth trend is a model in which the level is fixed but the slope is allowed to follow a stochastic pattern (Koopman, et al, 2000). This approach allows for the isolation of the changes / steps that policy changes might have caused on the path of the level average values in the two variables.

The trend component in the 'smoothly trended' model will not contain $\eta_t$, consequently, the variance: $\sigma_\eta^2$, will be 'zero'. The stochastic slope will still have $\zeta_t$, it will have a mean and a variance: $\sigma_\zeta^2$. Finally, the seasonal formulation is trigonometric.

The preliminary investigation involved a normality test for the values that exceed 3.5 in the level and irregular residuals in the generalised Equations10 and 11 (see 3.7.5.5). The results in table 3.7.4 show that the spikes in the level average and the residual irregulars are statistically significant. Therefore, this estimation will use these spikes to disturb the smooth trend in the weeks that they occur.

Table 3.7.4

Normality Test for the large values in the generalised models of the log of SMP and CP

| $lsmp_t = \mu_t + \gamma_t + \varepsilon_t$ | | |
|---|---|---|
| Component | *Period* | *Value* |
| Irregular | 1999.wk14 | -3.7352<br>[0.0001] |
| Level | 1999.wk14 | -3.7708<br>[0.0001] |
| $lcp_t = \mu_t + \gamma_t + \varepsilon_t$ | | |
| *Component* | *Period* | *Value* |
| Irregular | 1999.wk52 | -4.8768<br>[0.0000] |
| Level | 1999.wk52 | -3.7867<br>[0.0001] |

Note: The two-sided p-value is in parenthesis

Equation 12 and 13 are the functional forms of the estimation.

$$lsmp_t = \mu_t + \sum_{j=1}^{[s/2]} \gamma_{j,t} + \sum_{j=1}^{h} \lambda_j w_{j,t} + \varepsilon_t \qquad (t = 1,\ldots,T) \qquad (12)$$

$$lcp_t = \mu_t + \sum_{j=1}^{[s/2]} \gamma_{j,t} + \sum_{j=1}^{h} \lambda_j w_{j,t} + \varepsilon_t \qquad (t = 1,\ldots,T) \qquad (13)$$

Where

$$\mu_t = \mu_{t-1} + \beta_{t-1}, \qquad \beta_t = \beta_{t-1} + \zeta_t, \qquad (t = 1,\ldots,T) \qquad (14)$$

$\gamma_{j,t}$, is the trigonometric seasonal formulation, which equation 5 defines; and '$w_{j,t}$ is an intervention (dummy) variable' (Koopman et al, 2000:142). The SMP intervention is applied in week 14, whilst that of the CP is done in week 52; both of them occur in 1999.

### 3.7.7.2 Results

The estimation diagnostics are satisfactory. It reports **very strong** convergence in the iterations for both models; therefore, a confirmation that they are reasonable. This sub-part can rely on the developments in the industry to conjecture possible reasons for the breaks in the level average of SMP after 1998; as well as the dramatic increase in the trend of CP during the same period.

#### *3.7.7.2.1 Summary statistics*

Compared to the results that were obtained in the generalised model in table 3.7.1, there is a dramatic reduction in the normality values in table 3.7.5. The smoothing effect accounts for the significant decreases in the number of price spikes: irregular values that occurred during that period. The seasonally induced serial correlation is still evident in the models; and is partly due to the independence of the seasonal component with the slope and the underlying average CP and SMP.

### *3.7.7.2.2 Estimates of Components and Graphics*

The estimated values for the variances of the disturbances are shown in table 3.7.6. It confirms that the two models have a slope component, which is the annual growth rate in the trend at the end of 2000. There is a remarkable difference between the slopes of SMP and CP. It is higher in CP than in SMP; this is also a confirmation that at the end of year 2000, CP grew at a much higher annual percentage rate than the SMP. The seasonal pattern is fixed, which means that the pattern of the seasonals is constant between the fifty-two weeks in the sample.

**Table 3.7.5**

**Summary statistics**

$$lsmp_t = \mu_t + \sum_{j=1}^{[s/2]} \gamma_{j,t} + \sum_{j=1}^{h} \lambda_j w_{j,t} + \varepsilon_t \qquad (t = 1,\ldots,T)$$

| Period | *Nos.* | *Std. Error* | *Normality* | *H(j)* | *r(j)* | *r(j)* | *DW* | Q | $Rs\wedge 2$ |
|---|---|---|---|---|---|---|---|---|---|
| 1998 - 2000 | 156 | 0.097 | 2.93 | 0.92 (34) | 0.29 (1) | 0.056 (11) | 1.34 | 20.23 (11,9) | -0.07 |

$$lcp_t = \mu_t + \sum_{j=1}^{[s/2]} \gamma_{j,t} + \sum_{j=1}^{h} \lambda_j w_{j,t} + \varepsilon_t \qquad (t = 1,\ldots,T)$$

| Period | *Nos.* | *Std. Error* | *Normality* | *H(j)* | *r(j)* | *r(j)* | *DW* | Q | $Rs\wedge 2$ |
|---|---|---|---|---|---|---|---|---|---|
| 1998 – 2000 | 156 | 1.22 | 1.29 | 1.40 (34) | 0.28 (1) | 0.077 (11) | 1.39 | 42.41 (11,9) | -0.07 |

Note: Nos, is the number of observations within the sample period. The standard error is the square root of the prediction error variance, which is the goodness of fit in the model: $Rs\wedge 2$ . The normality statistics is the Bowman – Shenton test, 'with adjustment of Doornik and Hansen (1994) based on third and fourth moments of the residuals and having a $\chi^2$ distribution with 2 degrees of freedom when the model is correctly specified' (Koopman, 2000:119). H (j) is the heteroskedasticity; the serial correlations are given by r (j)'s with the number of lags placed in parenthesis right below the values. DW is the Durbin Watson statistics and Q, the Box – Ljung statistics, a test for residual serial correlation.

| Table 3.7.6<br>**Estimated variances of disturbances** | | | |
|---|---|---|---|
| $$lsmp_t = \mu_t + \sum_{j=1}^{[s/2]} \gamma_{j,t} + \sum_{j=1}^{h} \lambda_j w_{j,t} + \varepsilon_t \qquad (t = 1,\ldots,T)$$ | | | |
| | *Irr* | *Slp* | *Sea* |
| January 1998 to December 2000 | 0.0080297<br>(1.0000) | 0.00013797<br>(0.0172) | 0.0000<br>(0.0000) |
| $$lcp_t = \mu_t + \sum_{j=1}^{[s/2]} \gamma_{j,t} + \sum_{j=1}^{h} \lambda_j w_{j,t} + \varepsilon_t \qquad (t = 1,\ldots,T)$$ | | | |
| | *Irr* | *Slp* | *Sea* |
| January 1998 to December 2000 | 1.0268<br>(1.0000) | 0.066351<br>(0.0646) | 0.0000<br>(0.0000) |

Notes: (1) 'Irr' = Irregular, 'Slp'= the slope and 'sea'= seasonal. (2) The q-ratios are in parenthesis. It is the relative variance, which is the same thing as the *signal to noise ratio*: $\sigma_\eta^2 / \sigma_\varepsilon^2$ but in terms of standard deviation, it is $\sigma_\eta / \sigma_\varepsilon$ (see Koopman, 2000:121).

The summarised anti-log analysis in table 3.7.7 provides the real effect in £/MWh of the growth in CP and SMP at the end of year 2000; and it confirms that CP grew at a relatively abnormal rate than SMP.

| Table 3.7.7<br>**Anti-log analysis for the trend and annual growth rates** | | |
|---|---|---|
| *Model* | *Trend* | *Annual Growth Rate* |
| Lsmp | £16.29 | 0.012 (63.26%) |
| Lcp | £29.96 | 0.580 (3014.49%) |

Notes: The slope, which is in parenthesis, is the annual percentage (%) at which the trend grows within each sample period.

The coefficients of the level interventions are statistically significant as shown in table 3.7.8. This means that the intervening disturbance in the trend path in week 14 for SMP and week 52 for CP, were strong enough to cause a step in the smooth level path.

| Table 3.7.8<br>Estimated coefficients of explanatory variables | | | |
|---|---|---|---|
| Lsmp | | | |
| Variable | Coefficient | R.M.S.E | t-value |
| Lvl. 1999.wk14 | -0.49478 | 0.11201 | -4.4171<br>[0.0000] |
| **Lcp** | | | |
| Variable | Coefficient | R.M.S.E | t-value |
| Lvl. 1999.wk52 | -6.4280 | 1.5621 | -4.115<br>[0.0001] |

Note: (1) the 2 sided probability of the t value is in parenthesis. (2) R.M. S. E = root mean square error.

This effect is clearer in the component graphics: figure 3.7.11 and 3.7.12. The smoothened trend of the estimates along the trend path is shown in figure A; the model absorbs the intervention very well.

There is a clearer distinct step in the trend of SMP after the intervention in week 14; thereafter SMP decreases permanently from the fixed level where it was at the end of 1998, onto a lower level. In contrast, figure 3.7.14A picks the CP level intervention in week 52 in 1999. One may be tempted to think that the violent upward surge in the smooth trend at the intervention point is a break; but there is no permanent step increase. It is possible that it is a result of a one-shot transient run, which might have occurred over a few half-hours.

**Figure 3.7.11**

**Component Graphics for the Model of the Log of SMP: Anti-log analysis—January 1998wk1 to December 2000wk52**

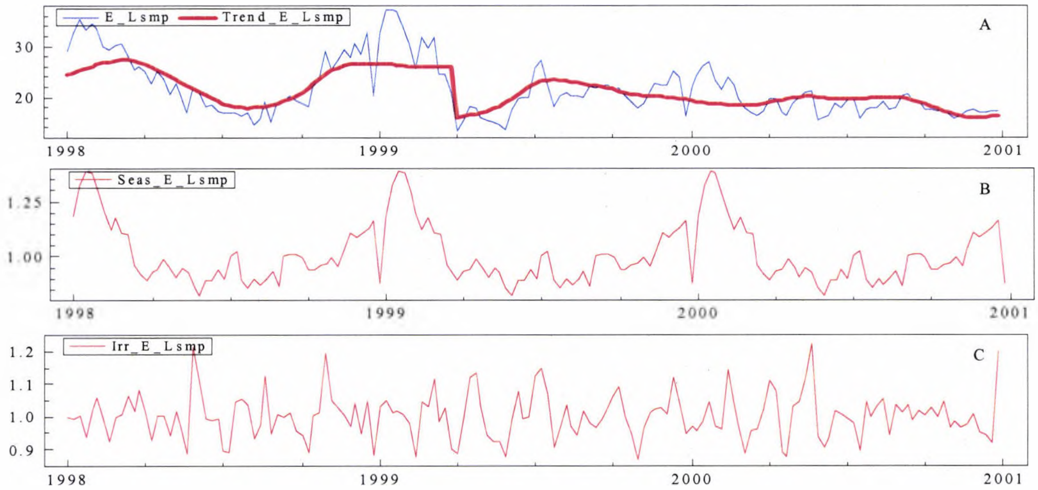
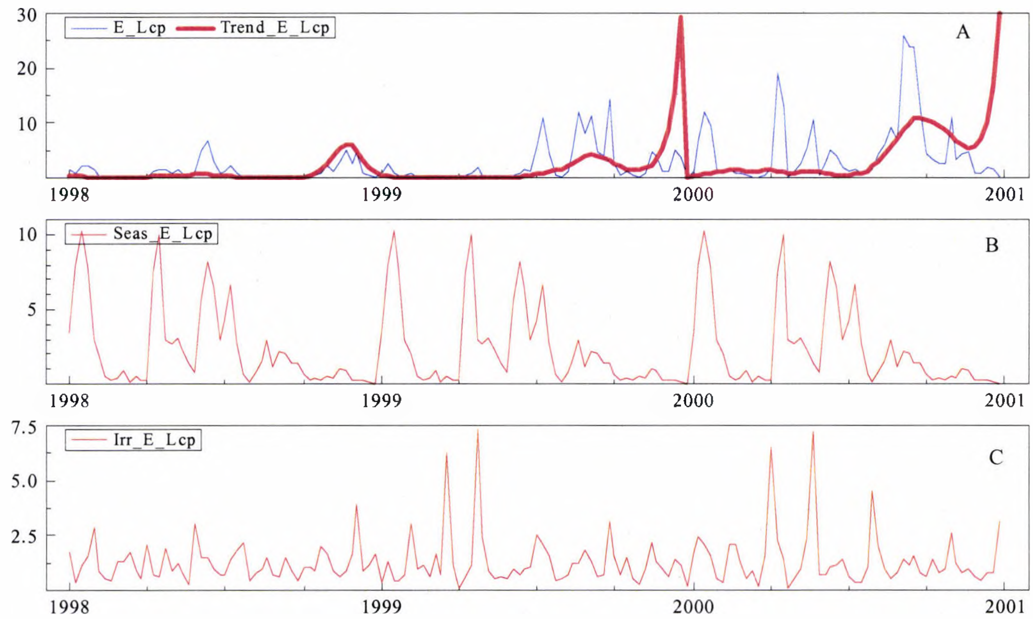


**Figure 3.7.12**

**Components Graphics for the Model of the Log of CP: Anti-log analysis—January 1998wk1 to December 2000wk.52**




*3.7.7.2.3 Summary*

By estimating smooth trend models for CP and SMP with level interventions in 1999, the results reveal a distinct break in the path of the level average value of SMP. There is no structural break in the CP; instead, as at the end of 2000, it grew at an annual percentage rate of approximately 2,951% more than the SMP.

### *3.7.8 Discussion of the implication of the investigation*

The growth in CP is striking and interesting both from a statistics point of view and for public policy. The natural questions that should follow from the estimation results are why was there a structural break in the SMP alone? Why did CP increase dramatically during the same period?

### 3.7.8.1 *Why did CP increase after 1998?*

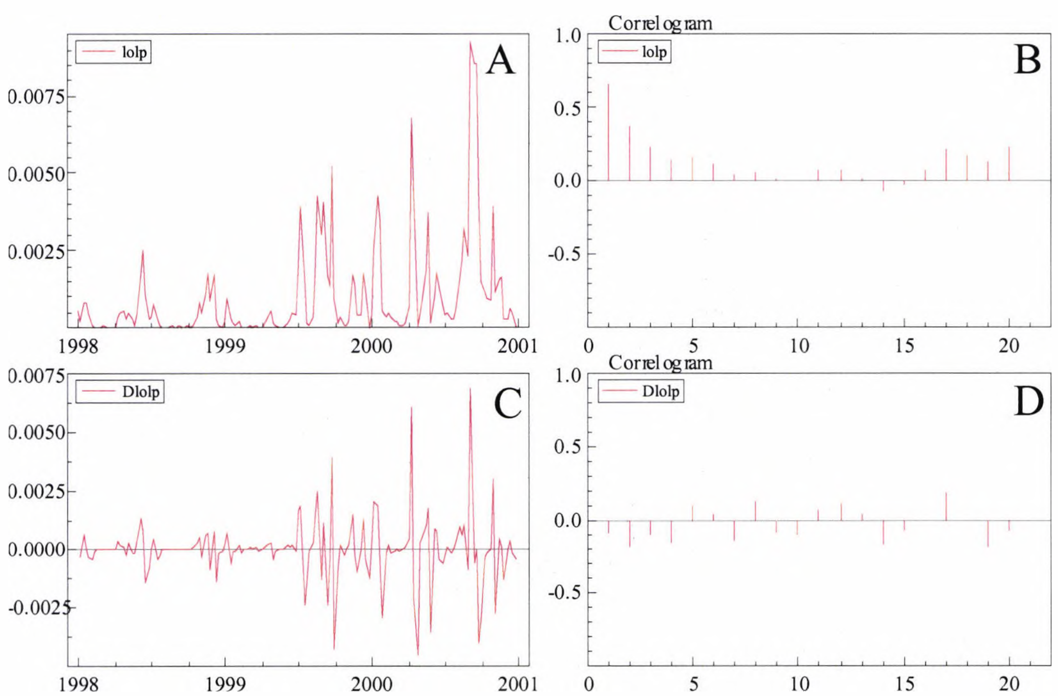
It seems that the best starting point to answer the question about what happened with the CP is its formulation rule.

$$CP = LOLP * (VOLL - SMP) \qquad (15)$$

Capacity is a common factor that will cause changes to the right hand side variables in equation 15. CP will increase if SMP decreases or if LOLP increases. Since SMP followed a constantly decreasing trend from week 14 in 1999, and the VOLL was fixed, the only driver of the higher CPs was increases in the LOLP. The visual analyses of the raw and first difference of LOLP, which figure 3.7.13 presents, confirm the corresponding increases during that period.

**Figure 3.7.13**

**Time series analysis of the actual observations and first difference of LOLP**



The first difference shows that the LOLP began to increase constantly in the third quarter of 1999. An estimation of the LOLP and its smooth trend after 1998 provides an insight into its contribution to the increases in CP.

Equation 16 is the univariate model.

$$llolp_t = \mu_t + \gamma_t + \varepsilon_t \qquad (t = 1,\ldots,T) \qquad (16)$$

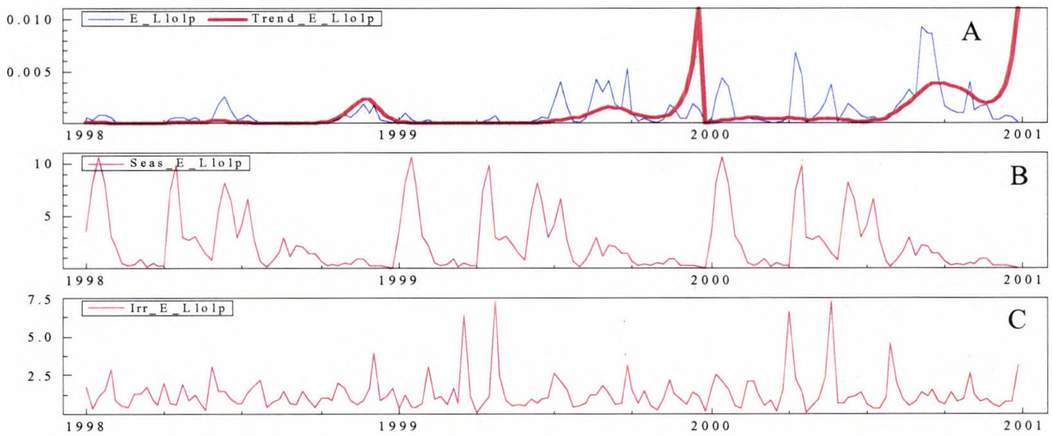
The anti-log analysis confirms that at the end of 2000, the LOLP grew at an annual percentage rate of 192.72% from 1998. The normality test confirmed the presence of large and statistically significant values in both the level average and the irregular residual, which interestingly occurs in week 52 in 1999. This coincides with the large values in the CP. Equation 17 is the form of the smooth trend of the model, in which the level value disturbance was applied in week 52.

$$llolp_t = \mu_t + \sum_{j=1}^{[s/2]} \gamma_{j,t} + \sum_{j=1}^{h} \lambda_j w_{j,t} + \varepsilon_t \qquad (t = 1,\ldots,T) \qquad (17)$$

The anti-log results for equation 17 reports a 3033.68% growth in LOLP. Figure 3.7.14 is the component graphics for the smooth trend estimation of LOLP after 1998. The interesting aspect is that is exactly the same thing as the CP graph in figure 3.7.12.

**Figure 3.7.14**

**Component Graphics for the Model of the Log of LOLP: Anti-log analysis—January 1998wk1 to December 2000wk52**



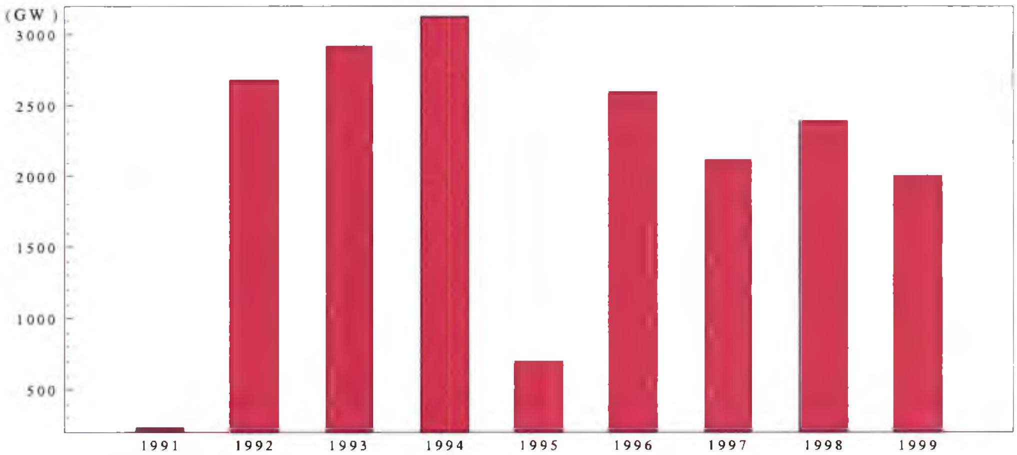
Higher values of LOLP signify the increased possibility that the system might not have sufficient capacity to meet demand increments within the day. The England and Wales' Grid had significant excess capacity and NGC was using more thermally efficient and operationally flexible plant to meet demand after 1998. Moreover, the second phase of divestment had occurred by end of 1999 and meant more owners of the marginal plants. This increase in LOLP suggests that Generators simply used capacity through it to manipulate CP. Bunn and Larsen (1992) predicted that this would happen in the pool. But it seems that one of the ways that the Generators manipulated LOLP in this way was by greater use of the GE inflexibility marker. Here again the DGES already identified that this might be an option back in 1991 (see OFFER, 1991). LOLP will be high even when there is excess capacity, if significant proportions of the plants that are used to meet demand are marked as inflexible. Therefore and subject to an investigation of the actual plant bids as well as the GE markings, it is reasonable to conjecture that after 1998, the Generators took advantage of the loopholes in the Grid Codes plant availability declarations to manipulate capacity.

##### *3.7.8.1.1 Entry: more thermally efficient and flexible plants*

The incumbent Generators made high profits and which, combined with the decreasing trend in the costs of entry gas-fired generation, signalled the massive entry of the *independent power producers* (IPPs), most of whom invested in gas plants. Apart from Sizewell B and the upgrading of part of the Scottish interconnector capacity (CC, 2001), approximately 18,702GW of CCGTs as figure 3.7.15 shows, came on-line across the network in the 1990s. By the year 2001 and compared to 'One' plant that was on-line at vesting in 1990, 31 gas-fired electricity generation plants had direct connections to the gas NTS (Transco, 2002).

**Figure 3.7.15**

**Combined-Cycle Gas Turbines (CCGTs) Commissioned in the 1990s**




** The 1997 figure consists of the 1996 / 97 1,360 from Didcot B and the 760 from Rocksavage

Data source: OFGEM/ 2001

By 1999, full retail competition had occurred in gas and electricity; the *initial portfolio* (IP) contracts that Government set in place between the Regional Electricity Companies (RECs) and the Generators: National Power (NP) and PowerGen (PG), at the industry's privatisation had expired. Most of the RECs that diversified and invested in gas plants jointly with some of the IPPs, began to use their own plants, which were predominantly located within the southern *local distribution zones* (LDZs), to meet their contracted positions. This is evidence that demand was being met with higher efficient and operationally flexible plants. Therefore, the LOLP ought to have decreased and not increased after 1998; it further supports that the high LOLP might have only been possible through the abuse of the plants availability declaration and the offer prices.

### *3.7.8.2 Why did SMP fall after 1998?*

The emphasis on the reduction in the pool prices after 1998 has largely been focused on the decreases in the PPP and PSP. Since SMP forms the greatest proportion of these two variables, any research that focuses on them will conclude that pool prices reduced after 1998. When consideration is given to the individual components that went into the prices setting, this study shows that SMP decreased after 1998, but that LOLP was a factor that drove the higher values of CP after 1998.

There are a number of reasons put forward, which cumulatively led to the reductions in pool prices after 1998; these are the same reasons for the reductions in the SMP. Mr. McCarthy, the then DGGE, upheld that the pool prices began to drop in 1998 as soon as the announcement of the model to trade outside the pool was made (see OFFER, 1998; OFGEM, 2000 a & b). Evans and Green (2003) find increases in the number of players, therefore competition, leading to the decreases in prices. Bower (2002) upholds that increased levels of mid-merit competition, which might have resulted from the divestment of NP and PGs plants as well as IPP entry, was an important factor. He goes further to conjecture that the removal of the gas moratorium in November 1999 was also an important driver for the reductions in prices. This study does not think that the removal of the gas moratorium had any economic effect on electricity prices. This is because the processes for securing a gas-plant licence, raising debt finance, organising a *Network Exit Agreement* (NExA)[22], construction of the plant and finally its commissioning, takes more than two years. It is actually more appropriate to consider that the ban on gas licensing, which the Government implemented between October 1998 and November 1999 was part of the panic response to the dynamic and possibly chaotic nature of the energy market at that stage.

Apart from the issues mentioned in the last paragraph, I learned from discussions with some of the participants at the Association of Electricity Producers (AEP) workshop in 2000, that the lower prices were all part of the NETA 'quiet'. There was only 5% of the aggregate consumption in the industry, which purchased through the pool, and the RECs usually entered into at the very least, one-year contracts with their customers. Therefore, consumer demand was not responsive to the volatile movements in the pool prices. Some of the participants explained that a number of them slowed down their participation in the market because they were concerned about the type of regulatory oversight that might follow after the implementation of NETA. The people that I spoke to were particularly concerned that if the CC upheld the inclusion of MALC into their licences that it might embolden the DGES to make more market rules licence conditions. They expressed the view that these markets abuses relating to operating clauses that are intrusive and might simply subject them to the DGES' *'good moods'*.

[22] 'Network exit agreements (NExAs) are agreements between Transco and other parties that set out terms and detailed provisions for taking gas off the system. NExAs generally include contractual requirements for the provision of certain information relevant to the use of the Transco system and may also define physical aspects of connections' (Transco, undated: 4).

Apart from the on-going MALC case at that stage, the DGGE was also considering options to harmonise the balancing periods between gas and electricity, presumably as a way to curtail the impact of changes in gas usage within the day on the security of the NTS. I had learned from discussions with staff at the Department of Trade and Industry (DTI) in April 2000 that there were technical constraints that limited the harmonisation of balancing periods between gas and electricity. But some of the colleagues that I spoke to at the AEP Workshop, felt that they would incur huge transaction costs to renegotiate contracts following any form of harmonisation of balancing periods that the DGES might put in place.

The other issues that they raised were related to the risks that they might face in the interaction between the New Gas Trading Arrangements (NGTA)[23] and NETA, which are discussed below.

*3.7.8.2.1 Operation risks associated with changes in gas trading arrangements*

NGTA was implemented with the first auction of the six-monthly entry capacity from the beach terminals onto the NTS, in October 1999. At the second auction, which was held in March 2000, capacity prices rose by over 1000% to what occurred at the pioneer rounds in October 1999 and with Transco over-recovering its annual revenue by about £84 million. These high prices that were realised at the primary auctions had a knock on effect on both the forward and over-the counter markets; it also marked the beginning of increased notional balancing point prices. OFGEM attributed the high prices to the effect of competition and the possibility that shippers might have bid very high prices to guarantee 'firm' rights to flow gas. But it acknowledged that there was some inefficiency in the design of the auctions, which might have contributed to the high prices. These included issues such as using SND methodology to determine the maximum physical capacity that Transco made available during each auction round and the application of reserve prices at some terminals.

[23] The NGTAs process included the introduction of an on-the-day commodity market (OCM), which EnMo, and independent firm operated; introduction of auctions for allocation of *firm* capacity rights from the beach terminals onto the gas NTS and the energy and capacity regimes.

At that stage, approximately 373 modifications had been made to the Network Code; work groups were steering forward other initiatives to fine-tune the regime to meet the developments in the industry. Clearly the market had not quite settled down, thus the participants were concerned about the viability of their operations both in gas and electricity; given the persistent changes that were being made to the capacity and commodity regimes.

*3.7.8.2.2 Threats of opportunistic behaviour post-Neta*

The energy market in the UK consists of active gas and electricity markets that have different balancing requirements. Gas faced an end-of-day balancing (EoD); whilst in electricity, Generators were expected to balance their injections with off-takes on a half-hourly basis. An out of balance shipper or Generator was subject to a penal cash-out, at the end of the balancing period within that market segment. The implication for shippers was that they could be out of balance within the day, if they brought in their fully day-ahead nominated gas by the EoD.

I also learned from the participants at the AEP Workshop, that there were concerns that an interactive NETA and NGTA might be biased in favour of the portfolio plants.[24] This concern did not seem justified because NETA was designed with the view that there might not be any systematic arbitrage opportunities between its sub-markets. There were no barriers to entry into the gas and / or electricity markets; and at that stage, the energy market participants already held both shipper and Generators licences. The view was that the difference between the balancing periods in gas and electricity would create profitable arbitrage opportunities for the dual-fuel plant owners. They could sell cheap gas as electricity within the day; provided that they were in balance at the EoD, they might not be exposed to penal charges in gas and over all, they would maximise profits from participating in both markets.

*3.7.8.2.3 Concerns about emergence of efficient competition post-NETA*

The small Generators, renewables and *Combined Heat and Power* (CHP) mainly embedded within the *local distribution zones* (LDZs), were also concerned that they might not be in a position to compete effectively in the interaction between NGTA and NETA. They raised the

[24] The theme of the AEP workshop was to assess and have a better understanding of the challenges that the participants might face in the energy market post NETA.

same issues which the DGES highlighted in his first consultation in 1994 on trading outside the pool (see OFFER, 1994). The broad issue of concern was with the design of the NETA, which included penal cash-out for energy imbalances. The renewable Generators were concerned that they would incur charges for imbalances, which technically, are outside their day-ahead planning. They cannot control the sources of their input. For example, if there is no wind, it will be impossible for a Renewable wind Generator to meet its notified injections. This means that such a Generator would be penalised to an imbalance that is not due to inefficiency. The small Generators felt that NETA was designed to penalise shareholders for *acts of god*. Following from that was the timing for gate closure that was initially set at 3½ hours before real time. Most of the industry participants felt that a shorter gate closure might curtail cash-out charges because the renewable Generators might have extended time during which they could confirm their positions. Some of the smaller traders, who might delegate their trading rights to third parties under the *consolidation arrangements,* were concerned that the lack of funding to procure at the very least 'read only' screens might inhibit them from participating effectively under NETA.

#### *3.7.8.2.4 Additional transaction costs post-NETA*

There were a lot of initiatives that were being taken forward in the England and Wales' energy market after 1998. By 2000, the Regulator embarked on an initiative to implement a form of transmission access methodology, which might be used to manage congestion post-NETA. On a general note, the industry felt that the Grid was not constrained to the level of incurring additional and indeed huge expenses to steer forward a transmission access methodology. The main concern was that apart from the costs they had already anticipated for changing over to new systems and staff retraining, there were still other costs that they might have to account for post-NETA. They felt that the renegotiations of contracts were imminent post-NETA; other costs that might arise from new initiatives such as transmission access methodology were not clear.

Based on the areas discussed above, most of the industry players thought that it was best to re-direct resources towards preparation for the implementation of NETA. Since there was very little dependence on the pool except for fine-tuning positions closer to real-time, few of them participated in the pool from 1999.

### *3.7.9 Conclusion*

This study has used the three-component variance decomposition approach as a tool to understand the trend-cycle, seasonal properties and irregular components in the SMP and CP between January 1994 and December 2000. The Maximum Likelihood technique was used in the Kalman Filter to obtain the estimates. The results confirmed the reliability of the estimations.

The trends in our estimated models are non-stationary; the slopes are fixed along the trend path, and the irregular components are high. Apart from January 1998.week1 to December 2000.week52 where there is a slight non-zero seasonal component, which was strong enough to cause the variations shown in figure 3.7.9B, the seasonal components in all the estimations are zero. Thus, the inherent features of our models are not common; it is usual to find models in which there is both a trend and seasonal variation but in which the $\sigma_{\varepsilon}^{2}$ is relatively small (see Franzini & Harvey, 1983:678). Most of the shortcomings revealed by the statistics are the result of the peculiar features of electricity data. It further supports the use of 'tailored' statistical techniques and models, which are significantly different from the standard ones and robust enough to capture the salient features of the series, to analyse electricity prices.

The investigation finds a structural break in the trend of SMP after 1998, but that the increases in the LOLP had caused CP to increase at an increasing rate during the same period. The decreases in the mean values of SMP might be due to the increased number of Generators that owned the mid-merit plants during that period. There might have also been some influence from the uncertainties that the industry participants faced regarding their operation and volumetric risk post-NETA. I was also aware that the DGES' was in constant communication with some of the dominant industry players, on the need to reduce their offer prices into the pool. This is consistent with (Green, 1999).

The CP result suggests that the Generators used capacity to manipulate prices. This conjecture is consistent with the earlier work by Wolak and Patrick (2001). The CP might have reduced after 1998, if there were initiatives put in place that reduced the VOLL; or possibly the modification of the way in which the LOLP was calculated. It is also possible that the CP might have reduced if the way the component of the offer prices was set was also modified. The present study shows that it is possible to use the three component variance decomposition approaches to understand the impact of the rules and the design of the pool on prices. None of the earlier studies on the pool that I reviewed in part 3.2 used this method; therefore, it is an original contribution in the literature on production and allocative inefficiency in the pool.

Finally a further bivariate and multivariate analysis of the unobserved components in the pool prices setting variables may provide more information about the regime. The next section examines the causes of SMP spikes.

# 3.8

# The Economics of the Day-Ahead Equilibrium Price:

## market forces or monopoly power?

*Abstract*

*This study is an investigation of the causes of the spikes in the system marginal price (SMP) between January 1994 and December 2000. It finds that the spikes did not reflect the demand and supply as well as the system security conditions. Therefore, it conjectures capacity manipulations, based on the high values of loss of load probability (LOLP) and capacity payments (CP) during the half-hours that the spikes occur. The very wide deviation between the day-ahead gross demand and declared availability also supports the conclusion of capacity manipulations. This study concludes that there were other exogenous factors that played major roles in the market-derived SMP; therefore, it upholds inefficiency in the price setting process.*

*Key words: Capacity Payment, Declared Availability, Electricity Market, England and Wales, System Marginal Price, Uplift*

## Introduction

A number of factors determine the level of competition after an electricity market reform occurs. Some of these include policies that are adopted for treating the stranded assets that may become uneconomic post-privatisation. How the vertical integrated structure is broken up and the horizontal merger arrangements that will be put in place will define the ownership of the marginal plants and the capacity mix. The tariff levels, trends and policy for dealing with the proportion of costs that the wholesale suppliers are allowed to pass on to their captive customers will influence best practices and the efficiency of the distribution companies. Finally the methodologies that will be adopted for determining the maximum physical capacity on the network, combined with the mechanism that is chosen for allocating the resource to the Grid Users, will exacerbate monopoly power.

The main downside to the initial policies in England and Wales was the creation of a duopoly market at vesting in 1990. The 'ownership' of the marginal plant was what mattered; therefore, the length of time it took to divest National Power (NP) and PowerGen's (PG) mid-merit plants influenced the emergence of efficient competition and the pattern of prices. On average, few dominant players controlled the sub-markets along the load duration curve.

Earlier studies on the pool data are firm specific, although they used different approaches, their main focus was on the investigation of National Power and PowerGens profits (see part 3.2). During the pool regime, the Regulator used deviations between his estimate of the Generators' avoidable and input costs, as a measure of the profits that they might have made from participating in the pool. This study is the first non-cost based examination of the price spikes in the pool. It uses the SMP as a proxy for the Generators' commercial strategies and examines the deviation between declared availability and gross demand. Based on the results, it conjectures reasonable commercial explanations for each of the half-hours that a spike occurs. Therefore, this paper provides an insight into the effect of SMP determination process on the traders' ingenuity.

The rest of this section is structured as follows: 3.8.2 summarises what we expect to see in the results, 3.8.3 covers the approach that we adopt for the investigation and in 3.8.4, presents definitions of spikes. We discuss the results in 3.8.5 and 3.8.6 concludes this aspect of the empirical data analysis section of the thesis.

### *3.8.2 Propositions*

Based on the developments in the industry (see section 3.3), this study expects to see the following in the results:

- SMP spikes will not reflect market conditions.
- Relative to 1994 there will be spikes in 1995. In addition, in 1995, there will be high spikes in January and December, reflecting the nuclear plant outages and the relatively low average reserve margins in that year.
- Reduction in the number of spikes after 1998.

### 3.8.3 Methodological approach

First I count SMPs equal to and greater than £60/MWh; figure 3.8.1 shows the monthly frequency of those SMPs.

As expected the figure shows an evolutionary pattern in the emergence of spikes; this suggests that the traders developed the skills with which to manipulate the market as the industry matured. The highest number of SMPs greater than or equal to £60 /MWh occurs in 1998; as expected, and compared to the preceding year, the number of spikes reduces dramatically in 1999. The highest growth rate against a preceding year of approximately 328% occurs in 1997 whilst a decline of approximately 700% occurs in 2000. On average and compared to the period between January 1996 and March 1999, spikes cease to exist in year 2000. This may be a reflection of the increased number of participants created through the second phase of divestitures. It suggests greater competition leading traders to make offers closer to their true costs to guarantee being in-merit. This is consistent with Green and Evans (2003) as well as Bower (2002).

**Figure 3.8.1**

**Monthly SMPs greater than or equal to £60/MWh**

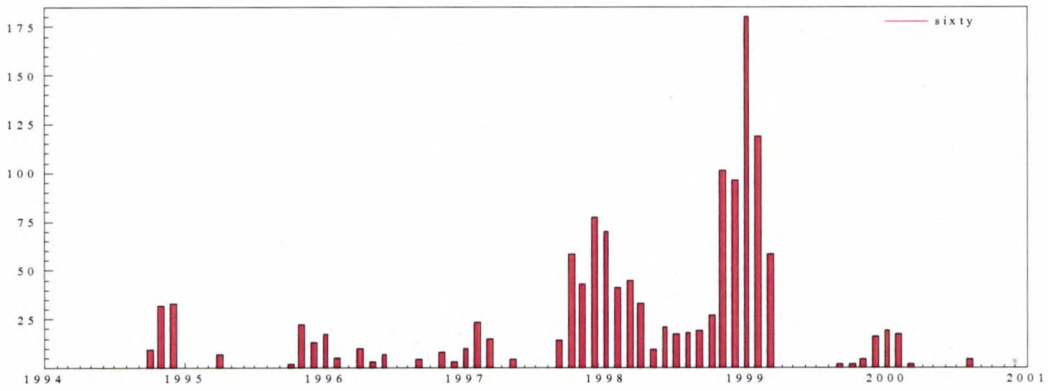
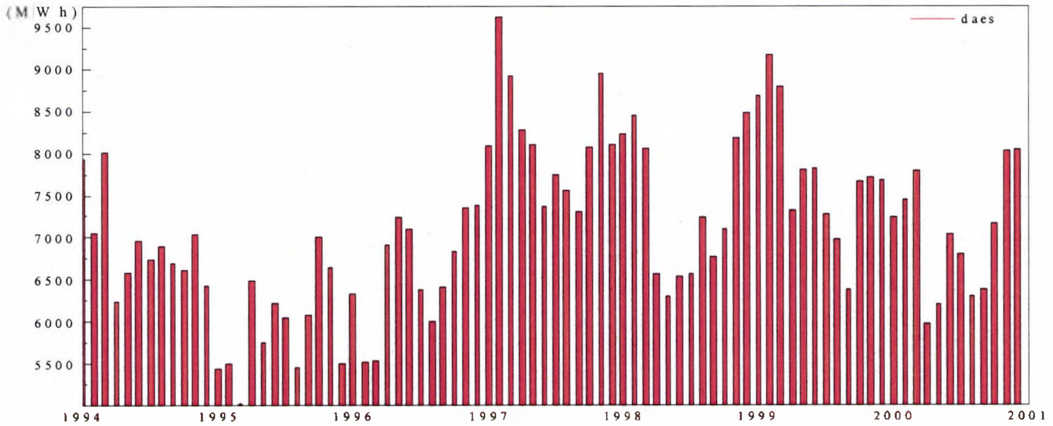



Figure 3.8.2 shows that the lowest average monthly day ahead excess supply occurs in 1995; and in relation to the other years in the full sample, this suggests that the reserve margin on the system was low in that year.

**Figure 3.8.2**

**Monthly day-ahead excess supply: Proxy for Reserve Margin**




1995 and 1998 provide a good basis for assessing the pools efficiency; that is, whether spikes reflected changes in the reserve margin and confirming its ability to incorporate changes in the reserve margin into prices. Assuming that the Generators did not manipulate the price rule, this study expects to find more incidences of spikes in 1995 when compared to the system situations in 1998.

The data exploration section in 3.6 establishes that the pool prices as well as the Generators' behaviour are time, volume and weather dependent. Based on these, it seems useful to investigate if the spikes are also weather dependent. To do that, a reasonable basis for comparison would be between the summer and winter months. Therefore, this study uses the months of June as a typical summer month and November as a typical winter month to compare seasonal trends in SMP spikes. This categorisation is consistent with the DGES' summer and winter months.

### *3.8.4 Definitions of SMP spikes and results*

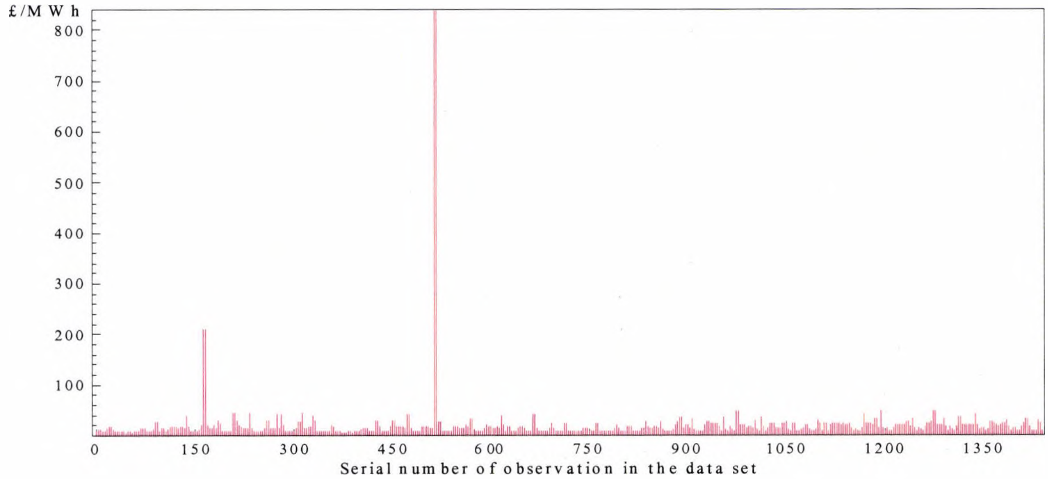
Under the pool a spike occurs if a price is *'three and half times the daily average'*. I adopt three definitions of SMP spikes:

Using the pool spikes benchmark, this investigation:

- Examines day-ahead declared, demand and system security for the highest observed SMP in the full sample: £836.16/MWh. It occurs during periods 35 and 37, which are between 5.30p.m to 6.30pm, on 11 April 1995. Figure 3.8.3 shows another abnormal spike: £211.24/MWh, which occurs between periods 19 to 22; these periods are between 9.30a.m to 11a.m, on 4 April 1995. It investigates whether the same factors were responsible for the spikes on these two days. .

**Figure 3.8.3**

**April 1995: The highest Observed SMP in the Full Sample**

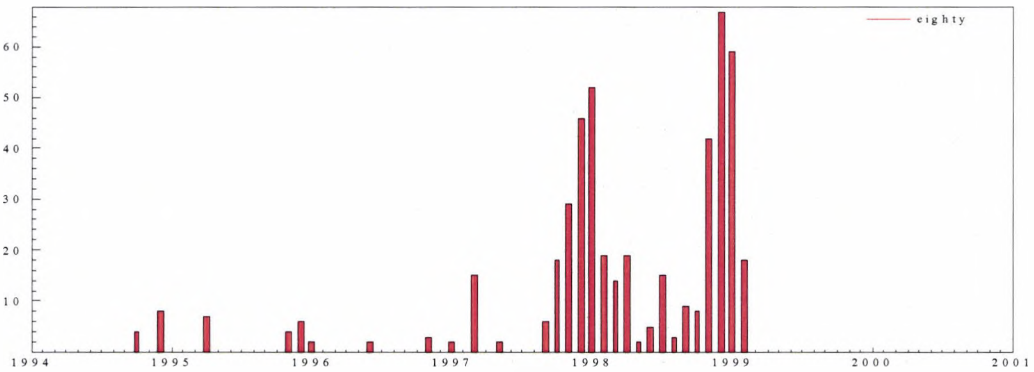


- Examines and compares the spikes in the months of June and November in 1995 and1998. Here again, the focus is on the deviation between the declared availability and gross demand as well as an examination of the system security position. The interesting aspect here will be to see if the low levels of reserve margin in 1995 caused any increases in the spikes. The highest number of spikes occurred in 1998; since it was the year in the full sample with the highest levels of reserve margin , this study investigates reasons why the spikes might have occurred. Table 3.8.1 shows that there were no spikes either in June or November 1995. These results suggest that although reserve margin was quite low in 1995, the capacity available on the system might have been sufficient to meet increments in demand. Put another way, it was economic for NP and PG to withdraw and / or mothball the excess capacities that they had on the system. The result in 1998 is a sharp contrast to what happened in 1995. Table 3.8.2 shows that a total of five spikes occurred in June; whilst eight occurred in November. Recall the loopholes in the rules for setting the SMP and the Grid Codes definition of operating plant availability (see part 3.6), I conjecture that these spikes might have resulted from the temporary intersystem transmission shocks; the traders' opportunistic behaviour is also a consideration.

- This study defines a spike as an SMP equal to or greater than £80/MWh. I obtained the cut-off threshold by approximating three and half times the annual mean for the 7.25 years in the full dataset. Figure 3.8.4 shows the monthly frequency of SMPs greater than or equal to £80/MWh; in tables 3.8.2 and 3.8.3, I tabulate the capacity payment (CP), loss of load probability (LOLP), gross demand, declared availability and actual availability for the half-hours in June and November 1998 when the spikes occur.

**Figure 3.8.4**

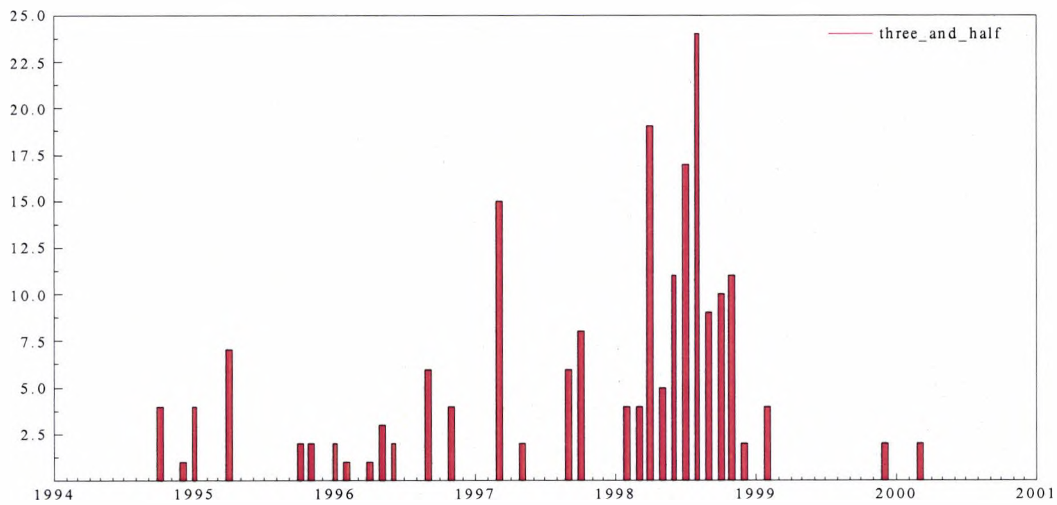
**SMPs Equal to or Greater than £80/MWh**



- The last definition that this study adopts for an SMP spikes is SMP that are three and half times the monthly average. Figure 3.8.5 shows the monthly frequency of SMPs that were three and half times the monthly average values.

**Figure 3.8.5**

**SMPs Three and Half Times the Monthly Average Values**




### *3.8.5 Discussion*

As expected, there are no spikes in 1994. There are some possible explanations, which are consistent with the expectations in this research (see sub section 3.3): (1) the knock-on effect of the transitional arrangements meant that the Generators were possibly still fine-tuning their commercial strategies in 1994. Only 15% of the initial IP CfDs expired in April 1993; that meant that they still had a large amount of physical deliveries locked into take or pay coal contracts. The initial structural arrangements for CCGTs also saw them exposed to very long-term take-or-pay gas contracts, some of which lasted up to 15 years ahead; also, they had associated 'off-take' agreements. (2) The DGES' price cap started in April 1994; it is possible that the need to maintain prices within the limit placed a downward pressure on offers. This is consistent with Wolfram (1999) who finds regulatory oversight placing downward pressure on pool prices. (3) The combined effect of the uncertainties, which the Generators had regarding the regulatory oversight plus the potential opportunistic behaviour of competitors that might happen after the NETA was implemented, influenced the prices trend after 1998. The DGES' ability to identify and name a Generator as inhibiting the emergence of lower pool prices might have influenced Generators' offers. This is also consistent with the effect of regulatory oversight on prices in Hong Kong's electricity market (see Lam, 1999). There was also an effect on the increased number of players brought about by NP and PGs second and 'voluntary' divestiture of their mid-merit plants; this suggests that fierce competition between Generators' to get called bid down prices. These are consistent with the views that industry participants shared with us at the Association of Electricity Producers (AEP) Workshop in 2000.

*3.8.5.2 Highest observed SMP*

The daily average SMP for periods 19 to 22 on 4 April and periods 35 to 37 on 11 April both in 1995, are approximately £30.94/MWh, seven times higher than the daily average and £65.61/MWh, eleven times higher than the daily average values for the respective days. Consequently, they pass the litmus test for spikes in the pool. Do demand, supply and system security positions on those days justify the spikes? The SMP, LOLP, CP, gross demand and Generators: declared, redeclared and actual availability and the table indicators for the half-hours are tabulated in tables 3.8.4 and 3.8.5.

*April 4*

Relative to the other periods within day figures 3.8.6 and 3.8.7 show high CP and LOLP during periods 19 to 22 (163 to 166 in the graph) on 4 April 1995.

**Figure 3.8.6**

**Capacity Payment on April 4, 1995**

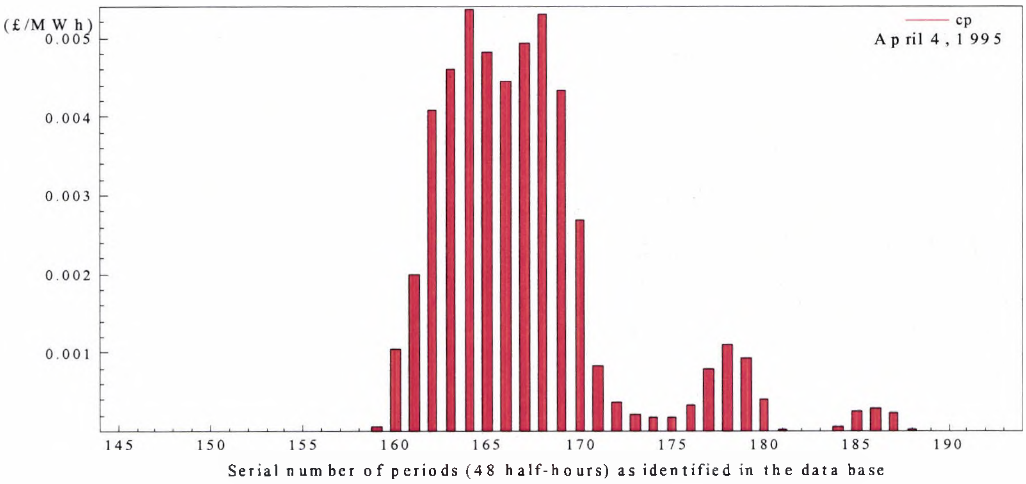
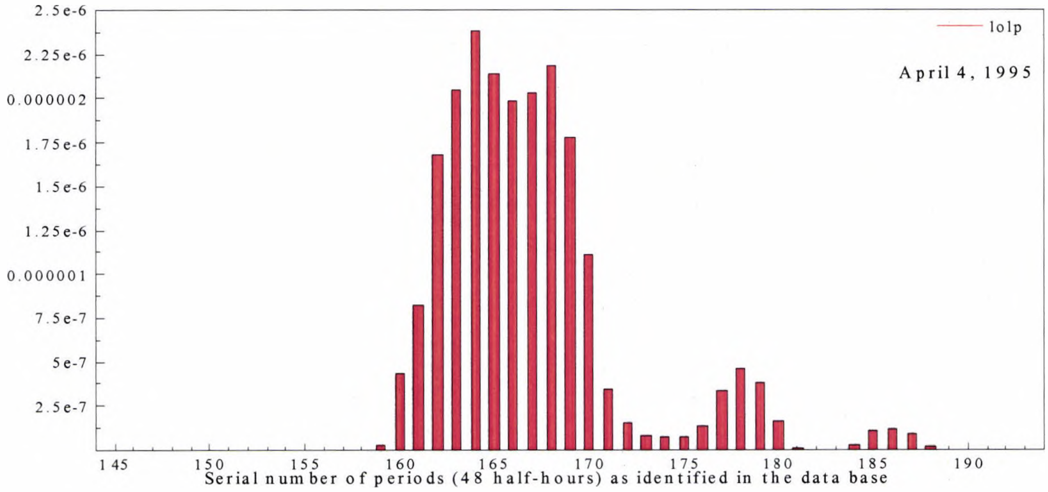



**Figure 3.8.7**

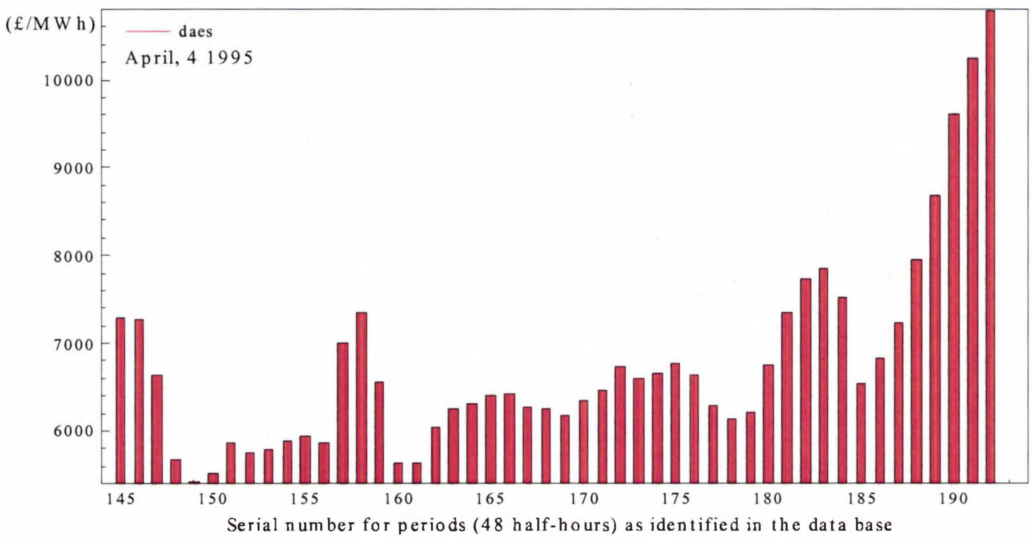
**Loss of Load Probability (LOLP) on April 4, 1995**




The daily average of excess capacity is approximately 6,590MW. The excess capacities for the periods are slightly lower than the daily average, presumably because it is not an early morning peak period.

**Figure 3.8.8**

**Day-Ahead Excess Supply on April 4, 1995**




The above shows no imminent threat to reliability of supply and system security between 9.30 a.m. and 11 am on 4th April 1995, leading us to conclude that there was no justification for the spikes.

*April 11*

In contrast to the situation during periods 19 to 22 on April 4 and in relation to the other periods within day, figure 3.8.9, and figure 3.8.10 show very low CP and LOLP during periods 35 to 57 (515 to 517) respectively [see tables 3.8.1 and 3.8.2]. Table 3.6.5, show over 30% excess capacity during periods 35 to 37 on April 11, 1995.

**Figure 3.8.9**

**Capacity Payment on April 11, 1995**

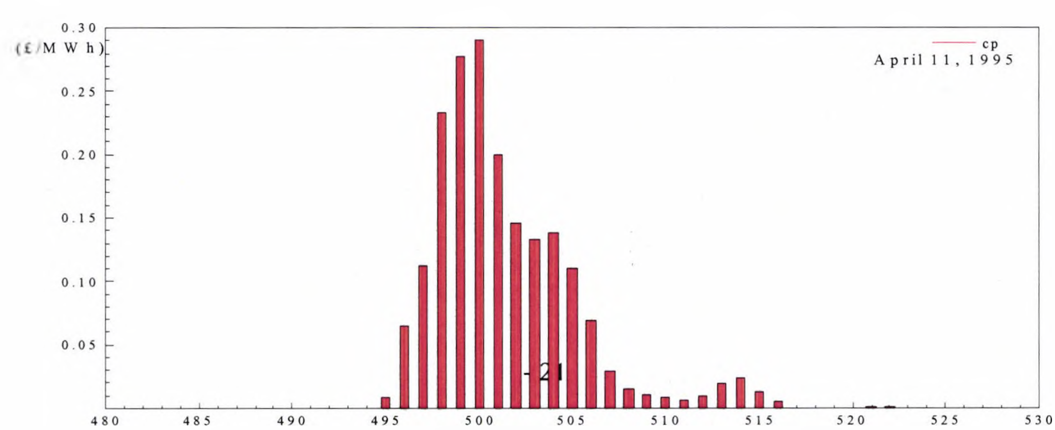
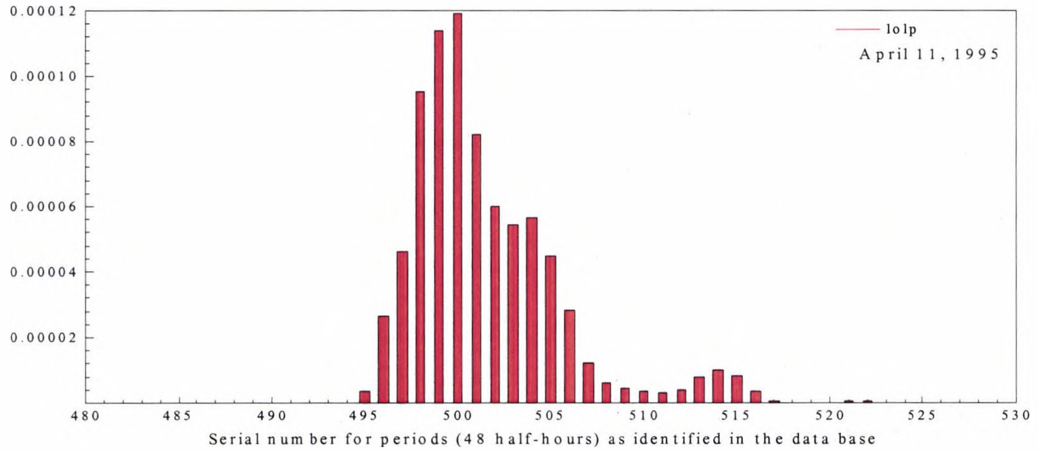


**Figure 3.8.10**

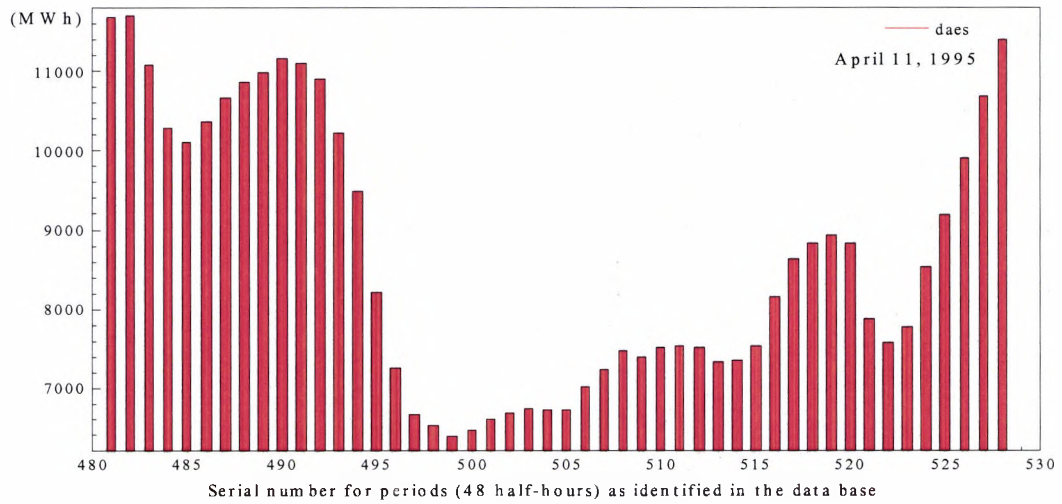
**Loss of Load Probability (LOLP) on April 11, 1995**




The excess capacity for periods 35 is slightly below, whilst those for periods 36 and 37 are far above the daily average value of 7,603MW.

**Figure 3.8.11**

**Day-Ahead Excess Supply on April 11, 1995**




Here again, we do not find any evidence that justifies the phenomenal: £836.16/MWh, spike in SMP.

*3.8.5.3 Comparative examination of the spikes that occurred on 4 and 11 April, 1995*

On average and compared to the three half-hours immediately preceding and after periods 19 to 20 on April 4, declared availability is excessively above gross demand; whilst reserve margin exceeds the notional 20% that is required to maintain safety and security on the transmission system (see NGC, 2000). This demand and declared availability situation is the same for periods 35 to 37 on April 11, 1995; although on average, capacity levels were lowest in 1995, there was over 30% reserve margins during the spikes on April 11.

*SMPs three and half times the daily average value: 1995 and 1998*

I first consider the results obtained within the years; thereafter, compare the spikes in the same month but between the two years.

Compared to June 1995, reserve margin decreased by approximately 9.7% in November 1995. In theory, increases in the daily average of SMP in November 1995 when compared to June will be reasonable. Furthermore, and in relation to 1994 since capacity decreased in 1995, it is also expected that there will be spikes in 1995. Compared to June 1995, what is outside the scope of this paper is the determination of whether the 31.9% increase in SMP in November is comparable to the level of capacity decrease.

Compared to June 1998, reserve margins increased by 1% in November; so in theory the increased levels of capacity should induce prices reductions. But instead of a decrease in price, the results show that the daily average SMPs increased by approximately 41.11%; there was also a 60% growth in the number of spikes in November.

Comparing the number of spikes within the same months [see tables 3.8.1 and 3.8.2] but between 1995 and 1998, I find that the daily mean of the reserve margin decreased by approximately 1.22% in June 1998 than in June 1995; whilst the daily average SMP increased by 1.62%. There is a 10.45% increase in the number of spikes that occurred in November 1998, when compared to the same period in 1995. The mean value of the SMP did not only increase by approximately 8.65%, the frequency of spikes also increased.

*3.8.5.4 Other definitions of spike*

There is tremendous variability in the outcomes across all of the definitions for SMP spike that are used in this section; nonetheless, there are some common patterns in these spikes. They occur in blocks and within very few half-hours; suggesting that in some cases they might be due to transient market shocks. My experience at the NETA simulation games was that on average, the traders' (including myself) submitted the same price across all the half-hours when we played the uniform SMP rule. However, when we played the discriminatory pay-as-bid (PAB) rule, the strategy was always to segment the load profile into peak and off-peak; and once that was done, to submit prices in an ascending order along the load duration curve. That is, on average, we submitted the highest prices for the peak periods and the lowest prices for the off-peak baseload. This strategy is consistent with Bower and Bunn (2000). The settlement information always showed similar prices in blocks and it was quite easy to identify each trader's bids and offers, essentially because we always had to have some type of odd decimated prices to enhance the correct identification of which trader owned what bid or offer. Based on that experience, it is fair to conjecture that the spikes in the dataset, which are identical were submissions made by one Generator.

All the spikes occurred during the table A indicated half-hours; also spikes were more prevalent during the winter. The result leads this study to conclude that spikes are time and weather dependent, meaning that they will be high when the system is under demand stress. Provided that a system has the notional 20% reserve margins required to maintain the system security position; traders do not engage in anti-competitive practices, and there are no geographical transmission and capacity constraints, spikes should not occur at nights.

A higher capacity payment coincides with the periods of the higher incidences of price spikes; this means that they are higher in the months of November than in June. Since more power is demanded during the winter than the summer, the higher capacity payment reflects the increased number of Generators that have to be paid to make their plants available.

The results in this study confirm that SMP in the England and Wales' pool did not reflect market fundamentals. It is also possible that the peaking plants such as Edison Mission Energy, Redditch, Indian Queens, Eastern and Brigg, were responsible for these spikes (see OFFER, 1991; OFFER, 1998).

In some half-hours of the spikes, demand is lower than that which occurred during the two preceding or future half-hours; in these situations, the declared availability is also higher. This also confirms that the spikes, which at the very best, might have been due to traders' opportunistic strategies, could not have reflected the market fundamentals.

On average, reserve margin was higher in June than in November; this is reasonable since more power is needed for heating purposes during the winter than what is required for cooling during the summer months. Although the reserve margin throughout the full sample was approximately thirty-four percent averages, there was at least a twenty-percent average reserve margin during each of the half-hours that a spike occurred.

### *3.8.6 Conclusion*

This study has shown that the SMP spikes did not reflect the demand, supply and system security situations; and there might have been some exogenous factors that led to them. The spikes were also evolutionary, meaning that they worsened as the traders' learnt the market rules and possibly re-defined their commercial strategies. It seems that as the Generators perfected their knowledge about the market rules, they began to use some of the loopholes in the declaration of capacity to manipulate SMP. For example, they might have relied on using things like the no load, start-up price, incremental bids and greater than or equal to (GE) inflexibility declaration (see OFFER, 1991; 1994 & 1999). It seems that the verification of the data on the Generators' specific price and quantity offers as well as a further analysis of the incremental bids, availability profiling and use of GE inflexibility markers for the days and periods in which the spikes occurred, will provide additional insight into their effect on the spikes. This is outside the scope of the investigation in this paper.

# 3.9

# Estimating the *time-of-day* (TOD) System Marginal Price (SMP)

*Abstract*

*This study uses ordinary least squares (OLS) structural regression to estimate time-of-day (TOD) system marginal price (SMP). On average, the results suggest that there are other factors other than Uplift and gross demand that determined the SMP. Based on the Grid Codes definition of operating plant availability combined with the loopholes in the components of the capacity prices that the Generators made into the pool, this research conjectures that the TOD SMP in the pool reflected the traders' commercial strategies.*

*Key words: Electricity demand, England and Wales, Pool, System marginal price, Uplift*

**Introduction**

The last section showed that spikes in SMP did not reflect market fundamentals. This sub-section asks what a competitive time-of-day (TOD) SMP would have been, if Generators' did not exercise monopoly power. It uses *ordinary least square* (OLS) regression in structural modelling to estimate the linear relationship between SMP, the dependent variable and Uplift and gross demand, as the independent variables, over the forty-eight half-hours within the day. It examines how these variables moved together during the full years: 1994 / 95, 1997 / 98 and 1999 / 2000. Finally, it provides a firm grounding with which to forecast SMP beyond the year 2000; therefore, gaining a useful insight into whether the *Competition Commission* (CC) upholding no inclusion of a *market abuse licence condition* (MALC) into the Generators' licences in 2000 would have emboldened them to continue to act anti-competitively.

The rest of this paper is presented as follows: 3.9.2 presents the methodology; 3.9.3, the formulation of the model; 3.9.4 the estimation results and 3.9.5 concludes this part.

### *3.9.2 Methodology*

Granger et als (1979)[25] structural regression modelling approach is the methodological basis for this study, which examines the equilibrium market price for the most expensive capacity that was used to meet demand over the forty-eight half-hours within the day. Most electricity markets adopt a two-part pricing regime (see Armstrong et al, 1998; Tirole, 1998; Laffont & Tirole, 1993) as a way to enhance the viability of the Generators off-peak costs of production. This is because the cost structures of plants changes between the day and night time. For example, recall that in part 3.6 we found that pool variables were highest during the peak (day time, table A, peak load regime) and lower during the off-peak (night time, table B and baseload regime). In England and Wales' pool a discriminatory table A and B indicator was used to distinguish between the peak and off-peak half-hours.

The pool was a half-hourly auction mechanism that consisted of forty-eight periods. This study treats each period as a separate and distinct market. This approach is

[25] Granger et al (1979) use structural OLS regression analysis to examine the factors that determine hourly electricity demand for a sample of households in Connecticut, USA.

consistent with the basis for defining target and geographic markets in the application of fair-trading acts and competition law into the energy industry in the UK. In addition to this, it allows the use of structural estimation; thereby curtailing the effect of heteroskedasticity and serial correlation that usually characterises economic time series. Consequently, the approach enhances the reliability of the estimates.

This approach allows us to gain a better understanding of the variations between the peak and of-peak SMPs. In addition, since electricity like other economic time series is inherently heteroskedastic, a structural modelling approach allows us to assume that variables within each half-hour are clusters; therefore, we can assume independence between half-hours. This will make it possible to curtail heteroskedastic-induced mis-specifications and bias in our estimates.

I examine price differentials between the peak winter and off-peak summer seasons. Consistent with Granger et al (1979), December, January and February make up the winter, whilst July and August make up the summer samples in this study. Each sub-sample consists of all the half-hourly observations within it.

I estimate the model for the year's 1994/95, 1997/98 and 1999/2000; these years are chosen because of the following:

- 1995, is the year with the lowest average reserve margin in the full sample; therefore, it allows for the investigation of the effects of capacity shortage on SMP.
- 1998, is the year with the highest number of SMPs that are equal to and above £60/MWh. It is also the year that has the highest number of spikes based on the three definitions of SMP spikes in the last section.
- 2000, enables the examination of the threats of a change in regime on the Generators behaviour.

### *3.9.3 Formulation of the Model*

SMP was the price for the last MW of power that was purchased to meet demand and from the most expensive plant. It was indeed the equilibrium price since it was obtained at the intersection of the National Grid Company's (NGC) forecast of gross demand and the aggregation of the Generators' declared availability. SMP is a good approximation of the Generators commercial strategies; and put another way, if one plots the co-ordinate SMP and declared availability points, the curve can also be a proxy for the industry's supply curve, since a movement along such a curve will depict the changes in supply and the results from corresponding changes in price.

Once the assumption of a supply curve holds, one can rely on economic theory to analyse the factors that can affect the production of electricity. The generalised two inputs Cobb-Douglas production relates input to output; and the functional relationship is defined as:

$$X = aA^{\alpha}B^{\beta}$$

Where X is the rate of output to be produced, A and B are the respective rates of the two inputs produced and $\alpha$ and $\beta$, the respective elasticity's of the output to the input. In theory, the quantity of goods that a producer may be willing to bring to the market will depend on the price of the goods, technology and alternative uses of capital.

In electricity, Generators respond to changes in demand by changing their declared availability; therefore, more power is supplied during the peak than off-peak periods. The technology that is used to produce electricity during the peak is different from the requirement for the off-peak periods. Finally, the need to maintain the safety and security of the system also influences the quantity of real power as well as the transmission services that the *system operator* (SO) can call on to help maintain the network in its acceptable tolerance limits.

The last paragraph means that the quantity of power, which the Generators supply, will depend on the production technology, demand and the SOs requirement for balancing services. Since I assume that SMP is a proxy for the quantity supplied, then it is dependent on the technology, demand and balancing services. The costs of technology, which is the type of plant that is used and the associated fuel cost, is not quite a stand alone cost in the dataset. But it forms part of the Generators offer prices; that is, they will include it in their capacity offers for real power, which in this context will be part of what goes into the SMP determination process. The Generators that had the facilities to provide transmission services contracted directly with NGC, and the total cost of all balancing services is charged to Uplift.

The explanation in the last paragraph leads to a possible functional linear relationship between SMP, Uplift and gross demand (gd). Therefore, the regression equation is:

$$sqrt(smp)_i = \beta_0 + \beta_1 * sqrt(uplift)_i + \beta_2 * sqrt(gd)_i + \varepsilon_i \qquad (i = 1\ldots, N) \qquad (1)$$

Where $\varepsilon$ is the error term, which captures all the effects of the omitted variables that may affect SMP; it has a mean and a variance and is serially uncorrelated. Therefore it is defined as:

$$\varepsilon \sim N(\mu, \sigma^2)$$

$sqrt(smp)$, is the cost of the last increment of power used to meet demand.

*Uplift*, is the cost that NGC incurred to keep the system in balance. It consisted of the costs for resolving constraints and balancing services.

$sqrt(gd)$, is the gross demand, is NGCs inelastic forecast of aggregate gross demand. It is based on historic seasonal normal demand (SND) profiles by pumped storage, large customers that consume 250MW and forecast deviation customers.

The linear regression defines SMP as a function of Uplift and gross demand. It has a disturbance term that captures all the omitted variables in the model.

The standard errors in the model are White (1980)-corrected, which is a direct test for heteroskedasticity. STATA version 8 is the software that is used to run the estimation. Using a robust variance in the modelling allows for the relaxation of the classical assumption of a constant variance; instead to allow the variance to vary by each observation. This approach gives the same point estimates, as that which could be obtained by running the model without a robust variance. Nonetheless, when a robust variance is used, the standard error differs because the confidence intervals are adjusted; thereby giving more reliability of the parameter estimates; and the F statistics become a Wald test. Overall, using White (1980) -corrected standard errors helps this research to curtail the effects of model mis-specifications; it particularly enhances the reliability of the estimates and policy inferences that are based on them.

### *3.9.4 Results & Discussion*

The estimation results are shown in tables 3.9.2 to 3.9.7. The $R^2$ reports the combined explanatory power / proportion of the variability in SMP that Uplift and gross demand explains. There is no clear pattern across the three years. It is higher in the summer of 1995 than in the winter of 1994 / 95; and suggests a higher explanatory power during the summer than the winter in that year. But the opposite holds true in the other two years, where on average; the $R^2$ is higher in the winter and lower in the summer. The F-statistics measure the null hypothesis of $R^2 = 0$ against the alternate $R^2 \neq 0$. Its associated *p-value* provides the plausibility of obtaining a value as extreme as or more than the observed and the significance of $R^2$ at the 5% level of significance. The statistically insignificant, which correspond to small values of F; and which evidences $R^2 = 0$, are boldly highlighted.

The coefficient of the independent effect of Uplift and gross demand is reported when each one is held constant but the other varies. They are all positive, thus suggesting that SMP increases with a marginal change in gross demand and / or Uplift. This makes sense because an increase in demand will mean an upward movement along the LDC; and depending on the original location of the system along its LDC, an increment in consumption can cause an upward movement into a higher load regime. Once that happens, a completely different capacity mix would be required to meet demand.

There is a remarkable effect of Uplift on SMP when gross demand is held constant. In relation to demand, on average, there is larger number of half-hours in which the effect of Uplift on SMP is statistically insignificant. These results suggest that Uplift has no effect on SMP during such half-hours. Put another way, there are clearly other factors apart from Uplift that determined the SMP during these half-hours. One striking factor about the no effect of Uplift on SMP is its evolutionary pattern. Against our expectations, the number of half-hours in which there were no effects increases progressively from 1994 / 95 and throughout the sub-samples analysed; by 1998, there is no effect throughout the forty-eight half hours within the day. This is a clear suggestion that there were other exogenous factors that determined SMP; and based on the literature on the industry's evolution, a conjecture that the most important drive of SMP is the Generators opportunistic strategy appears a reasonable conclusion. It is also a sensible basis; Generators spent the early years in the reform learning market rules and developing their commercial strategies. Their manipulation skills improved as the industry evolved and it was possible that as a result, they devised sophisticated ways to manipulate prices and capacity, which the market-derived prices could not capture.

The standard errors for the coefficients have been calculated based on White (1980) robust variance estimator; thus the variances vary by observation. The standard errors for Uplift are quite outstanding when compared to that of demand; and in relation to gross demand, it confirms the lower trust that can be placed on the estimates for Uplift. It also confirms that Uplift exhibits greater variability than gross demand.

The *p-value* for each coefficient reports the possibility of getting extreme values than the observed; it tests the null hypothesis of the coefficients of Uplift or gross demand being equal to 'zero' against the alternate hypothesis of a relationship between each of the independent variables and the SMP. As expected, the individual coefficients display dual effects; and on average, they are statistically significant in some half-hours and insignificant at 5% level of significance in others. Uplift and / or gross demand does not affect SMP at all, during these half-hours in which they are statistically insignificant; as a result, it confirms the acceptance of the null hypothesis of no effect during those periods. This finding reinforces the view that pool prices did not reflect market conditions but simply the Generators commercial strategies. This view is consistent with the earlier studies on the pool data, which for example Wolfram (1998), Green (1994) and Wolak and Patrick (2001) have carried out.

The preliminary investigation of the strength of the association between the variables revealed that they are positively correlated as shown in table 3.9.1; and confirm multi-collinearity. The maximum correlation within the sub-samples is the one between SMP and gross demand. The corrective approach to multi-collinearity problems is dropping one of the correlated variables so that the results will not be biased. One thing to note here is that the multicollinearity is not a result of wrong model specification; instead, it is due to the nature of the dataset. The strong correlation between SMP and demand is because demand goes into the estimation of SMP. Moreover, SMP increases (decreases) with increase (decreases) in demand. The highest correlation of 0.69 is between Uplift and demand, which occurs during the summer of 1995. This figure is the limit that is set for the level of correlations that will not cause difficulties in estimations. Lind et al (2000) confirms that the lower and upper limits are -0.70 and +0.70. Consequently, it is still possible to rely on the estimates from this model for policy prescriptions.

### *3.9.5 Conclusion*

This study has used structural OLS regression to model TOD SMP for the winter and summer of 1994 / 95, 1997/98 and 1999/2000. On average, Uplift and gross demand explains the variations in SMP more in the summer of 1994 / 95 than the winter in 1995; but from 1997 / 98 as well as 1999 / 2000, the goodness of fit in the winter is higher than that, which occurs in the summer. The null hypothesis of no combined effect of Uplift and gross demand on SMP is rejected during the half-hours when the probability of the F statistics exceed 0.05 at the 5% level of significance. There is no pattern in the trend of the independent effect of Uplift and gross demand when one is held constant and the other varies. This implies that in any market that is modelled after the England and Wales' pool regime, Uplift and / or gross demand may not have any marginal effect on the SMP, both during the peak day time and the off-peak night time. Uplift and SMP will be correlated especially since capacity is the factor that drives both of them. On the other hand, it is possible that since SMP is set on the day-ahead, lower values might induce Generators to manipulate capacity more on the day, so as to earn higher income through Uplift.

| Table 3.9.1<br>**Correlation matrix for SMP Uplift and gross demand** | | | |
|---|---|---|---|
| | *SMP* | *Uplift* | *Demand* |
| *1994 / 95 winter* | | | |
| SMP | 1.0000 | | |
| Uplift | 0.5013 | 1.0000 | |
| Demand | 0.4826 | 0.4513 | 1.0000 |
| *1995 Summer* | | | |
| SMP | 1.0000 | | |
| Uplift | 0.6407 | 1.0000 | |
| Demand | 0.6951 | 0.6851 | 1.0000 |
| *1997/98 winter* | | | |
| SMP | 1.0000 | | |
| Uplift | 0.3395 | 1.0000 | |
| Demand | 0.6151 | 0.3408 | 1.0000 |
| *1998 Winter* | | | |
| SMP | 1.0000 | | |
| Uplift | 0.1222 | 1.0000 | |
| Demand | 0.5089 | 0.3344 | 1.0000 |
| *1999/2000 Winter* | | | |
| SMP | 1.0000 | | |
| Uplift | 0.4971 | 1.0000 | |
| Demand | 0.7762 | 0.3323 | 1.0000 |
| *2000 Summer* | | | |
| SMP | 1.0000 | | |
| Uplift | 0.3700 | 1.0000 | |
| Demand | 0.6425 | 0.4519 | 1.0000 |

# 3.10

# The Relationship between Reserve Margin, Electricity Prices and Gross Demand: A Quantitative Analysis

*Abstract*

*This research uses ordinary least squares (OLS) regression analysis to model the quarterly relationship between reserve margin, the components of pool selling price (PSP) and declared availability. As expected, the coefficients of system marginal price (SMP) and capacity payment (CP) are negative and statistically significant. Therefore, suggesting that increase in reserve margin reduces electricity prices. In contrast Uplift exhibits a dual effect; it is statistically significant in some quarters and insignificant in others. This result suggests that there are other factor other than uplift that determine reserve margin some times; meaning that for example within the day constraints do not affect reserve margins. Finally, the results from this analysis suggests that policy advisors should be aware that the within the day system security costs may not always reflect the thermal efficiency and operational flexibility of plants as well as the excess capacity, on a network.*

*Key words: Capacity Payment, Declared Availability, Electricity Market, England and Wales, System Marginal Price, Uplift*

**Introduction**

There are a number of theoretic and empirical studies that investigate production and allocative inefficiency in the England and Wales pool regime (see part 3.2). In contrast, there is little in the literature on the relationship between reserve margin and prices. But discussions about sensitivity of prices to reserve margin in the evolution of the de-integrated and privatised England and Wales' electricity industry have been documented in the pool literature since 1992. Bunn and Larsen (1992) use a system dynamics modelling approach to investigate amongst other issues, how the regulatory regimes, economics and market structure might affect reserve margins as the reformed industry matured. Also Bunn, Larsen and Vlahos (1993) investigated how fragmentation might affect the investment decisions, of the *Independent Power Producers* (IPPs) in the industry. Both studies show that fragmentation and the dash for the relatively cheap gas will induce large-scale entry of IPPs. Consistent with economic theory of price determination in free markets, (see Sloman, 2000), they expected that the excessive entry-induced capacity would lead to significant decreases in prices. But Bunn and Larsen (1992) find that there is a non-linear relationship between the loss of load probability (LOLP) and prices. Given the rules for setting prices, they predict an incentive for Generators to use capacity to manipulate the *loss of load probability* (LOLP) to gain higher CPs.

Apart from Bunn and Larsen (1992) and Bunn, Larsen and Vlahos (1993), I am not aware of any other study that empirically examines the relationship between reserve margin, prices and declared availability in the pool. This paper fills this knowledge gap.

NGC (2000) defines '*plant margin*' as 'the amount by which the installed generation capacity exceeds the peak demand' (page 4 2). Under the pool they calculated plant margin by comparing *Generator-registered capacity* (GRC) with the highest forecast winter demand, based on *average cold spell* (ACS) winter peak conditions; this included station transformer demand, transmission and distribution losses. I learned from discussions with NGC staff on 4 June 2004 that a number of technical issues mean that plant margin is much more than the deviation between demand and declared availability. I also understand that the closest approximation to plant margin in the dataset is the day-ahead excess supply (DAES), which is the difference between declared availability and gross demand. This research uses it as the proxy for reserve margin.

The rest of this paper is structured as follows: 3.10.2 is the theory that underlies the study; 3.10.3, the propositions and the model is in 3.10.4. The results of the robust regression are presented in 3.10.5; the discussion in 3.10.6 and 3.10.7 concludes this part of the empirical section.

### *3.10.2 Theory*

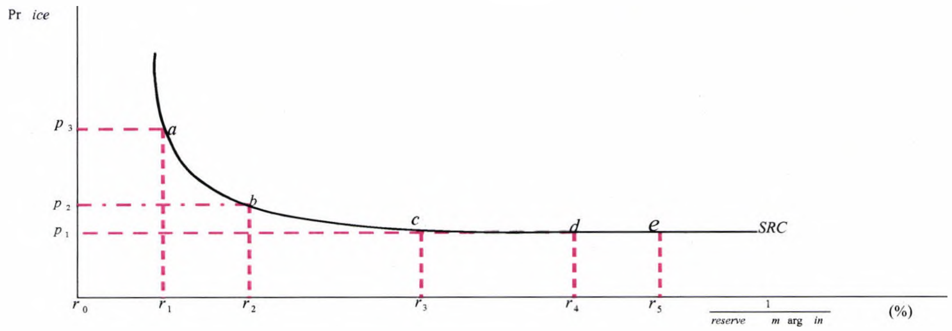
This section uses the invisible hand theorem to formulate expositions about the relationship between price and reserve margin. It makes the following assumptions:

- Generators are free to enter or exit the market
- Firms make rational decisions to invest in capacity.
- GRC changes when new capacity comes on-line; declared availability and reserve margin also change but subject to changes in gross demand.
- Demand is inelastic.
- There is a lag between firms' decision to invest and commissioning of plants; this is due to the timing required to raise debt finance, conclude contracts for input and network exit agreements and the construction of the plants.

Figure 3.10.1 shows the relationship between price and reserve margin. The *security reliability curve* (SRC) is a line that connects the reserve margin at each given price.

**Figure 3.10.1**

**Relationship between reserve margin and electricity price**




Movement along the SRC depicts changes in price. But changes in the system marginal price, LOLP, number of Generators, capacity mix, thermal efficiency and technical flexibility of plants, will cause SRC to shift inwards or outwards.

Suppose this system is located at point $a$, where the reserve margin is $r_0r_1$ and the corresponding commodity price is $P_3 / MWh$. This network has a very low reserve margin; and prices can take any abnormal values whenever the reserve margin on this system is anywhere between $r_0$ and $r_1$. Assuming the following demand and all the other factors that can cause the SRC to shift, are held constant; and there are two developments, which can cause a change in the quantity of capacity available on the system. Generators always make very rational decisions to increase or decrease investment in plants. They can decide to increase investment in response to expected growth in demand. On the other hand, lack of profitability on idle capacity can cause them to withdraw, mothball or close a generation station.

Suppose that the Generators simultaneously invest in plants, upon commissioning, they increase their registered capacity with the SO. This increase in aggregate capacity causes the system to move to a point such as $b$; the overall effects can be seen as follows: (1) GRC increases and subject to changes in demand, Generators will be able to increase their declared availability. (2) The system moves to a new level of reserve: $r_0r_2$, which results from the increase in the reserve margin by approximately $\overline{r_1r_2}$. (3) Price will reduce by $\overline{P_3P_2}$ and the new price level will be $P_2 / MWh$.

Suppose there is further fragmentation and the Generators decide to diversify; this includes decision to make additional investment into plants. If the additional plants come on-line at approximately the same time and causes a further increase in the systems capacity mix. Lets now consider that all the factors, which can cause SRC to shift, are held constant; but the combined investments create a geometric increase in capacity to the extent that the system moves to a point such as $c$. Here again, there are three effects: (1) GRC and declared availability increases. (2) Reserve margins increase and the network will now have $r_0r_3$ of reserve. (3) Price decreases to $P_1 / MWh$. The excessive increase in new capacity causes prices to decrease; however, in relation to the decrease that occurred when the system moved from point $a$ to $b$, there is a smaller change in price: $\overline{P_2P_1}$.

Finally, if firms make further capacity investments' following the plants coming on-line, the system moves to a point such as $e$. Again, keeping all the other factors that can cause the SRC to shift constant, the effect at $e$ is seen only in terms of a change in reserve margin, which increases from $0r_3$ to $0r_5$. But price remains constant at $P_1 / MWh$.

Firms can decide to exit the industry for a variety of reasons, if for instance lack of profitability causes Generators to withdraw, close, decommission or mothball inefficient plants. These will reduce GRC as well as the level of the reserve margin on the network; it will also cause prices to increase. If capacity withdrawals persist until reserve margin in the model is between $r_0$ and $r_0 r_1$ prices will take infinite values, which will reflect the higher probability of the system not meeting increments in demand.

The hypothetical model shows that additional capacity after $r_0 r_3$ will no longer lead to reductions in price. This implies that there is a notional optimal level[26] after which increases in capacity will not cause comparable decreases in price.

The model also shows that prices are higher when the level of reserve margin on the system is low; therefore, in relation to off-peak, prices will be higher during the peak periods. There is also an intuitive inference that a higher impact of a change in capacity on prices will occur during the periods of lower reserve margins than when the system has surplus capacity. This is reasonable since the slope of the SRC in the model is larger at a point such as $a$ than what it will be at, for example, $d$. This means that one should expect a higher effect of a marginal change in capacity on prices during the peak periods, for example in the winter months, than what might occur during off-peak such as the spring months.

Finally, the portion where price remains constant after the system exceeds the notional planning margin is not observed in real electricity markets. Instead, what exists is that once the system attains this notional threshold, the Generators will devise strategies that

[26] The Central Electricity Generating Board (CEGB) sought to achieve a plant margin of 24% several years ahead which it called the '*planning margin* (i.e. the plant margin for planning the need for future generation). An international Review of Adequacy Standards for Generation and Transmission Planning (CIGRE Report No. 37 – 92 (AG) 02 (E), November 1992) shows that large utilities world-wide seek planning margins up to 30%' NGC 2000 p 4 2).

they can use to earn higher prices for the residual demand that they supply to the market. Or better still, for their idle capacity; this is also reasonable if one considers that generation business is solely debt financed. On average the Generators would all have to honour their debt obligations when due; and this happens whether they are in or out of production. By the time the system exceeds this notional limit, there will be a lot of idle capacity; and capacity and prices manipulation will be the only way for the Generators to earn some income on their investments.

### *3.10.3 Proposition*

Prices of economic goods reflect scarcity; therefore, it usually increases (decreases) with decreases (increases) in supply. On the other hand, the suppliers will respond to increases in demand by bringing more goods to the market.

Therefore, this research proposes that:

*The coefficients of SMP, CP and Uplift will be negative; this will show that increases in reserve margin, which in this case is excess supply, decreases prices. But declared availability will have a positive coefficient and in this case, it will confirm that reserve margin increases with increases in capacity.*

### *3.10.4 Formulation of the model*

A reliable analysis of reserve margin for a market where prices are set *ex-ante* should incorporate gross demand and declared availability. This study regresses *reserve margin* $(rm)$against *pool selling price* $(psp)$, this is the price, which the *regional electricity companies* (RECs) and all large industrial consumers pay to take-off power from locations across the Grid. *Declared availability*$(da)$, is the aggregate day-ahead availability by Generators and *gross demand* $(gd)$, *National Grid Company* (NGC) forecasts of gross demand by pumped storage, *non-daily metered* (NDM) sites and large industrial consumers.

The OLS equation is:

$$sqrt(rm)_i = \beta_0 + \beta_1 * sqrt(psp)_i + \beta_2 * sqrt(da)_i + \beta_3 * (gd)_i + \varepsilon_i \qquad i = 1, \cdots, T \qquad (1)$$

$\varepsilon$ is the random term that captures the effect of all the omitted variables in the model. It is normally distributed; it has a mean, a variance and is serially uncorrelated; and therefore, defined as:

$$\varepsilon \sim N(\mu, \sigma^2)$$

This research aims (1) to validate the theory in 3.10.2; and through that, establish the relationship between the dependent and independent variables in equation 1. (2) To ascertain how a marginal change in reserve margin might have affected SMP, CP, Uplift and Generators' declared availability.

I first ran preliminary results of equation 1; on average, the annual and quarterly estimation yielded $R^2$ in excess of .90. This appeared a good result, since it meant a high explanatory power of the model. But demand, declared and reserve margin were highly correlated in excess of $-0.70$ and $+0.70$. Whilst this is the nature of the dataset, demand appeared rather redundant. Since part of the main objective of the model is to assess the relationship between reserve margin and declared availability, it seemed appropriate to drop demand. Therefore equation 2 is the re-written linear relationship.

$$sqrt(rm)_i = \beta_0 + \beta_1 * sqrt(psp)_i + \beta_2 * sqrt(da)_i + \varepsilon_i \qquad (2)$$

Under the pool, PSP was calculated as the sum of: (1) system marginal price $(smp)$, the cost of meeting demand at the margin; (2) capacity payment $(cp)$, the payment made to Generators for making their plants available; and (3) Uplift, the *ex-post* costs for maintaining system security, which included the system operators (SO) costs for *'constraining on'* and / or *'off'* plants, costs for procuring within day transmission services[27] and availability payment. PSP was calculated as:

[27]This consisted of non-real power such as reactive power, reserves (spinning and non-spinning) and black start.

$$psp = smp + cp + uplift$$

I decomposed PSP into its components and re-wrote the linear relationship and equation 3 became the quarterly regression model.

$$\begin{aligned} sqrt(rm)_i = \beta_0 + \beta_1 * sqrt(smp)_i + \beta_2 * sqrt(cp)_i + \beta_3 * uplift_i \\ + \beta_4 * sqrt(da)_i + \varepsilon_i \end{aligned} \tag{3}$$

$$i = 1, \ldots, T$$

And

$$\varepsilon \sim N(\mu, \sigma^2)$$

$sqrt(rm)_i$: is the square root of DAES, the proxy for reserve margin . It is the dependent variable calculated as the difference between declared availability and gross demand. This study examines its relationship with system marginal price (SMP), capacity payment (CP), Uplift and declared availability.

$sqrt(smp)_i$: Is the square root of system marginal price. It is the cost of the last megawatt of power, required to meet demand during each half-hour. SMP ought to reflect Generators' short run avoidable costs and capacity scarcity. This means that it increases (decreases) with decreases (increases) in capacity.

$sqrt(cp)_i$ : Is the square root of capacity payment (CP). CP is calculated as $(VLL - SMP) * LOLP$; it reflects the technical and operational availability of plants; therefore, it was expected to increase (decrease) with decreases (increases) in operational availability and / or withdrawals or mothballing of plants.

$Uplift_i$ : Is all the costs that NGCs incurred to maintain system security in a given half-hour. Uplift reflected the number and severity of balancing actions; therefore, it increases (decreases) with increases (decreases) in the number of threats to the system operations.

$sqrt(da)_i$ : Is the square root of Generators' declared availability. Demand drives declared availability, but the quantity of capacity that a Generator declares is subject to its registered capacity (GRC).

Decomposing PSP enables the investigation of the systematic patterns and changes between reserve margin and SMP, Uplift and CP. It particularly allows through the Uplift, which was an unhedged component in the pool, the examination of the effect of the activities in the downstream commodity market and what happens on the transmission system. Table 3.10.1 lists the goodness of fit for our preliminary run of equations 2 and 3, which shows that the decomposed PSP provides a better explanation of the variation between the dependent and the independent variables.

**Table 3.10.1**

**Goodness of fit for equations 2 and 3**

| Year | *Equation 2 Non decomposed PSP* $R^2$ | *Equation 3Decomposed PSP* $R^2$ |
|---|---|---|
| 1994 | 0.39 | 0.42 |
| 1995 | 0.19 | 0.28 |
| 1996 | 0.32 | 0.41 |
| 1997 | 0.37 | 0.46 |
| 1998 | 0.40 | 0.53 |
| 1999 | 0.41 | 0.52 |
| 2000 | 0.48 | 0.61 |

I reviewed the residuals and tested the validity of equation 3. Table 3.10.2 summarises the residual statistics; it confirms the normality assumption, which was made about the error term.

| Table 3.10.2<br>Summary statistics of equation 3 residuals | | | | | |
|---|---|---|---|---|---|
| *Year* | *N* | *Mean* | *Std. Dev.* | *Min* | *Max* |
| 1994 | 17520 | 1.21e-10 | 6.13 | -23.81 | 26.01 |
| 1995 | 17520 | -7.78e-10 | 7.95 | -39.76 | 77.52 |
| 1996 | 17568 | -1.70e-09 | 7.27 | -37.23 | 27.43 |
| 1997 | 17520 | -1.02e-10 | 6.08 | -31.63 | 33.34 |
| 1998 | 17520 | 1.46e-10 | 6.12 | -35.44 | 43.98 |
| 1999 | 17520 | 5.69e-10 | 6.40 | -31.11 | 29.24 |
| 2000 | 17568 | -3.64e-09 | 5.71 | -30.34 | 25.27 |

Notes: 1. These residuals are obtained by running equation 3 without a robust standard error and the assumption that the observations within each period are clusters. 2. The residuals have a 'zero' mean and range from the minimum to maximum values. 3. The fitted values range from an over estimated = absolute value of the minimum to an under estimate = maximum positive value.

The cumulative probability distribution of the residuals revealed that they are approximately symmetrical around zero in each of the years; therefore if provides an indication that there are as many over estimates as underestimates.

The result of the Cameron and Trivedi (1990) decomposition test for heteroskedasticity, skewness and Kurtosis (imtest) is in table 3.10.3. It confirms the violation of the constant variance in the error term; and suggests that the estimates may be biased. To curtail the effect of heteroskedasticity I re-ran the model, but this time, with a robust variance (see Huber 1997, 1997; Wooldridge, 2002; Carrol et. al, 1998; STATA; page 337) (STATA, V. 8—R: 331-341 & U23: 270-276). I also designated the 'period', that is, the forty-eight half-hours within the day, as the cluster variable.

*Summary*

Reserve margin has a linear relationship with SMP, CP, Uplift and declared availability. The standard error in the estimation will be White (1980)-corrected; and the period will be used as the cluster variable when running the model in STATA v.8. The next section presents the robust regression results.

| Table 3.10.3<br>Cameron & Trivedi's decomposition of IM-test | | | | |
|---|---|---|---|---|
| *Year* | *Source* | $\chi^2$ | *df* | *probability* |
| 1994 | Heteroskedasticity | 1309.85 | 14 | 0.0000 |
| | Skewness | 518.81 | 4 | 0.0000 |
| | Kurtosis | 18.78 | 1 | 0.0000 |
| | Total | 147.43 | 19 | 0.0000 |
| 1995 | Heteroskedasticity | 7313.47 | 14 | 0.0000 |
| | Skewness | 1470.72 | 4 | 0.0000 |
| | Kurtosis | 4.72 | 1 | 0.0298 |
| | Total | 8788.90 | 19 | 0.0000 |
| 1996 | Heteroskedasticity | 724.37 | 14 | 0.0000 |
| | Skewness | 358.00 | 4 | 0.0000 |
| | Kurtosis | 100.30 | 1 | 0.0000 |
| | Total | 1182.66 | 19 | 0.0000 |
| 1997 | Heteroskedasticity | 3332.18 | 14 | 0.0000 |
| | Skewness | 8213.93 | 4 | 0.0000 |
| | Kurtosis | 9.99 | 1 | 0.0016 |
| | Total | 4164.10 | 19 | 0.0000 |
| 1998 | Heteroskedasticity | 4207.02 | 14 | 0.0000 |
| | Skewness | 939.19 | 4 | 0.0000 |
| | Kurtosis | 3.40 | 1 | 0.0651 |
| | Total | 5149.61 | 19 | 0.0000 |
| 1999 | Heteroskedasticity | 1504.74 | 14 | 0.0000 |
| | Skewness | 214.27 | 4 | 0.0000 |
| | Kurtosis | 15.06 | 1 | 0.0001 |
| | Total | 1734.07 | 19 | 0.0000 |
| 2000 | Heteroskedasticity | 1570.38 | 14 | 0.0000 |
| | Skewness | 576.49 | 4 | 0.0000 |
| | Kurtosis | 41.71 | 1 | 0.0000 |
| | Total | 2188.58 | 19 | 0.0000 |

Note: 'The information-matrix test is a conditional moments test with second-, third-, and fourth-order moment conditions'; whilst the heteroskedasticity test is very similar to the general tests for heteroskedasticity that White (1980) proposed (STATA, R. 8, vol.3 N-R: 363).

### 3.10.5 Results

Tables 3.10.4A – G are the results of the robust regression. The F-statistics gives the overall significance of the model. It tests the hypothesis that all the independent variables have no effect on reserve margin; that is, the hypothesis of the slope coefficients (excluding the intercept) in the regression being jointly equal to zero. Its *p*-value shows the probability of obtaining an *F* as large as or more than the calculated figure. The table shows that the independent variables are jointly statistically significant at the 5% level.

As expected, SMP and CP have negative coefficients; this confirms that an increase in capacity causes prices to decrease. Surprisingly, on average, Uplift is statistically significant but it had quarters in 1994 and 1995 when it was statistically insignificant. These results imply that a market with a similar capacity mix as England and Wales' and in which the price rules are the same as that used in the pool, may have periods when balancing costs will not reflect the excess capacity on the network. Unplanned outages or simply the Generators capacity manipulations can cause that. The idea of capacity manipulations leading to higher prices is consistent with some of the reports in earlier studies on the pool data (see for example, Fehr and Harbord, 1993; Green, 1994; Wolfram, 1998; Wolfram, 1999; Wolak and Patrick, 2001). This is also a familiar argument that the *Director General of Electricity* (DGES) made in investigations into causes of high prices in the pool (see for example, OFFER, 1991; OFFER, 1994).

## Table 3.10.4 Regression results of Reserve margin and SMP, CP, Uplift and Declared Availability

| A | Dependent variable = Reserve margin | | | | | |
|---|---|---|---|---|---|---|
| Year | | All | Q1 | Q2 | Q3 | Q4 |
| 1994 | N | 17520 | 4318 | 4368 | 4416 | 4418 |
| | SMP | -4.70 | -2.63 | -4.12 | -4.77 | -4.37 |
| | | [-19.14] | [-7.93] | [-11.42] | [13.23] | [-14.12] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | CP | -1.88 | -4.35 | -5.48 | -4.26 | -2.70 |
| | | [-6.91] | [-6.32] | [10.09] | [-6.61] | [10.94] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Uplift | **0.14** | -1.53 | -0.43 | **-0.06** | 0.75 |
| | | [0.91] | [-6.99] | [-2.92] | [-0.58] | [6.56] |
| | | (0.369) | (0.000) | (0.005) | (0.568) | (0.000) |
| | Declared | 0.29 | **-0.05** | 0.22 | -1.13 | 0.49 |
| | | [19.42] | [1.45] | 5.75 | [-2.69] | [14.04] |
| | | (0.000) | (0.154) | (0.000) | (0.010) | (0.000) |
| | $R^2$ | 0.42 | 0.41 | 0.46 | 0.60 | 0.56 |
| | F-statistic | 231.15 | 237.14 | 208.28 | 217.11 | 139.79 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

| B | Dependent variable = Reserve margin | | | | | |
|---|---|---|---|---|---|---|
| Year | | All | Q1 | Q2 | Q3 | Q4 |
| 1995 | N | 17520 | 4318 | 4368 | 4416 | 4418 |
| | SMP | -2.74 | -0.89 | -2.27 | -5.42 | -6.40 |
| | | [11.32] | [-8.08] | [-3.83] | [13.21] | [-19.34] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | CP | -3.71 | -2.87 | -4.95 | -4.20 | -1.82 |
| | | [11.33] | [16.20] | [-11.94] | [-7.71] | [-4.78] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Uplift | 0.99 | 0.63 | **-0.24** | -2.07 | 0.41 |
| | | [8.85] | [10.44] | [0.92] | [7.47] | [3.30] |
| | | (0.000) | (0.000) | (0.362) | (0.000) | (0.002) |
| | Declared | 0.30 | 0.58 | 0.30 | 0.58 | 0.86 |
| | | [8.34] | [10.99] | [9.88] | [15.62] | [21.39] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | $R^2$ | 0.28 | 0.29 | 0.37 | 0.63 | 0.53 |
| | F-statistic | 74.92 | 146.23 | 103.62 | 136.13 | 211.64 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

Notes: 1. All = all the observations in the year. N = number of observations within the year and / or quarter that is used in the regression. Q1 = January to March; Q2 = April to June; Q3 = July to September and Q4 = November to December. SMP = System Marginal Price and CP = Capacity Payment. 2. Used White heteroskedastic consistent standard errors to calculate the t-statistics in square parenthesis. 3. Statistically insignificant coefficients at the 5% significance level are boldly highlighted.

## Table 3.10.4 Regression results of Reserve margin and SMP, CP, Uplift and Declared Availability (cont.)

| C | Dependent variable = Reserve margin | | | | | |
|---|---|---|---|---|---|---|
| Year | | All | Q1 | Q2 | Q3 | Q4 |
| 1996 | N | 17568 | 4366 | 4368 | 4416 | 4418 |
| | SMP | -3.76 | -4.44 | -4.07 | -4.77 | -4.13 |
| | | [-12.72] | [-17.79] | [13.53] | [13.57] | [-12.23] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | CP | -3.87 | -5.10 | -2.78 | -3.03 | -2.63 |
| | | [-17.07] | [-21.84] | [-7.13] | [-10.46] | [-9.20] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Uplift | 1.16 | 1.87 | **-0.10** | 0.57 | 0.60 |
| | | [12.41] | [12.82] | [-0.28] | [3.07] | [7.32] |
| | | (0.000) | (0.000) | (0.779) | (0.004) | (0.000) |
| | Declared | 0.27 | 1.01 | 0.11 | 0.24 | 0.44 |
| | | [12.93] | [25.60] | [3.36] | [9.03] | [19.28] |
| | | (0.000) | (0.000) | (0.002) | (0.000) | (0.000) |
| | $R^2$ | 0.41 | 0.50 | 0.55 | 0.61 | 0.46 |
| | F-statistic | 202.00 | 356.83 | 146.49 | 146.49 | 179.51 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

| D | Dependent variable = Reserve margin | | | | | |
|---|---|---|---|---|---|---|
| Year | | All | Q1 | Q2 | Q3 | Q4 |
| 1997 | N | 17520 | 4318 | 4368 | 4416 | 4418 |
| | SMP | -3.92 | -3.89 | -3.60 | -5.43 | -3.41 |
| | | [-12.20] | [-16.38] | [-7.44] | [-9.03] | [-8.47] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | CP | -5.14 | -3.87 | -15.13 | -9.96 | -4.83 |
| | | [-8.18] | [8.86] | [-16.35] | [-10.98] | [-8.68] |
| | | (0.000 | (0.000) | (0.000) | (0.000) | (0.000) |
| | Uplift | 1.49 | 0.88 | 2.84 | -3.95 | 1.44 |
| | | [7.22] | [6.36] | [3.06] | [-5.43] | [7.15] |
| | | (0.000) | (0.000) | (0.004) | (0.000) | (0.000) |
| | Declared | 0.52 | 0.42 | 0.52 | 0.41 | 0.55 |
| | | [21.37] | [13.94] | [14.31] | [8.81] | [11.21] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | $R^2$ | 0.46 | 0.52 | 0.49 | 0.48 | 0.50 |
| | F-statistic | 172.34 | 180.84 | 468.46 | 56.90 | 130.61 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

Notes: 1. All = all the observations in the year. N = number of observations within the year and / or quarter that is used in the regression. Q1 = January to March; Q2 = April to June; Q3 = July to September and Q4 = November to December. SMP = System Marginal Price and CP = Capacity Payment. 2. Used White heteroskedastic consistent standard errors to calculate the t-statistics in square parenthesis. 3. Statistically insignificant coefficients at the 5% significance level are boldly highlighted.

## Table 3.10.4 Regression results of Reserve margin and SMP, CP, Uplift and Declared Availability (cont.)

| E | Dependent variable = Reserve margin | | | | | |
|---|---|---|---|---|---|---|
| Year | | All | Q1 | Q2 | Q3 | Q4 |
| 1998 | N | 17520 | 4318 | 4368 | 4416 | 4418 |
| | SMP | -2.77 | -2.19 | -2.91 | -4.33 | -2.66 |
| | | [-9.96] | [-5.59] | [-7.56] | [-8.77] | [-8.48] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | CP | -7.82 | -7.44 | -7.08 | -7.86 | -6.26 |
| | | [-15.71] | [-14.90] | [-19.38] | [-17.08] | [-8.22] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Uplift | 3.36 | 3.64 | 3.33 | 2.86 | 2.22 |
| | | [10.34] | [14.77] | [14.70] | [6.75] | [5.95] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Declared | 0.61 | 0.33 | 0.54 | 0.39 | 0.71 |
| | | [26.32] | [5.40] | [17.59] | [10.47] | [27.98] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | $R^2$ | 0.53 | 0.39 | 0.48 | 0.42 | 0.49 |
| | F-statistic | 425.87 | 129.72 | 188.85 | 151.31 | 292.12 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

| F | Dependent variable = Reserve margin | | | | | |
|---|---|---|---|---|---|---|
| Year | | All | Q1 | Q2 | Q3 | Q4 |
| 1999 | N | 17520 | 4318 | 4368 | 4416 | 4418 |
| | SMP | -2.86 | -2.35 | -5.18 | -5.35 | -6.33 |
| | | [-10.76] | [-9.00] | [-10.83] | [-16.59] | [-12.43] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | CP | -6.22 | -6.74 | -2.75 | -4.88 | -3.41 |
| | | [-25.64] | [-10.89] | [-6.69] | [-21.88] | [-6.02] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Uplift | 2.23 | 2.50 | -2.84 | 1.58 | 1.48 |
| | | [19.73] | [9.27] | [-5.44] | [19.69] | [5.65] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Declared | 0.48 | 0.57 | 0.17 | 0.72 | 0.44 |
| | | [22.95] | [12.81] | [6.77] | [20.22] | [14.50] |
| | | (0.000) | (0.000 | (0.000) | (0.000) | (0.000) |
| | $R^2$ | 0.52 | 0.44 | 0.53 | 0.60 | 0.52 |
| | F-statistic | 394.18 | 206.93 | 157.33 | 202.85 | 114.13 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

Notes: 1. All = all the observations in the year. N = number of observations within the year and / or quarter that is used in the regression. Q1 = January to March; Q2 = April to June; Q3 = July to September and Q4 = November to December. SMP = System Marginal Price and CP = Capacity Payment. 2. Used White heteroskedastic consistent standard errors to calculate the t-statistics in square parenthesis. 3. Statistically insignificant coefficients at the 5% significance level are boldly highlighted.

**Table 3.10.4 Regression results of Reserve margin and SMP, CP, Uplift and Declared Availability (cont.)**

| G | Dependent variable = Reserve margin | | | | | |
|---|---|---|---|---|---|---|
| Year | | All | Q1 | Q2 | Q3 | Q4 |
| 2000 | N | 17568 | 4366 | 4368 | 4416 | 4418 |
| | SMP | -5.31 | -5.07 | -5.28 | -5.27 | -5.85 |
| | | [-19.61] | [-16.03] | [-9.84] | [-11.74] | [-15.69] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | CP | -3.05 | -3.95 | -3.06 | -3.29 | -2.15 |
| | | [-14.70] | [-6.93] | [-9.18] | [-17.00] | [-7.71] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Uplift | 0.93 | 1.48 | 0.93 | 1.06 | 0.56 |
| | | [11.86] | [6.04] | [6.67] | [14.48] | [7.33] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | Declared | 0.46 | 0.50 | 0.39 | 0.39 | 0.53 |
| | | [32.76] | [12.44] | [20.65] | [12.67] | [24.58] |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | $R^2$ | 0.61 | 0.60 | 0.56 | 0.56 | 0.54 |
| | F-statistic | 799.65 | 164.99 | 201.46 | 98.48 | 273.16 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

Notes: 1. All = all the observations in the year. N = number of observations within the year and / or quarter that is used in the regression. Q1 = January to March; Q2 = April to June; Q3 = July to September and Q4 = November to December. SMP = System Marginal Price and CP = Capacity Payment. 2. Used White heteroskedastic consistent standard errors to calculate the t-statistics in square parenthesis. 3. Statistically insignificant coefficients at the 5% significance level are boldly highlighted.

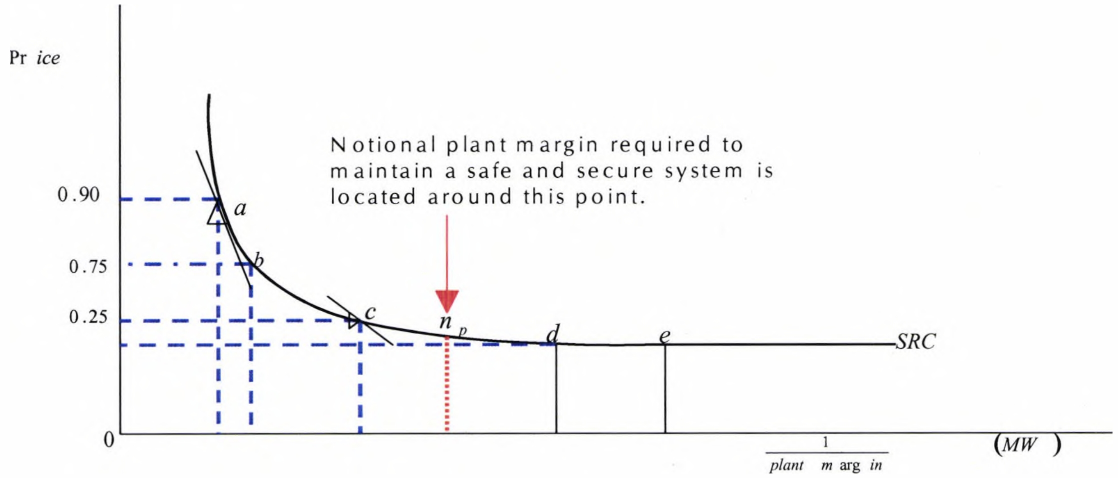
### *3.10.6 Discussion*

The result of SMP and CP having negative coefficients and being statistically significant across the full sample is consistent with the proposition for the study. This suggests that a marginal increment in capacity will increase reserve margin and decrease SMP and CP. But it is important to note that the real magnitude of the effect of reserve margin on prices will depend on the ease with which the Generators exercise monopoly power.

Based on this theory, the effect of reserve margin ought to be highest during the peak than at the off-peak periods. This makes sense if we consider that the differential of the SRC in figure on 3.10.2 is higher at point a, which corresponds to the period of low reserve margins than at point c, where reserve margin is higher.

**Figure 3.10.2**

**Plant margin and prices**




This is consistent with the economic theory of scarcity and price determination in competitive markets (see Sloman, 2002); it also supports the decision to conduct quarterly analysis.

But the results show enormous variability in the magnitude of the change between the quarters. Also there is no consistent pattern in the effect, as in some years, a change in reserve margin causes a higher effect on prices during the third and fourth and not the first quarter in the year. Here again, this suggests capacity manipulation.

Contrary to our expectation, Uplift exhibits a dual effect; on average, it has positive coefficients, but it is negative in some quarters. The positive coefficients suggest that Uplift increases even with increases in reserve margin and the negative coefficients show that it is possible for increases in reserve margin to decrease Uplift. The coefficients are statistically insignificant in Q3, 1994, Q2, 1995 and Q2 in 1996; this suggests that reserve margin could have had no effect on Uplift in those quarters.

There are possible explanations for the dual effect, which Uplift exhibits. Based on economic theory of price determination in free markets, we expect Uplift to decrease (increase) with increases (decreases) in capacity and the thermal efficiency plus operational flexibility of plants. But the location of the constraint boundaries, combined with the actual position of the system on its security reliability curve, determines the real effect of a marginal change in reserve margin on Uplift. The owners of plants, which are located behind constraint boundaries, are more likely to offer the capacities from such stations at higher prices. Such costs will increase Uplift charges. The DGES found that National Power (NP) consistently offered its Fawley units, which were located behind a long-term transmission constraint, at prices between £45 and £80/MWh (see OFFER, 1991: 44). Therefore, it is possible that such factors might have meant that Uplift costs were higher than competitive rates.

As expected, in 1994, there is a low responsiveness of reserve margin on each independent variable when the other is held constant in 1995. This implies that the low reserve margin did not really cause high variability in prices. On average, the lowest effect of reserve margin on prices also occurs in the first quarter of 1995, which suggests that there was a low responsiveness on prices to the sudden outage of a nuclear plant in January 1995.

CP has the highest responsiveness to reserve margin and SMP has the lowest. This result reinforces the earlier report in the data exploration in part 3.6, that SMP was the most stable, whilst CP the most volatile component of the PSP. This is consistent with the results in section 3.6 as well as in the earlier study by Wolak and Patrick (2001).

On average, declared availability has positive coefficients. The only negative coefficients occur in Q1 and Q3 in 1994 but are statistically insignificant only in Q1. This implies that reserve margin could have had no effect on declared availability during that quarter. Nonetheless, the annual coefficient for 1994 is positive and statistically significant. Therefore, despite the effect in the first quarter, on an annual basis, a marginal increase in declared availability increases reserve margin. As expected, there is a quarterly effect in the relationship between reserve margin and declared availability, which suggests that declared availability exhibits weather seasonality. This also provides further support that the need for power for heating or cooling purposes drives demands between the quarters in the year. This is consistent with Granger et. al (1979); they find significant increases in the volume of electricity that the households consume during the winter and peak day time periods than over the summer and off-peak periods.

*Summary*

The level of reserve margin is not all that may be required to reduce within day balancing costs, particularly if Generators' exercise market power. The next sub-section concludes this part of the empirical data analysis.

### 3.10.7 Conclusion

This paper provides the first investigation of the relationship between reserve margin s, SMP, CP, Uplift and declared availability in the England and Wales' pool regime.

First, provided that the other inter-system transmission shocks that may affect prices are held constant, one may expect to find a negative relationship between reserve margin and prices; and subject to demand and investment in capacity, to find a positive relationship between reserve margin and declared availability. Second, the effects of reserve margin to the components of the pool selling price, as well as the declared availability, exhibit time and weather seasonality.

The surprising result is Uplift having positive coefficients and being statistically insignificant in some quarters. This is an interesting result for public policy on how best to curtail balancing costs, because the wrong policies will mean that final consumers will bear the costs of Generators inefficiencies. Therefore, this paper lays the foundation to consider forward looking timely initiatives that may be used to relieve temporary and long term constraints on electricity networks.

# 3.11

## Policy Implications

This sub-section asks what a country that adopts the England and Wales' model can do to ensure that its customers pay the *'right'* price for electricity demand.

The five quantitative parts in this section examined the patterns in the prices setting variables in the pool between January 1994 and December 2000. In theory, market efficiency requires that the changes in the industry variables be incorporated instantaneously into stock prices and for price mechanism to allocate resources efficiently. This assumes a perfectly competitive market. In electricity generation, which is an imperfect market, it means that inefficiency might be minimised if Generators do not exercise monopoly power. The empirical results from the five quantitative analyses show that the pool pricing mechanism was inefficient. I conjectured that the loopholes in the rules for setting the system marginal price (SMP), the definition of plant availability in the Grid Code, the allowed components of the offer that the Generators made and the market structure enhanced anti-competitive practises.

Section 2 finds that the Regulators 'invisible hand' influenced the Generators' behaviour; consequently, it restrained higher prices to what was observed in the pool. The regulatory input that went into the reform and the resulting market outcome confirms that public policies are necessary for competition policy to succeed in electricity markets. These polices include but are not limited to the Regulators close surveillance of market operations, the increased use of competition law to set codes of conduct regarding price fixing, vertical and horizontal mergers and other aspects of fair-trading, which would specify acceptable agreements between undertakings (see Kora, 1997; Wish, 1993). Apart from these, it appears to me that getting the rule for setting the prices right is also an important driver of the outcome in markets (Green, 1998b). The incentives given to the SO to balance the system at the least possible cost; the distribution of the marginal plants and the policies that are used to encourage the Generators to balance their injections with off-takes, are also integral parts of the price rule.

*3.11.2 Price Rule*

The data exploration in part 3.6 showed that capacity payment was the most volatile component of the pool selling price (PSP). The functional form for capacity payment: $CP = LOLP(VOLL - SMP)$, made the loss of load probability an important element in the calculation of CP. In earlier research, Bunn & Larsen (1992) found a non-linear relationship between LOLP and electricity prices; and suggested that the Generators would use capacity to manipulate LOLP. The consequence, which they predicted, was that the CP would be high and not reflect the true capacity position in the network. Part 3.8 reveals that about 192% in annual growth in LOLP contributed to roughly 195% growth in CP at the end of week 52 in December 2000. This estimate was based on the sub-sample from January 1998week.1 to December 2000.week52.

The developments in the industry did not justify the abnormal growth rates in LOLP and CP, which I report in the paragraph above. The independent power producers (IPPs) had made significant investment in plants, most of that had come on-line. The excess capacity had contributed to over 35% in the annual average reserve margins between January 1994 and December 2000. Following the increase in the Generators registered capacity (GRC); the aggregate weekly average of declared capacity grew at an annual rate of approximately 0.90% between January 1998.week1 and December 2000week.52. Approximately 18,702GW of gas-fired plant came on line during the 1990s; as a result more thermally efficient and operationally flexible plants were also used to meet demand. By the end of 1999, National Power (NP) and PowerGen (PGs) had divested additional portions of their mid-merit capacity. The divestment increased the number of Generators that owned the marginal plants; and meant increased competition between the Generators to set the SMP. The Grid had transient constraints, however they disappeared as the system changed within the day (CC, 2001). These are consistent with the intuition that Bunn and Larsen (1992) provided, that capacity manipulations was the only way by which LOLP could have grown at the observed rate, after 1998.

In part 3.6, I find that after 1998, the average value of Uplift followed a constantly increasing rate. This implies that during the period, NGC incurred more costs to resolve constraints, for start-ups and to procure transmission services, all of which it used to keep the system within its tolerance limits. Finally, the study finds a structural break in the path of the level average SMP in week 14 in 1999. This suggests that the Generators might have changed their commercial strategies, presumably they began to offer capacities at rates that might have been closer to their true costs. Given the system situation that I have already mentioned in the last paragraph, the patterns of the SMP, CP and Uplift, particularly after 1998, confirm that the Generators used capacity: supply function (see Green & Newberry, 1992) to control these price variables. By doing that, they controlled and sustained PPP and PSP above competitive rates throughout the life of the pool.

Most of the earlier empirical studies on the pool found evidence that Generators used capacity to manipulate prices (Fehr and Harbord, 1993; Green, 1994; Wolfram, 1998; Wolak and Patrick, 2001). The extended results in this section, which provide a complete picture of the trend of the prices throughout the regime, confirms that capacity and price manipulations were an inherent feature in the England and Wales' pool. The Generators using capacity to manipulate pool prices are consistent with the theory of strategic behaviour of agents in the oligopoly market, which Tirole (1998) exposits. Kreps and Schienkman (1983); Schienkman and Brock (1985) and Friedman (1971) provide supporting theories and arguments that intuitively explains how the repeated and interactive nature of the commodity auctions in electricity generation lead to socially inefficient outcomes some of the time. This is the basis for the aggregate costs of generation in electricity markets being above competitive levels (see Fehr & Harbord, 1993).

The theories of natural monopoly and competition confirm that market prices should decrease (increase) with increases (decreases) in the quantity demanded or supplied (see Sloman, 2000). Since all the segments in the electricity industry are sub-additive, they exhibit increasing returns to scale; therefore, in theory, price decreases should follow increases in quantity supplied, this makes sense because the average cost (AC) of production will decrease with increases in the number of customers / geographic density. The slight downside to the increasing returns to scale in generation and which is very relevant in competitive regimes is that the growth in one firms' share of the market (where demand is inelastic) can cause other competitors to leave the industry. This is because subject to a firms plant portfolio, those with smaller numbers of customers will have higher ACs. Although the desire to be in-merit will often limit the economic objectives of Generators, profit maximisation will mean that the owners of the marginal plants will make higher offers for their residual and inframarginal capacity.

In practise and like part 3.6 shows, electricity demand is time and weather dependent. Also higher cost is used to meet load as demand increases along the load duration curve (LDC). Once in the market, the merit order methodology in the England and Wales' pool, meant that the most expensive and usually the least efficient plant set the SMP; but the full capacity of the least cost plant was fully dispatched first (see Fehr and Harbord, 1993). The ideas of economies of scale and the number of Generators that might enter the market were irrelevant when setting the SMP. This explains why in part 3.6, I found that prices increased in an ascending order of magnitude from the base to peak load regime. The theory of economies of scope / production does not quite offer an intuition into how high prices can go in electricity markets, even when there is excessive network capacity. On the other hand, the performance of electricity markets is not solely determined on the interaction of demand and supply or the system security position. This means that the policies that influence the emergence of prices require an approach that is different from the theory of competition, which in this case and as I believe, has to be separated from the expectation that a price mechanism will maximise social welfare.

Given the evidence of inherent inefficiency in the pool; an emerging market that adopts the England and Wales' model and prices setting may improve social welfare by using rules that are sufficiently robust that it restrains the Generators from systematically manipulating capacity to earn higher rents.

### 3.11.2.1 Capacity payment

*Flexible, robust and forward-looking rule.* Capacity payment should be a part of the price rule in electricity markets; and the emphasis in the regulatory reform should be directed at having the *'right'* rule and implementation methodology that can encourage the Grid users to make appropriate levels of network investment. How will CP be structured? It should be robust, flexible and forward-looking. I am assuming by saying this that the objective will be to see that prices reduce as network capacity increases. In that case, an efficient regulation that approximates price mechanism can achieve this. Consistent with the invisible hand, if CP rule is flexible, then the Regulator can effect its timely modification to reflect changes in the network capacity mix as the industry evolves.

CP can be a part of the price rule right from the inception of a reform; or provided that its rule is determined as part of the regulatory reform, it can be set up later. Before vesting there should be a reliable assessment of the system capacity. This should include the status of the available plants, the nature of the Grid (possible constraint boundaries) the potential for demand growth and the ease of entry of IPPs. I assume the adoption of the *planning margin*[29] used in the UK. I do not specify a benchmark because the energy balance on a system, nature of the Grid, size, capacity mix, will determine its acceptable plant margin (see IEA, 2002). I assume that the requirement to maintain a transmission system within its acceptable tolerance limits calls for a notional plant margin within which the SO can plan the need for future generation.

If the plant margin on the system is below say its notional level, but there are high potentials for growth in demand, CP should be paid right from the time the regulatory reform is vested. Paying CP right from the inception of a reform will be a compensatory incentive to the owners of the available plants; it may also encourage entry of IPPs. The forward-looking aspect and by which it can approximate a price mechanism comes from its

[29] The Central Electricity Generating Board (CEGB) sought to achieve a plant margin of 24% several years ahead which it called the *'planning margin* (i.e. the plant margin for planning the need for future generation). An international Review of Adequacy Standards for Generation and Transmission Planning (CIGRE Report No. 37 – 92 (AG) 02 (E), November 1992) shows that large utilities world-wide seek planning margins up to 30%' NGC 2000 p 4 2).

flexibility. A flexible rule would mean that changes could easily be made as the reformed industry develops and the patterns of some of the market variables emerge. For example, the system capacity mix and the load profile. It is also over time that the tools for forecasting marginal changes in demand can be done accurately. The Regulator should modify the price rule as soon as the system attains its acceptable notional planning margin; but have a phased scheme with which to withdraw CP. Phasing out CP will give the Generators the opportunity to plan an adjustment to the income changes.

The arrangement that I propose in the paragraph above suggests that CP can be a transitionary arrangement. The next question is how long should the transition period be. The length of the transition will vary between systems; in particular it will be subject to the size of the network, and will be equal to the time that it takes for the network to attain its notional planning margin. This implies that the transition period is indirectly dependent on the rate of investment in the generation plants as well as in the wires and cables. Because if the focus is placed only on generation and there are no lines to transport power from places of lower cost and excess supply to those of higher cost and excess demand, the resulting local monopolies would charge higher prices for their capacities.

At the assessment stage, and before vesting, if a system has significant capacity relative to the anticipated growth in demand and there are clear signals that environmental changes may induce more entry into the industry, capacity payment should not really be a distinct part of the price rule. This is where I think that England and Wales created the price distortion. At the time that the privatised industry was vested in 1990, the massive entry of IPPs was imminent because the cost of fuel had started to decrease at an increasing rate in the 1980s. It was clear that the 'dash', lower costs of entry for CCGTs and the diversification of the Regional Electricity Companies (RECs) into generation business, would result in massive investment in gas-fired plants. The *'initial portfolio'* (IPs) contracts, which the Government set in place between the Generators and the RECs that was backed against British Coal was set to expire in 1998, at which time full retail competition was initially scheduled to commence. There was also the case for the changes in the gas trading arrangements, which was thought to start in 1998 as well. All of these meant that there were possibilities from the onset, or as the market emerged in the early 1990s that investment patterns would be biased in favour of more thermally efficient and operationally flexible gas plants. The knock-on effect would be reductions in the LOLP; in theory, the earlier studies and simulations on this aspect of the development in the pool inferred that the LOLP would follow a constantly decreasing trend between 1990 and the closure of the pool. See for example, the first predictions by Bunn, Larsen & Vlahos (1993); and the further insight provided by Larsen and Gary (1998).

The issues that I raise above were obvious at the time of vesting because they originated from the initial policies for vesting the industry. If we consider the developments that were not obvious in 1990 but which emerged as the market changed, we will still conclude that LOLP ought to have decreased after 1998. For instance, by the expiration of the IP contracts in 1998, the RECs that had diversified and invested jointly into gas plants, which were located closer to their load and within the LDZs, preferred to meet their demand from their gas plants. The transmission network use of system charges (TNUoS), which was based on long run incremental cost (LRIC) pricing combined with the higher cost of British Coal had made Generators re-locate their plants in the south. The debt convents on CCGTs made then to construct their plants as flat loads[30] had changed and they could operate on a two-shift basis at about 50% of their load factors and move up the load duration curve into mid-merit (CC, 2001).

As long as there is equitable distribution of the marginal plants, and the potential for a rapid entry is identified, initiatives should be directed at securing and promoting competition; this can be done by facilitating non-discriminatory entry into all the portions of the LDC. Once the reform is implemented, the market dynamics, which although directed by public polices and fair trading acts, should be allowed to stimulate and equilibrate entry and exit into the industry. Moreover, it is reasonable to expect that the Generators in the pool would have bid aggressively to be in-merit rather than have idle capacity and earn nothing.

It seems to me that the problem in the England and Wales' pool was that the Regulator did not adjust the *value of loss load* (VOLL) as the industry evolved, newer plants came on line

[30] At the inception of the regime, CCGTs were subjected to long term take-or-pay contracts, which stretched up to fifteen year ahead with, associated off take agreements. They were also required to operate at approximately eighty-percent of their load factors. The only way that they could achieve these targets was to operate as flat loads.

and the inefficient ones were withdrawn or mothballed. In addition to that the 192% annual growth in LOLP between January 1998.week1 and December 2000 week52, suggests that the way in which the *loss of load probability* (LOLP) was calculated was not modified to accurately reflect the increased levels of the thermal efficiency and operational flexibility of the plants that were being used to meet demand.

*Interaction of energy market with the finance sector.* One lesson from the England and Wales' pool is that the debt covenants which the Generators had at the onset of the regime and the merit order dispatch methodology influenced entry of plants into segments of the LDC. This confirms that the sectoral lending policies in the finance industry for the energy market participants will affect investment decisions, and the development of the capacity mix as well as the emergence of prices. The indirect effect will be on the length of time that it will take for the system to attain its notional plant margin limit. This is where the Government can use public policy to influence the speed at which the capacity mix develops and to ensure an equitable investment across all segments of the LDC. I assume by saying this that improvement of social welfare underpins energy polices. Therefore, Government should work closely with the Central Banks to ensure that the sectoral lending polices of the investment banks are favourable to energy market participants.

*Equitable investment along the load duration curve (LDC).* From an economic point of view, IPPs will avoid the sections of the LDC where plants would run for a small number of times in the year. In the UK's case, if there was a serious inter-system transmission shock or *'act of god'* that caused a significant loss of baseload capacity, the security of the system might have been jeopardised (see also the illustration in 3.6).

The energy balance on a network will determine the severity of the potential threats to the system operations that arises if there is a high investment bias across the load regime. For example, consider a hypothetical system that is predominantly hydro-based. Hydro plants in this system are guaranteed to run most of the time in the year. If I assume that the IPPs make little investment into thermal plants and that severe draught occurs, which dries up water resources; the effect will be seen in the plant margin, security of the transmission system, reliability of supply and the price levels. Plant margins will reduce, if it is low and at such levels that price can take infinite values, the thermal Generators will offer their capacities at abnormal rates. This was the case in Norwegian pool, where the lack of rain meant that the system cold not cope with the winter demand during the last quarter of 2002. There were abnormal and very volatile prices in the pool, with some of the retail customers paying an average of approximately $2,745 for their electricity demand for the three months (Time, 2003). In an extreme situation the number of threats to the system operations will increase. Supply may not be guaranteed; it may be that this system will not have any opportunity to import significant quantities of electricity to meet its aggregate demand. If this happens, equitable distribution may be achieved only by load shedding.

Capacity payment in this hypothetical hydro system may be load regime based. The supporting policy initiatives should be directed at encouraging IPPs to construct thermal plants. Since the thermal plants may be called on a few times in the year, it may be economic to implement a financial scheme that pays these Generators' a flat annual fee right from the inception of a competition regime. That is, the owners of the thermal plants should receive financial compensation whether the plants are in or out of production. Consistent with a two-part pricing scheme, which is common in network utilities (see Laffont and Tirole, 1993; Armstrong et al, 1994) this payment can be designed to have a fixed and a variable element. The Generators can be paid the fixed element, which would reflect the long run costs all year round. The fixed element can be calculated as a proportion of the Generators' highest expected dispatch under abnormal conditions and subject to its registered capacity (GRC). The variable portion will be a linear function of the actual MW of power that the thermal generator injects onto the Grid at any time in the year. That is, it will reflect the short run avoidable costs; and is dependent on actual production.

### *3.11.3 Ownership of marginal plants*

The univariate analysis of the unobserved properties of the SMP shows that the level average value had a permanent decrease in week 14 in 1999. One lesson about the regulatory oversight in the pool was that the Regulator used price controls and daily surveillance of market operations, to restrain prices. There was also a price influence from the potential that the Regulator could refer any generator that he perceived to inhibit the development of competition to the Competition Commission (CC). Green (1999) records that the Regulators' reliance on strongly worded communications to the Generators also helped to keep their offers low. The impact of the regulatory oversight on price recorded in the UK is consistent with what happened in Hong Kong (see Lam, 1999). This calls for a dedicated team in the Regulators' office who would monitor the daily market operations.

In part 3.6 I gained an insight into the type of capacity manipulation that the Generators might have used to earn higher SMPs. It was in 1991 that the Regulator found that the Generators were likely to use the loopholes in the Grid Codes definition of an operating plant and the components of their offer prices, to manipulate SMP and the table indicated half-hours (see OFFER, 1991). This immediately tells me that it was clear, within a few months of the pool starting, that it would be difficult to achieve efficient competition within the duopoly structure (see also Fehr and Harbord, 1993). In part 3.8 I gained an idea of the effect industry structure and anti-competitive conduct has on prices, and which I conjecture would equate to a high social cost. The analysis in that part of the thesis also shows that the demand and supply as well as the system security position did not justify the spikes. However, in part 3.9, I find results that suggest that there might have been other exogenous factors that determined the time of day SMP. I conjectured again that the Generators opportunistic strategies were the main determinants of time-of-day (TOD) SMPs.

The interesting aspect of the spikes, price volatility and the ratio of table A to B ('Stretch'), was their evolutionary emergence, which suggests that they worsened as the market developed. The inference I made from this revelation is that as soon as the Generators learned the rules, they perfected sophisticated strategies with which they manipulated prices. This makes me believe that the trend of prices and manipulation strategies will emerge slowly in any market. Consequently, it might be unwise to use the early days of a reform as a basis for assessing the effectiveness of competition or the success of the reform. This slow emergence of price suggests, consistent with Hogan (2001a[31]) that the regulatory oversight should also be evolutionary. Deregulation does not mean 'no' or 'less' regulation. Instead, it means that as the market evolves, the Regulator would need to use a different approach to deal with the emergence of issues such as refusal to supply, price and capacity manipulations, predatory pricing, cross subsidisation and merger issues amongst other

[31] Hogan is quoting Steven K. Vogel—Freer Markets, More Rules: Regulatory Reform in Advanced and Industrial Countries.

things. These issues are usually difficult to deal with and enforce as markets mature (CC, 2001). England and Wales' regulatory reform did quite well here; the Regulator carried out a series of consultations with the industry and changed the capacity and commodity regimes in line with developments in the industry.

Free market mechanism does not provide any guidance as to the true price level that one can reasonably expect in network utilities, at least not in electricity. At the onset of the pool, prices were initially much lower than the values expected at the design stages; and these low prices were not sustained (OFFER, 1991; MMC, 1996; CC, 2001). Given a capacity level, I am not aware of any empirical study that recommends what prices should be in electricity markets. Throughout this research, I have found that different factors influence the development of competition, the emergence of the pattern of prices and the number of Generators that would exist in the industry as well as their associated GRC. This also applies to the Generators' willingness to declare plants available to run and their decisions to profile availability within the day.

When a uniform price rule is used and plants are scheduled to run in merit-order, the important issue is the 'ownership' of the marginal plant. The initial policies for vesting in England and Wales did not handle this properly; and the creation of a duopoly non-baseload market enhanced market power. The Regulator delayed carrying out a timely divestment of National Power and PowerGen's mid-merit plants, after he identified in December 19991 (OFFER, 1991) that the market structure might affect the development of a competitive regimes. It seems to me that this delay (which I doubt if it was induced by 'regulatory capture' from the incumbents rather than the public interest) influenced the emergence of price patterns. Prices might have been lower than the observed if the number of the owners of the marginal plants were increased earlier than in 1996 and 1999.

The earlier studies on the pool data such as Bower (2002) and Evans and Green (2003) attribute the lower prices after 1998, to increased competition. This suggests that there will always be a positive price effect if the number of Generators that own the marginal plants is increased. This means that the distribution of the marginal plants, indeed the entire system capacity mix should be an important consideration when designing regulatory reforms. Subject to the available network capacity, polices should be directed at restricting collusion.

It appears to me that the way in which Britain could have created a more competitive market in 1990 was if the Government had decided to divide the non-baseload capacity between at least five equally-sized Generators (see Rudkevich, 1997). Tacit collusion could have been constrained between these five firms because their non co-operative equilibrium might have delivered significantly different outcomes from the duopoly structure. If that had been done, it might have been possible that from 1990 SMPs that were closer to the Generators one-year avoidable costs might have been sustained in the pool.

*Bilateral contracts.* The good aspect of the reform in the UK though, was the inclusion of the bilateral contract market, which complemented the pool. One lesson for emerging markets is to ensure that a bilateral contract market is integrated into the industry reform (see Bolle, 1992; Newberry, 1995; Green, 1994). In addition, there are incentives that would be put in place right from vesting to encourage the Generators to lock in a significant proportion of their physical positions into long-term contracts. If the contract market emerges as the sub-segment that facilitates effective threats of entry, it might help to restrain high spot market prices. Over all since Generators would always control prices in both the spot and bilateral contract markets, the latter will help to keep prices and output levels stable. The prices volatility that parts 3.6 and 3.7 reports in pool prices and the rate of spikes that part 3.8 reveals in the SMP could have been worse if Generators did not lock over 80% of their aggregate demand into long-term contracts. Put another way, prices could have been much higher if more than 90% of the aggregate power consumption were purchased at the pool price.

### *3.11.3 Government participation in the ownership of plants*

The quantitative analysis of plant margin, SMP, Uplift and demand reveals two things. Firstly, increases in plant margin might decrease prices, provided that the Generators behave normally. This is consistent with the economics theory of scarcity, which provides an idea of the direction of the price effect of capacity shortages, but not an indication of the magnitude that would have an adverse effect on social welfare. The second thing, which though surprising was the unresponsiveness and statistical insignificance of Uplift to plant margin in some quarters in the estimation, carried out in part 3.10. Again, the result confirmed that there might have been other factors than plant margin that determined Uplift during these quarters and that the effect of plant margin was statistically insignificant. It seems logical to expect that balancing costs should reduce with increases in plant margin. In reality the potential that technical problems can lead to plant outages and / or for inter system transmission shocks to cause transportation constraints, means that Uplift cost may not always be responsive to plant margin. In order words, as long as the reserve margin on a system is not 'zero', plant outages and transmission constraints can occur even when there is excess capacity on the Grid.

The constantly increasing rate of Uplift after 1998 suggests increases in the threats to the system operations during that period. Since public policy and not price mechanism determines the conduct of Generators and the performance in the industry, it seems to me that a different type of regulatory approach is necessary to control capacity manipulations within the day; and which if achieved, will reduce balancing costs.

The Government uses public policy to curtail windfall gains and excessive prices, which collusion and pre-emptive practises makes worse in electricity markets. None of these problems relate in any way to the capacity mix or to the economies of scale and scope. That is, they do not affect firms' AC of production. The increases in the number of players that own the marginal plants or the increased thermal efficiency of the plants used to meet demand do not also mean that the costs of generating power will reduce. Or that the Generators will change their commercial strategies to reflect comparable efficiency savings that they make. It also does not mean that transmission shocks will disappear and that within the day balancing costs will reduce. Apart from having an equitable distribution of the capacity mix on the network, it seems to me that the Governments direct participation is another way to strengthen initiatives to curtail capacity manipulations *ex-ante* if there are prices that are set before real time. This will also be the same *ex-post* if the system security costs forms part of the price that the demand side pays to purchase electricity from the pool.

One of the reasons for deregulation is that public firms are inefficient; therefore, competition policy is expected to lead to the right value discovery for commodities. The five quantitative parts in section 3 have shown repeatedly that the pool price rule, in a privatised regime, was inefficient. Although the direction and / or causes of the inefficiency between the pool and what might have happened post-deregulation, for instance under the

Central Electricity Generating Board (CEGB) might differ, in my view, this is evidence that the ownership of electricity firms does not matter. This is because there can be market inefficiency both in the private and public regimes. Based on this, I think that the important thing is that optimising social welfare should just mean that any firm that can provide electricity at the cheapest possible price should be allowed to do so whether it is a private or a public firm. My argument is based on the theory that price in oligopoly or monopoly markets can never be at competitive levels but will be below monopoly rates (see Tirole, 1998). Judging that in practise, one firm can produce at a cheaper rate than another and that the Government will continue to spend significant amounts of money regulating the industry, it seems sensible that in the competitive regime, a regulated public generation can be allowed to own a portfolio of plants too. This should be created right at the time of vesting.

The regulated public firm can charge a price that is below the monopoly rates. This logic means that the desired social equilibrium under competitive regimes can be achieved within the combined private and regulated public monopoly. This type of regime will give the Government the opportunity to continue its dual role as a paternal director of the market operations; and still participate directly in promoting social action initiatives.

The regulated public utility should compete with IPPs in the wholesale market to set commodity prices and quality standards. The holding company should not be given an open-ended subsidy; instead, its managers' will be given financial incentives to run the company efficiently. They can also be made more accountable to the taxpayers, the shareholders of the firm; their terms of engagement can be set similar to those given to their counterparts in the private firms. So for example, the managers can be made to pursue profit maximisation and be subject to appropriate penalties if they fail to meet the set targets.

The referenced IPP prices can be used to calculate the firm's commodity price. But it should adopt an AC pricing rule, to which a period varying spread (if appropriate) can be added. The firms' pricing may be allowed to increase annually subject to the Retail Price Index (RPI). Since on average, the price of the public firm will not vary with market shocks, the SO can use output from the Government-owned plants to meet demand whenever IPP prices are abnormal and / or when sudden outages occur. The additional benefit of this recommendation for the Governments' direct participation by using publicly owned utilities is that the more efficient private firms may stimulate the public firms to improve productivity.

There can also be incentives to support reductions within the day security costs, which can be put in place to complement the public-private regime. For instance, policy initiatives for timely resolution of network constraints can be instituted with associated schemes of strict compliance on all Generators to submit a periodic plant maintenance schedule, which will be subject to an independent engineering verification. I believe that this will serve as a preventative measure against sudden outage and complement energy balancing and capacity regimes on the NTS.

I anticipate that the policy measures that I prescribe here could help curtail the effect of the Generators use of capacity to manipulate prices.

# SECTION 4

## **Sub-Saharan Africa (SSA):** Towards Regional Power-Pools

*Abstract*

*The growing dimension to have regional power integration arrangements and exchanges in East, West and Central Africa that are similar to the regime in the Southern Africa Development Community (SADC) is closely tied to electricity deregulation. But there is no evidence that these regimes will improve social welfare in sub-Saharan Africa (SSA). I examine the factors that will inhibit the development of contestable capacity and commodity regimes in these regional markets; and use the experience from the regulatory reform in England and Wales to develop a model for the emergence of regional power pools. This is a theoretic and generalised model, based on some assumptions. It can be adapted to suit country specific economic and political situation, as well as the characteristics of a network such as the size, capacity mix, storage installations, energy balance and transmission and distribution capacities. Nonetheless, policy advisors that are interested in enhancing the success of competition policy in the electricity industry in the region can use my model to do so. This section opens an area for further research, which is identified.*

*Key words: Africa, Deregulation, Electricity, sub-Saharan Africa, Regional Power Pools; Southern Africa Development Community (SADC)*

**Introduction**

The objective of this paper is to examine the factors that would inhibit the development of efficient competition in the proposed regional power pools in *sub-Saharan African* (SSA). These issues are particularly important given the renewed pressure from the World Bank and the IMF on these countries to meet their privatisation commitments. It seems that the Heads of Government in East, West and Central Africa consider that the continued existence of the regional power integration regime in the Southern Africa Development Community (SADC) is a sign that it is a success. Therefore, they see regional power pools as a possible way to optimise natural resources and reduce the aggregate cost of providing electricity to their population. They also consider that it might be the best arrangement, for the member states, who can support each other in times of emergency. They anticipate that it would be a long-term solution to the problem of electricity capacity shortages in the region.

*4.1.2 Method of study and literature review*

This is an economic regulation and competition case study, which we use to build a theoretic model. It involves the critical evaluation of the impact of the socio-politics and economics of the SSA region, on the success of its electricity regulatory reforms.

Joskow (1997) uses the same approach to present an economic summary of deregulation of electricity systems in some of the states in the United States of America (USA). Vickers and Yarrow (1991) use it to examine the economics of privatisation between the developed, market-based economies in Western Europe and the *less developed countries* (LDCs). Green (1999) Wolfram (1999b) and Kwoka (1997) use it to evaluate developments in the England and Wales' pool.

The studies that we report in the paragraph above cover the electricity systems in the advanced countries in Europe and the USA. In the African Continent, O'Leary et al (1998) examine the factors that led to the emergence of the regional power-pooling regime in the SADC. These highlight the technical and operational issues that the *Southern Africa Power Pool* (SAPP) needs to facilitate to develop the regional power integration regime. Turkson & Wohlgemuth (2001) examine the possibility of using *distributed generation* (DG) to increase growth and sustainable power supply in SSA. They conclude that DG would be a path towards the flow of foreign investment into the regional electricity industry. Bell (2003) takes the view that there are inherent characteristics of developing countries that make it impossible to replicate the regulatory reforms that appear to work in most of the mature and advanced countries in Europe and the USA, into these economies. As a result, he recommends using options contracts that discourage opportunistic re-negotiations on the part of the investors in regulatory contracts in developing countries. He is of the view that the process for such contracts would provide signals about the true valuation of the profitability of investments before the agreements are negotiated. Sulaiman & Ghebreysus (2001) use probit modelling to investigate the determinants of privatisation between 1970 and 1994 in SSA. On average, their sample results suggest that political pressure and the belief that macroeconomic resource re-allocation might be achieved, are the main drivers for privatisation in the region.

None of the literature that I reviewed has explicitly focused on assessing the impact of the inherent characteristics of the region on the success of competition policy. My contribution to literature is the evaluation of the barrier to the development of the regional power pools. Also, I introduce a model, which these countries can follow to enhance successful transition to competition regimes.

*4.1.3 Overview*

The World Bank estimated that SSA needs approximately $18 billion investment to develop its electricity sector; some of these have either been damaged from years of civil war or vandalised. There are cases where the electricity infrastructures are obsolete; on a general note, the Governments' do not have the funds to finance the network capacity expansion to the levels that will facilitate increases in supply density. Consequently approximately 70% of the population do not have access to electricity supply.

The World Bank expects that the region can raise approximately $5 billion out of this estimate from its utilities and other domestic sources. Other international and multi-lateral organisations can support the project to the value of approximately $3 billion. The region can also look to private sources, presumably foreign investors, to finance the residual $10 billion (Turkson & Wohlgemuth, 2001). These member states are predominately agrarian based economies; they do not earn significant income from the sales of their products abroad; moreover, they have large international debts. The World Banks breakdown of how the region can raise its debt finance suggests that these countries should look to foreign investors to develop their electricity sector. Most of these countries are not active participants in the *International Bond Market* (IBM); consequently, they do not have *sovereign credit ratings* (SCR). This makes it difficult (if not impossible) to see how these countries can have access to the many foreign investors that trade instruments in the IBM. Therefore, we conjecture that the foreign investments that these countries expect from regional power pooling reforms are an illusion.

It is well known that SSA has not made any progress with its electricity privatisation projects since the 1980s. Apart from Mozambique and Zambia, the first set of privatisation projects, which were carried out in SSA between the 1980s up to the end of 1995 failed (see Bennell, 1997). Before that, most of these countries depended on importation to meet their electricity demand. These imports were based on bilateral contracting and an understanding between the member states that they would support each other in times of emergency. At that time, most of the Governments pursued equitable distribution of power supply; they were not interested in the right value discovery for commodity. Therefore, they focused on the expansion of the number of customers that had access to electricity supply; they did this by developing rural electrification schemes, which the World Bank and the IMF supported. Publicly owned and vertically integrated monopolies carried out the entire roles involved in the electricity industry: production through to end-user supplies services. The final consumers did not pay the full cost for their electricity demand because the Treasury subsidised the public monopolies.

In the past 4 decades, we have seen the publicly owned and vertically integrated electricity utilities run down because most of the Governments do not have the finance that is required to meet the daily operations of these corporations. Also, the inability of the managers to curtail the costs of production that they can reasonably curtail means that these utilities persistently produce and distribute electricity inefficiently. The general belief is that the ownership structure of the public utilities is a primary cause of the managers' inefficiency (see for example Crain & Zardkoohi, 1978). In other words, there is a relationship between ownership and the performance of the managers in a firm. There are also studies that find a positive effect of the environment within which a company operates on managers efficiency (see for example Estrin & Perotin, 1991). These studies suggest effective competition induces managers to curtail the operating costs that they can control (Shleifer, 1998). There are also studies that assume that both the public and private firms can be inefficient in a politically unstable environment and in which endemic corruption means that firms spend large sums of money on bribes to the extent that it is a 'fourth factor of production' (Sulaiman & Ghebreysus, 2001).

The World Bank and the IMF adopted deregulation-based lending policies in the 1980's by which they compel the Heads of Governments of the developing countries to deregulate their electricity systems. Stiglitz (2002) reports that these lending policies were not based on the economics of the projects; instead, there was an underlying political motivation. Also, it seems that the perceived success some of these member states claim to have made from deregulation of the telecommunication industry (see for example Shirley, 1992; Gebreab, 2002), gives them a sense of confidence that it can also work in electricity.

The question is whether this type of competition policy is the right way for these countries to achieve sustainable levels of power supply. Given that the path to regional power pools requires transition from vertically integrated monopolies, it raises the question whether unbundling, privatisation and deregulation of electricity networks are 'one size fits all? Is it an initiative that would be more beneficial to some of the participating member states and not to others? If so, is there a path that they can all follow to ensure some level of harmony in the growth and development of the electricity industry within their member states; in particular that competition policy in electricity succeeds?

This section examines some of the issues that the paragraph above presents. To answer them, I summarise in 4.2, the emergence and development of the SADCs power-pooling regime. The outcome suggests that regional power pooling might not improve social welfare of the population in SSA. 4.3 examines the factors that might inhibit private and foreign investment in the deregulated electricity markets in SSA. I investigate these problems from socio-political and economic perspectives. Based on these barriers to foreign investments, which I identify, I conjecture that merely adopting competition policy regimes would do nothing to improve the state of the regional electricity industry. The constraints to foreign investments that I identify are inherent features of the countries in SSA. I believe that there is a need to solve their electricity problem. Therefore, in 4.4, I assume that all the electricity systems of the member states are at the same stage in their development, and use the experience of the licensed-based regulatory reform in England and Wales' electricity, to develop a model, which they can use to work towards regional power pools. This model is not only useful as a path towards the development of regional markets; it can also be a guide for introducing competition regimes within the internal markets of members' states. I accept that there are some factors that may limit the application of our theoretic model in SSA. Therefore, in 4.5, I conclude the study by highlighting some of the initiatives that the member states can put in place now to improve productivity in their electricity sector.

## 4.2 Towards Competition Policy in Southern Africa Development Community (SADC)

The Southern Africa Development Community (SADC) has an abundance of natural resources such as hydro, coal and gas. The Rivers Congo, Zaire and Zambezi are located in the northern part of the region; if the joint hydro resource from the three rivers are put to efficient use, it could provide sufficient power for the sub-region and for export to other parts in the Continent. Coal is the main fuel that is found in the south; there are deposits in the Republic of South Africa (RSA), Botswana and Zimbabwe. There is also some hydro resource in the south on the Orange River in RSA; a small quantity of nuclear-fired generation at the Koeberg Plant, and some natural gas in the Kudu gas fields, which is located in the ocean of Southern Namibia (Stage & Fleermuys, 2001:428).

Despite the large quantity of natural resources that I mention in the last paragraph, the region does not have the necessary generation and transmission infrastructure. Therefore, it has many pockets of capacity shortages that are particularly located between the north and the south, for instance in places like Zimbabwe and Zambia, in central Southern Africa. There are also shortages in Angola; although most of the problems in Angola are as a result of its years of Civil War which damaged its electricity infrastructure.

**Figure 4.1**

**Member States of Southern African Development Community**


Source: www.afrika.no/norsk/Land/S_rlige_Afrika/

The history of co-operation in the regional development, trade and mutual trust between the member states in the SADC started before 1958. It was then that the interconnector that links Nseke in the Democratic Republic of Congo (DRC) and Kitwe, Zambia was constructed to supply the electricity generated in the Congo to the Copper Mines in Zambia. The objective to enhance mutually beneficial development and relationship between the nation states led to the construction of further electricity infrastructures thereafter. The Kariba Dam and the hydro power stations were constructed to provide electricity between Zambia and Zimbabwe. The interconnector between: Ruacana and Windhock was built in 1976 to link the northern hydro power station on the Kunene River and Van Eck power stations; the CahoraBassa and Apollo, in 1976; the Francistown and Gaberone, in 1983; and the Matiba and Bulawayo in 1995 (O'Leary, 1998). These developments encouraged transportation of electricity between the geographic areas; and the ambition to foster harmony in the regional development meant that they needed a way to stimulate a rapid investment into more generation plants, transmission and distribution cables as well as interconnectors.

In 1993, SADC and the World Bank conducted a joint study, which found that the region could save approximately $1.6 billion in its total cost of meeting electricity demand, if it consolidated its regional resource in a power-pooling arrangement. It seems to me that this study plus the additional pressure from the World Bank to liberalise electricity systems motivated the Heads of seven member states to sign up to an *Inter-Governmental Memorandum of Understanding (IGMOU)* in September 1995, which introduced a *loose*[32] *regional* power integration regime. The *Southern Africa Power Pool* (SAPP) an integral part of the project was created shortly thereafter. The regime, which was based on bilateral contracting and not law, was similar to the Scandinavian Nordel / Nord Pool, the Western Europe's UCPTE; and before the restructuring of the US electricity market in 1996, the Mid-

[32] 'Loose pools emphasise the constant exchange of information in order to maximise both the economic and reliability benefits from trading and system autonomy. These pools do not use central despatch of power plants, often relying instead on long-term bilateral contracts for electricity supply between generators and customers. These contracts are supplemented with short-term contracts and other deals under the overall agreement. Loose pools may provide central services, however, including producing continuous, real-time data to match generation and demand, developing indicative expansion plans, and implementing emergency procedures. Loose pools also establish detailed common design and operational standards to ensure system security and reliability and to facilitate trade' (O'Leary et al 1998:2).

Continent Area Power Pool (O'Leary et al, 1998). The co-ordination centre is located at Harare in Zimbabwe; and the regime is governed through a dispute resolution tribunal, energy ministers and officials, a technical administrative unit (TAU) and an executive and management committee. There are also sub-committees: planning, operation and environmental, that support its governance (SAPP, 1999).

Following the implementation of the power-pooling regime, a population of approximately 200 million people that were spread across over 9 million square kilometres were in theory, connected to a unified electricity network. However, the region expected that by introducing the regime, it would attract significant foreign investment with which to develop network capacity. Africa received approximately 0.8% of foreign direct investment (less than 1%) between 1994 and 2003, which is clearly far less than the SADC requires to develop its electricity sector. This lack of finance has limited completion of a number of projects that SAPPs TAU earmarked. These include (1) the rehabilitation of the Inga Hydro Power project, which needs to build a second interconnector that would link the Democratic Republic of Congo (DRC) to the South-Western part of the RSA, but which would pass through Angola and Namibia. Plus the reinforcement of an existing transmission line from Inga through Zambia, Zimbabwe and Botswana to South Africa[33]. (2) There is also a need to upgrade the 210MW interconnector that runs between DRC and Zambia to approximately 500MW and (3) the extension of the Zambia / Tanzania interconnector to Kenya. (4) Finally, they identified that a new interconnector is required, which would run between Mozambique and Malawi. This interconnector should have the capacity of between 50MW and 100MW of electricity (SAPP, 1999).

The lack of interconnectors that I mention in the last paragraph is the reason why countries like Angola, Malawi and Tanzania are not yet connected to the high voltage system, and as a result, they hold observer status in the SAPP. There is clearly excess capacity in the aggregate load; for instance, ESKOM has sufficient capacity with which it can meet the demand in the region if there was network capacity. Also it has not been possible to wheel power easily across the region from the places of excess and relatively low cost generation for example in South Africa, to the places of higher costs and excess demand like in Mozambique. This means that the SADC has not maximised its abundant natural resources in places like the

[33] A 1998 feasibility study conducted by SNEL, ESKOM and the national power utilities in Angola and Namibia estimated that it could cost approximately $461 million to execute this project.

Rivers: Zambezi, Congo and Orange; and the renewable wind energy from places like Walris Bay and Luderits (Stage & Fleermuys, 2000), to generate electricity. The consequence is that despite deregulation and the regional power regime, only approximately 20% of the regional population have access to power supply.

### *4.2.2 Barriers to the development of the power project and SAPP*

The last section summarised the emergence of SADCs power integration arrangement. Next we discuss some of the factors that have limited its development since 1995.

*Dominant monopolist*: South Africa's ESKOM is the net exporter of electricity in the region. It controls 95% of generation and transmission as well as 75% of distribution in its domestic RSA. It also controls approximately 90% of generation in the regional SADC market. In 1992, ESKOM policy was to expand capacity and we observe that it extended its market share into the west and central Africa (ESKOM, 1992). Its planners estimate that even if it does not make additional investment, it has enough capacity to meet growth in demand until after 2010 (Hansen, 1998).

*Accumulated debts*: The agriculture sector employs most of the working population; and on average, the member states do not earn enough foreign exchange from the sales of these agricultural products to meet their international debt commitments. This has resulted in most of the utilities not meeting their debt obligations under the power project. They are also heavily indebted to ESKOM for the power that most of them import onto their Grids. As is well known, these countries cannot honour these outstanding debts in the near future; and they still depend on ESKOM to meet their domestic demand.

There is also the sympathetic case of a long-term power supply agreement that Portugal executed with South Africa in 1969. This contract was for Mozambique to export power to South Africa based on a fixed price agreed in 1969. In our view, this arrangement was possibly based on expected growth and development of the Cahora Bassa Dam. Mozambique generates approximately 2 ½% of its installed 2,075 megawatts at Cahora Bassa Dam, which is not sufficient for it to meet its local demand; consequently, it imports power from ESKOM. The outstanding 1996 contract means that in reality, the two countries have what may be best described as a *'sale and buy back'* arrangement. Although Mozambique owes ESKOM for its imports, the unadjusted costs of the 1969 agreement, means that it purchases electricity at approximately10 times what it sells to South Africa (see Economist 29 March to 4 April 2003). In real terms, Mozambique's debt to ESKOM is overstated.

We observe that ESKOM has adopted a debt for equity swap policy, presumably a way to securitise its exposure to its debtors. For instance apart from its outstanding debt, Zimbabwe's ZESA needs to raise over $200 million to meet its contracted position with ESKOM. Meanwhile, the latter proposed a ownership swap for the formers outstanding debt. This policy dimension will see ESKOM increase its asset base and since it is the more efficient utility, there is hope that, subject to the availability of transmission and distribution lines, electricity supply may increase in the region. The downside though, is that competition will be inhibited if ESKOM becomes a part owner of all the utilities in the region.

*Lack of harmony within the internal markets*: There is no harmony in the development of competition within the internal markets. The RSA and Zambia have the most developed electricity systems in the region. In addition, RSA operates its own tight pool regime: the National Power Pool (NPP), which is modelled after the England and Wales' pool arrangement. Lesotho started its own generation only in 1999; before that, it depended on ESKOM to meet its demand. Table 4.1 shows the development within the internal markets. The diversity in the states of development means that the pricing, tariff structure, planning and quality maintenance standard, differs between the member states.

**Table 4.1 (Source: Masawi, 1999)**

Development of Internal Markets of SADC Member States

| Member State | Unbundling and IPPs | Third party access | Tariff Reform | New regulatory framework | Reorganisation of distribution | Utility commercialisation |
|---|---|---|---|---|---|---|
| Angola | | | X | X | X | X |
| Botswana | | | | | | |
| Lesotho | | | | | | X |
| Malawi | | | | | | X |
| Mozambique | X | X | X | X | X | X |
| Namibia | X | X | X | X | X | X |
| South Africa | | | X | X | X | |
| Swaziland | | | X | X | | X |
| Tanzania | X | X | X | X | | X |
| Zambia | X | X | X | X | X | X |
| Zimbabwe | X | X | X | X | | X |

*Lack of transmission infrastructure*: All the countries are theoretically members of the SAPP. In practice, the lack of interconnectors limits the active participation of Angola, Malawi and Tanzania; consequently, as table 4.2 shows, that there are seven operating and 2 observer members (SAPP, 1999).

**Table 4.2**

Membership of SAPP (SAPP Annual Review Report: April 1997 – March 1999)

| Country | Utility | Abbreviation |
|---|---|---|
| Angola | Empresa Nacional de Electricidade | ENE* |
| Botswana | Botswana Power Corporation | BPC** |
| Democratic Republic of the Congo | Societe Nationale d'Electricite | SNEL** |
| Lesotho | Lesotho Electricity Corporation | LEC* |
| Malawi | Electricity Supply Commission of Malawi | ESCOM* |
| Mozambique | Electricidade de Mocambique | EDM** |
| Namibia | NAMPOWER | NAMPOWER** |
| Republic of South Africa | ESKOM | ESKOM** |
| Swaziland | Swaziland Electricity Board | SEB* |
| Tanzania | Tanzania Electricity Supply Company | TANESCO* |
| Zambia | ZESCO Limited | ZESCO** |
| Zimbabwe | Zimbabwe Electricity Supply Authority | ZESA** |
| Portugal / Mozambique | Hidroelectrica de Cahora Bassa | HCB*** |
| Zambia | Cooper Belt Energy Company | *** |

Note: *Not Operating; **Operating; ***Observer

Source: SAPP Annual Review Report: April 1997 – March 1999

*Unresolved disputes / unstable governance*: The regime is saddled with accumulated disputes that are difficult to resolve. The rule of law and enforcement procedure differs between the member states; the central governance and the regulatory contract are weak. The persistent changes in management personnel, which arise when the Heads of the Ministries that oversee the liaison between the member states and the regional projects change and new people are brought in. The weak institutional structure influences continuity of policy initiatives and prevents enforcement of penalties on the member states that breach the terms of engagement in the pool.

### *4.2.3 Summary*

Apart from the RSA most of the countries in SADC are still evolving through a traditional society. Therefore, they are a long way from the 'take-off' stage at which time that they may be able to overcome the hurdles of traditions and resistances (see Todaro, 2000; Ghatak, 1986; Donaldson, 1984). That is when they may have the capacity to generate their own investment and advance technologically at sufficiently high rates that economic growth and development will be virtually self-sustaining (see Bannock et al 1998). ESKOM is a dominant monopolist both in its domestic RSA and the regional market. Its has the economic and network strength to prevent similar and potentially efficient firms from entering the regional market. Its debt recovery policy will enable it to strengthen its power to operate independently of any other competitor that may enter the market (Kora, 1997).

## 4.3 Limitations to the success of regional markets

The previous section discussed the emergence of the regional power project in the SADC. In theory, provided that the socio-politics and economics of the country are sound, liberalisation can enhance the flow of foreign investment into sectors in the economy. The inherent features of the SADC were barriers to significant private and foreign equity in its regional power pooling arrangement. Consequently, some of the member states that ratified the IGMOU are not yet connected to the integrated network; and therefore are not active participants in the SAPP. There are also internal problems that have inhibited the development of contestable commodity and capacity regimes within the SAPP. The Heads of Governments incurred high transaction costs to unbundle their systems, deregulate and join the regional pool, with the objective that it may be a route towards the improvement of power supply. But they have not achieved that objective.

Many of the countries in Sub-Saharan Africa (SSA) have similar economics and socio-politics to the SADC member states. Also the Heads of Government in SSA countries and their policy advisors expect that similar regimes in the East, West and Central Africa will help them set the right tone for foreign investment into their electricity sector. Next, we focus on whether foreign investors will have any reason to behave differently and invest in these proposed pools.

The England and Wales' regulatory reform, which this thesis discusses in section 2, has some institutional and contractual conditions that any country must have for its competition policy to succeed. We believe that they are the same factors that would determine the quantum of foreign investment that SSA can attract whenever they launch their power pools. I ask what are the barriers to entry of foreign investors' post-power pools? The most common of these are:

- Political stability
- Low levels of corruption
- Energy Policies
- Good economic indicators
- Participation in the international bond market / access to debt finance
- Regulatory reform, regimes and contracts

### *4.3.2 Political stability*

Many of the countries in SSA have very unstable socio-political environments. There are a few well known issues, which are undemocratic and in most cases anarchist regimes; changes in the government in power; civil wars; assassinations, labour unrest, naturalisation programmes, politically induced riots (see Asteriou & Spiropulous, 2000). All of these will affect the social structure in these countries; as Mauro (1995) finds, the impact of political instability is higher in developing countries than in the developed and industrialised countries. It limits foreign investment and consequently slows down the pace at which a country develops. The political instability in SSA has worsened in the past 4 decades. Are there long term plans that these countries can use to combat political instability?

I believe that it would be a very difficult enterprise for the following reasons. Before the colonial invasion of SSA, the SSA people were rooted in their cultural, tribal and ethnic heritage. They were founded and governed through their local tribal Kings; these Royals used institutional frameworks, which included rudimentary controls and enforcement procedures to organise the people within their empires (Lienhardt, 1962; Freeman-Greenville, (eds).

1962; Biobaku, 1962; Ijediogor, 2004 a & b). The local tribal Kings were warriors and in those days, the wars that they fought were mainly directed at expanding territorial boundaries.

As Newbery (1962) shows, the SSA people were traders; they had a very effective trading relationship with Europe. The imperialists might have discovered the potential in these regions through the trading relationships. This may be why Rodney's (1985) exposition suggests that the constant ravaging of Africa for over 500 years was an economic necessity to Western Europe.

Following the colonial invasion, it seems that the imperialists paid little attention to the idiosyncrasies of the people that they invaded. They created and / or amalgamated boundaries; in some cases, separated multiple tribes and ethnic groups. In reality, the imperialists simply created new boundaries that consisted of tribe(s), ethnic group(s) and religious sect(s) that did not share similar ideologies or culture. Yet in the new social structure, they *must* learn to live together and foster the level of unity that is required to enhance a society's advancement to sustain economic growth. That is, the imperialists merged tribes, ethnic groups and religious sects that did not necessarily co-exist harmoniously and peacefully.

In the new social order there are over 1000 ethnic and tribal groups within some boundaries who have to grapple with the challenges of the western type of politics. The 'political parties' and 'democratic governments' are 'organisations' that represent and promote the interest(s) of tribe(s) / ethnic origin(s) and / or religious sect(s). The visions of most of the 'democratically elected' government personnel are significantly different from that which exists in advanced and industrialised countries. In the latter, similar interest groups usually seek ways to develop ideas and social classes. Consultations with the people that they govern and modifications of public policy initiatives, which ultimately improve social welfare, are some of the ways by which they do this.

Also, cultural, ethnic and religious inclinations of leaders dominate political ideologies and public policies that the 'democratic governments' initiate. Each tribe, religion, or ethnic group persistently seeks to be the leader and expects that the others be seen, but better still, not heard. This is one of the main causes of ethnic and religious rivalry in Nigeria; and in the past few years has motivated consultations on the best way for the country to rotate its presidency between the major tribes and / or regions.

The situation that I describe in the paragraph above is partly why power struggle makes tribal and religious clashes the most common form of political instability. Holding a political office enriches the office-holder and, on a smaller scale his tribe. As a result, many of the citizens are prepared to fight for it and a significant number of their young, jobless male cousins are particularly happy to pull triggers for them to achieve these desired objectives (Economist 19 – 25, 2003).[34] Also the "the preservation of ethnic hegemony is perceived as a condition for physical survival [sic]; therefore, killing becomes a moral duty[35]'. This partly explains why anthills of riots and assassinations, which are usually induced by politics of tribal or religious rivalry, persist and terrorise the pockets of very *fragile democracies* that exist in the region. There are also pockets of anarchist regimes and persistent labour motivated unrest that influences regional development. The labour unrest is the only means that the population has to express their sense of outrage at issues such as low earnings; and sometimes months of unpaid wages. They find these issues difficult to cope with, particularly in the face of the grand corruption that makes it possible for the top Government personnel and some of the privileged citizens that are closely 'connected' to the corridors of power, to live in luxury.

### *4.3.3 Corruption*

The corruption that this study considers is the Business International Corporation (BI) definition, which Mauro (1995) uses, in his empirical analysis of the effect of corruption on long-term investments and economic growth. There are all forms of 'questionable payments' that occur in the course of business negotiations. The common categories of corruption are petty corruption, corrupt management practices and grand corruption (Lovei & Mckechinie, 2000); there are also studies that generalise corruption into two broad classes: legislative and administrative corruption (Cartier-Bresson, 2000).

---

[34] The writer was reporting the democratic government and the elections in the Federal Republic of Nigeria (FRN).

[35] In discussing the Rwanda and Burundi ethnic factionalism, the writer borrows this from Rene Le Marchands (The Central Africa Crisis: a closer look).

Questionable business payments are made in many countries throughout the world. The difference between what happens in SSA and in some of the industrialised countries such as in Europe and the USA, is the severity, and the opulence, with which individuals demand and accept these payments, termed *'bribes'*. Moreover, these countries often have no property rights, no sovereign legal systems, and consequently no means to enforce penalties that can help to restrain the practise (Barhan, 1997; Clarke & Xu, 2002). The level of abject poverty and the weak nature of the civil society, may be the cause of the several levels of corruption in these SSA countries (Ngwane et al, 2000; Makenete et al, 1998; Tjonneland, 1998; Bardhan, 1997).

Is corruption really bad for any economy? This is another area where economists have diverse views. Some economists argue that corruption is 'an endogenously generated price mechanism that corrects disequilibria and restores optimal allocation in the market (Clarke et al, 2000:1; referring to an earlier work by Lui, 1985[36]). There are others who oppose this view; instead, they argue that bribery causes allocative inefficiency; they uphold that it is more distortionary than taxation (see Shleifer & Vishny, 1993:600; Rose-Ackerman, 1996). There is also evidence that it hinders economic growth and development; here again, the effect is worse in the developing than the developed countries (see Mauro, 1995; Asilis & Juan-Ramon, 1994).

My experience in Nigeria, which is the pattern in many of the countries in the SSA, is that bribery facilitates the speed with which goods and services are delivered. Businesses conclude contract negotiations faster if they are willing to pay the necessary bribes to a chain of interested middlemen. Bribes also curtail the transaction costs of negotiating businesses because it occurs at 'those points where the political, bureaucratic and economic interests coincide' (Cartier-Bresson, 2000:1). It's endemic nature, which to an extent makes it 'legal' (since it is publicly accepted) makes the citizens prepared to do anything within their power to remove any obstacle that potentially prevents them from receiving bribes. Therefore, it is known to motivate politically induced assassinations, coups and civil unrests. Even though it has this inherent potential to distort the social structure, politicians and some of those that hold high Government positions choose to carry on in a very oblivious manner. It suits them to do so, because otherwise, they would be distracted from pursuing their own private

[36] I have not yet seen this paper.

objective, which is simply to embezzle as much wealth as they can whilst they are in office, because the position (in their view) is usually very transient.

The situation that we explain in the paragraph above shows that firms often include top Government personnel and / or those that have access to the corridors of power, onto their payrolls. It is also known that bribes constitute the largest proportion of a company's 'entertainment budget', which is quite aside from unduly raising project costs and supply invoices.

Is it possible to combat bribery in SSA? Remember that this is a region that is endowed with an abundance of natural resources particularly oil and gas. As Ades & Di Tella (1999) put it, there is a positive correlation between the presence of oil and the level of corruption in a country. The severity of corruption varies across SSA; the highest levels are recorded in the countries that have more oil and gas than the others are. This is consistent with the world's corruption rating. In relation to the other countries in the region, Nigeria, which has more oil and gas, has the worst rating than its counterparts (see http://www.transparency.org). Assuming that the view of Ades and Di Tella holds, we believe that it will be a very difficult task to resolve, given the untapped quantity of oil and gas in the region. How will this situation affect foreign investment into the regional power pool?

Most countries in Europe and USA consider bribery as a criminal offence; therefore it is subject to legal prosecution. Corporate liability includes parent firms being directly responsible for the actions of their subsidiaries. These legal systems put pressure on the subsidiary firms in the developing countries, where giving and accepting bribes is a norm, to 'behave properly' in conducting their business transactions. Nonetheless, it is well known that the viability of firms in many of these countries in SSA depends on connections and the ability to pay the right bribes. The threat of legal proceedings for any foreign firm that is suspected of engaging in bribery related activities, aside from those that are actually caught doing so, prevents foreigners that have the capacity to operate electricity utilities from entering the regional market. The firms that are already established in the region face a different kind of challenge. They still pay bribes; and the viability of their businesses hinges on the 'good moods' and continuity of their contacts in the 'corridors of power'. Although it is well known that these foreign firms pay huge bribes to operate in the region, it is difficult for their home Governments to prove these. This is because, over time, they learned the rules, developed mutually beneficial 'bribery relationships' with the 'cronies' in the country that they have a very low risk of losing the returns on their investments. New entrants require time to build this level of relationship; even when they enter, there are no guarantees that they will succeed.

Remember that the electricity industry requires massive capital investment, which in many cases takes many years to amortise; and bribery related losses are not insurable risks. Therefore, it seems to us that given a choice, electricity utilities that wish to enter into developing countries would prefer to invest in some of the emerging economies in Europe, South America and Asia, where they will have reduced pressure to protect themselves against allegations of bribery related activities.

### *4.3.4 The politics of energy policy*

The Heads of Government control energy policies, which cover issues such as allocation and / or, procurement of inputs, for instance oil, gas and solid minerals and the setting of prices and tariffs. They also decide appropriate levies on solid minerals, subsidies and performance standards for firms in the energy sector. I established in sections 2 and 3 of this thesis that public policy and not price mechanism determines the structure-conduct and performance of electricity markets. This means that the people who influence these public policies must believe in the initiatives that they steer forward from conception, implementation, monitoring and the development of regimes. Ideally, Government commitment requires that on average, all the cadres: central, regional and local boroughs, should sing from the same 'deregulation hymn sheet' for the regulatory reform to succeed.

There is a mixed feeling amongst some of the Government personnel and citizens regarding the benefits of electricity deregulation. For instance, in Nigeria, some citizens see the privatisation of its National Electricity Power Authority (NEPA), the vertically integrated monopoly utility, as the only way for the country to have a reliable electricity supply. Many of the people that I speak to cite the success that the country made in deregulating its telecommunications. They take a very narrow view that the criteria for assessing the improvement to social welfare is that more people in the urban cities have access to telecommunication accessories (even when they cannot afford to maintain the associated bills). They believe that the growth in connections that has been recorded in telecommunications since the Government auctioned telecommunications licences in 2000 could not have been possible if the Nigeria Telecommunication Plc (NITEL), were still the monopoly service provider. Citizens in some of the other SSA countries that have also deregulated their telecommunications sector, which has given easier access to telephone services to some people in the cities, uphold the same view.

Nonetheless, there are some citizens who take a holistic view and have a different opinion from what I report in the paragraph above. This group considers the micro economic situation in evaluating the potential impact of electricity deregulation on social welfare. In South Africa, we find that the municipalities and local governments rely heavily on the spreads that they earn between purchasing electricity in bulk from ESKOM and selling it to the retail consumers. Keswell-Burns (1998) estimates that this spread is approximately 30%. They use this income to augment taxation and the periodic revenue allocation that they receive from the central Government, which they use to administer the Boroughs. These Borough Heads in conjunction with some other sub-Government officials and educated elite (who understand that electricity deregulation is significantly different from telecommunications) oppose the idea of introducing competition policy into electricity. The social structure is based on extended family dependency, which means that each worker in the population is a proxy and potential source of consumption to a family lineage. This makes it easier to appreciate that, apart from the loss of income to the Boroughs, there is this sympathetic case that deregulation will cause the source of living to some families to dry up. This expectation is consistent with the evidence from other parts of the world where electricity deregulation has occurred; hence making privatisation closely tied to retrenchment of labour. For instance, ILO (1998) reports that many employees lost their jobs in Europe following the deregulation programmes in the 1990s. Stiglitz (2002) finds that after Chile implemented the World Bank's structural adjustment policies, between 1982 and 1983, many people lost their jobs. Similarly, approximately 40% of employees had lost their jobs, in the first three years of electricity privatisation in Argentina (DOE, 1997). This implies that there will be many job losses in the deregulated electricity markets in SSA; and the increase in the number of the unemployed would worsen the abject poverty in a region, which does not have any functional social security system.

Also assassinations, stealing and armed robbery make many of these countries in SSA very insecure. Ngwane et al (2001) found that there is a multiplier effect of the increases in poverty levels in developing countries; it worsens anti-social behaviour. Based on that, I conjecture that post-deregulation, the crime and anti-social behaviour rates in SSA would increase; and would weaken the already very fragile social capital in the region. It makes sense, therefore, that some citizens would oppose competition policy in electricity.

It brings us to the concern about the long-term viability of any electricity utility. Persistent theft of infrastructure is a well-known challenge that electricity utilities face in the region. Also the growth in anti-social behaviour would include an increase in destruction of electricity infrastructure at every opportunity that arises. This cost, which affects the profitability of these firms, is quite difficult to quantify. Since these are not risks in so far as they are certain to occur, insurance firms are usually reluctant to issue polices to firms in the region. Therefore, we believe that electricity utilities, that have the capacity to provide the service in the region, will not consider making investments there.

### *4.3.5 Economics and debt situation*

Most of the countries in SSA have very poor economic indices: low GDP and growth rates; high rates of inflation and unemployment; and persistent balance of payment deficits. Moreover, 34 of these countries are listed under the World Bank *Heavily Indebted Poor Countries* (HIPC)[37] scheme. This means that the Bretton Wood Institutions accept that many of these countries in SSA require exceptional concessions to their debt situations. They are predominantly agrarian based; the agricultural sector employs approximately 60% of the working population and contributes between 40% and 50% of most country's GDP. Their debt as a percentage of export earnings as well as the level of their international debt in US Dollars, which table 4.3 presents, is also very depressing. Many of these countries are highly dependent on food and financial aids (see Makenete et al, 1998); they also depend to survive on the loans that they get from the international finance institutions. They depend very much on importation and the manufacturing sector is quite rudimentary. Their main source of foreign exchange being oil and solid mineral revenue (for those of them that are endowed

[37] The World Bank and the IMF initiated this scheme. It provides additional opportunities to some of the poorest countries in the world who already benefit from concessionary assistance through the International Development Association (IDA). This initiative enables them to obtain debt relief within an overall policy framework directed at poverty reduction.

with these). The top Government officials embezzle this revenue from natural resources, which means that the social structures in many of the countries are run down. All these features of the economics in the region symbolise abject poverty (Ngwane et al, 2001).

Many of the member states cannot afford to finance the reform of their electricity industry; whilst some of the international finance agencies commit to supporting them by providing financial assistance such as additional loan facilities, it will obviously worsen their international debt portfolios (Makenete et al, 1998; Tjonneland, 1998). Compared to the other countries in the region, RSA has outstanding economic indices; it is an active participant in the International Bond Market and compared to its regional counterparts (Senegal, Botswana and Mauritius), it has the highest sovereign credit rating. These are some of the reasons why RSA attracts most of the small investment that foreigners make to the region (Heese, 2000). I am also aware that some of the indigenous private entrepreneurs in SSA prefer to invest in RSA and not in their own home countries.

*Purchasing power of the final consumer.* Many of the citizens have difficulty paying the subsidised electricity prices under the vertically integrated public monopoly regimes. Private firms can only survive in market regimes if they are profitable (Sidak & Spulber, 1998). Assuming that foreign firms enter the market in the regional pool regime, they would expect to cover their average costs of production and make some reasonable return on their shareholders investment. It seems to us that compared to the period before, final consumers will face a much higher tariff post-deregulation, which they cannot afford. This is where Newberry's (2000) recommendation becomes very relevant: the case for privatisation should be made on the grounds that the final consumers can pay the post-privatisation tariff, because the market will fail if they cannot do so. This happened in Sweden, where the final consumers inability to pay the tariff and the financial constraints that the smaller industry players had from participating, threatened the success of the reform (ILO, 1998). Similarly the regulatory reform in the Republic of Armenia, which was directed at attracting foreign investment into the country's electricity sector failed. Kaiser (2000) reports that final consumers could not afford the post-privatisation tariff.

We conjecture that the citizens in SSA could not afford the market rates for electricity. The predictable outcome would be that the 'invisible hand' will re-allocate the resources of the firms and the consequence would be failure of the regimes. Since this failure is certain and not a risk *per se*, we conclude that it is a barrier to the entry of potential firms into its electricity industry.

| **Table 4.3 Summary of Debt in Sub-Sahara Africa** | | |
|---|---|---|
| *Country* | *Debt as a % of export earnings 1994 – 1996* | *Total debt US$ bn* |
| Angola | 260 | 10.6 |
| Benin | 324 | 1.5 |
| Burkina Fasso | 456 | 1.3 |
| Cameroon | 465 | 9.5 |
| Central African Republic | 439 | 0.9 |
| Chad | 359 | 1.0 |
| Cote d'Ivorie | 400 | 19.7 |
| Democratic Rep. Of Congo | 764 | 12.8 |
| Equatorial Guinea | 249 | 0.3 |
| Ethiopia | 1,377 | 10.0 |
| Ghana | 397 | 6.2 |
| Guinea | 449 | 3.2 |
| Guinea-Bissau | 3,509 | 0.9 |
| Kenya | 238 | 6.8 |
| Liberia | 414 | 2.1 |
| Madagascar | 557 | 4.1 |
| Mali | 624 | 3.0 |
| Mauritania | 473 | 2.3 |
| Mozambique | 1,1411 | 5.8 |
| Niger | 548 | 1.5 |
| Nigeria | 256 | 31.4 |
| Republic of Congo | 406 | 5.2 |
| Rwanda | 1,372 | 10.00 |
| Sao Tome & Principe | 2,132 | 0.2 |
| Senegal | 231 | 3.6 |
| Sierra Leone | 909 | 1.2 |
| Somalia[38] | 3,671 | 2.6 |
| Sudan | 2,131 | 16.9 |

### *4.3.6 Access to debt finance*

The International Bond Market (IBM) facilitates trading of a country's debts; it is also an avenue by which the participating countries can raise medium and / or long-term investments.

[38] This was before the genocide

There are independent agencies for instance Fitch, Standards & Poors and Moody who evaluate a country's credit worthiness; the indicative rates that they provide are known as sovereign credit ratings (SCR). These agencies use GDP; an index for qualitative variables such as the socio-politics, levels of corruption and some other business intelligence related inputs, to calculate SRC (Fitch, undated; Eaton et al, 1986; Cantor and Parker, 1995). This means that SRC is very important to an investor because it reveals the private attributes of the country and provides intuitive ideas about the safety of an investment; these can be points that the investor can rely on when negotiating contracts (see Bell, 2003). It also determines the quantum of foreign investment that a country receives (Cantor & Parker, 1996).

SSA expects that after these pools are launched, it will enhance the flow of significant foreign investment into its network capacity building. But there are only four countries in the region (Botswana, Mauritius, Senegal and South Africa) that are active participants in the IBM, and therefore, have an SCR. These four countries are located in the SADC and there are none in any of the regions that are proposing the new regional pool regimes. The countries that are absent from the IBM cannot trade their debt instruments in the world market, which limits the scope of getting foreign investors (Bird, 1988). Also many businesses in the region have limited access to finance and where they find willing creditors, the cost of capital is usually quite high (Konig, 1986) compared to what their counterparts in the other developing economies in Asia and South America may have access to. This implies that businesses have a lot of difficulty raising debt finance to meet their daily working capital requirements and for capacity expansion.

### *4.3.7 Regulatory reform, regimes and contracts*

The England and Wales' *licensed-based* regime has influenced deregulation reforms in many of the developing countries such as India, Uganda and Ghana (Bell, 2003). This thesis shows in section 2 that the regulatory reform and the initial policies that Britain adopted for vesting its electricity privatisation in 1990, were integral parts of its restructuring process.

For instance, the White Paper upon which its electricity industry was privatised: '*Privatising Electricity*' (1988), was a product of consultation between Government, legal institutions, academics and the industry. Thereafter came the Electricity Act (1989), which was the Statute that legalised the reform. The White Paper laid out the creation of a non-ministerial department, the electricity regulatory office: *Office of Electricity Regulator* (OFFER) headed by a *Director General of Electricity Supply* (DGES) who had the statutory responsibility to 'promote' competition in the generation and supply of electricity. The appointment of Regulators, and board and commission members are always transparent. The Regulators office was designed, as a non-ministerial department and the Director Generals are accountable to Parliamentary Committees for the operation of the regulatory office.

The energy regulators protect the independence of the regulatory office by publishing consultations, advice and policy decisions. Its role includes the licensing of grid users; the development of economic polices for the efficient regulation of the monopoly businesses of the System Operator (SO) and 'promoting' competition in generation and supply. It also sets tariffs; carries out the timely review of transmission and distribution prices and uses efficiency techniques to stimulate high performance standards between the distribution companies. In addition, it has the responsibility to identify and take the necessary steps to deal with the industry risks that might be against the public interest. There might be elements of 'regulatory capture' either in favour of the consumers or at times the regulated companies, but on average the process in England and Wales is seen all over the world as sovereign and an ideal.

In summary, the mature institutional framework facilitates the interaction between the institutions that facilitate reform processes. The Regulators decisions, legislation and judicial oversight continues to govern the success of the British regulatory reform and contracts (see Sidak & Spulber, 1998). As mentioned in section 2, this thesis identified that these processes require significant financing and professional expertise. Therefore, any country that expects to embark on regulatory reform should have the means of financing the process.

The success of the regional pools in SSA requires the organisation of similar regulatory regimes to that which exists in Britain. Given the rate at which the Heads of SSA countries and their policy advisors are steering forward the regional power pools, one would think that they already have mature financial markets, developed property rights and the sovereignty of the legal systems. In reality, the contractual frameworks such as regulatory offices and sector regulators are very weak in the places where they currently exist. I have already identified that this is a major hindrance to the development of the regime in SADC. I believe that the situation will not be any different in the other regional pools.

The level of corruption, which we have already discussed, and the absence of technical expertise raises additional concerns regarding the treatment of contractual problems, as well as what will happen to foreign investments if contractual terms change post-deregulation. It does not seem that Bells (2003) options contract can make any difference here because of the very weak legal system. The policy advisors and Heads of Government that are steering forward these regional pools ignore the enormity of the problems that they would be faced with after the pools are implemented.

### *4.3.8 Summary*

The inherent features of SSA countries are barriers to the entry of foreign investors whenever they implement the regional pools. Perhaps it is these constraints and the rampant corruption in the region that made Stiglitz (2002) (who understands the way politicians and public servants in the developing countries operate) refer to the privatisation programmes that these countries propose as *'Briberization'* projects. He contends that their direct effect would only be to enhance the personal wealth of the public servants and politicians involved in the processes and will not in any way improve the social welfare of the population. We think that these emerging regional pools will not be any different from the experience in SADC.

## 4.4 A path to Regional Market Arrangements

We have established that regional pools will not enhance the development of the electricity systems in SSA. We assume that there is a need to improve electricity supply in the region, therefore, this section makes some assumptions and uses the regulatory reform in England and Wales to develop a theoretic model, which SSA countries can rely on to move towards regional power pools. They can also use this model as a guide towards implementing competition policy within domestic markets of the countries that have not unbundled and deregulated their electricity systems.

The 1990s saw a growth in privatisation projects in SSA and in SADC, and the birth of the regional power integration regime. SADCs project has not met the objectives for which it was created; yet other Heads of Government in the East, West and Central African countries perceive its continued existence as a sign that it is successful. Therefore, they are consulting on modalities with which they will implement similar arrangements. We have already shown in the previous section that this policy dimension will be high transaction cost with no benefits to these economies. An added implication is that the private and publicly owned electricity utilities may be inefficient in the region. The next aspect of this study is to investigate a path, which these countries can follow to allow a smooth transition to competition regimes and eventual emergence of a successful regional power pool.

This path that we propose leads ultimately to a market arrangement. Remember that the idea of privatisation and deregulation is for price mechanisms to curtail production and allocative inefficiency. Also markets are preferred because the sustained pressures from the threats of take-over and decrease in firms' stock prices makes managers more efficient. The managers in the private firms have profit orientated and productivity goals; they are also penalised if they fail to meet their performance targets. Apart from Ghana, Nigeria, Cote d'Ivoire and Kenya, stock market trading is not well developed in the rest of the member states in SSA. This means that on average, the managers in the privatised regime in SSA will not have sustained pressure from the stock market like their counterparts in the mature economies, if the transition is just from the vertically integrated straight on to the regional market. The model that we propose, which involves a gradual transition from vertically integrated to a regional market is presented in figure 4.2; table 4.4 (in the appendix) shows the breakdown of some of the key milestones associated with each stage.

### *4.4.2 Assumptions*

*Starting point: Vertically Integrated Public Monopoly*: We assume that all the electricity systems start from the same vertically integrated and publicly owned monopoly utilities. The advantage of this starting point is that it makes clear what each country requires to achieve at each stage in the process. Also those who hastily organised the unbundling and deregulation of their systems can see what they have left undone; possibly what actions that may be required to remedy the effect of design inefficiencies in their current regimes on the development of the sector.

*Flexibility*. The regional markets should emerge slowly; it gives the countries the opportunity to stop at any stage if there are factors that affect its transition onto another stage. For instance, if there are financial constraints that make movement to the next stage impossible, a country can assess what is yet to be done and decide whether to continue, stop or modify the rules and design in line with the market challenges. The advantage here is that the key milestones that need to be completed before joining a regional market are clear from the outset.

*Timing.* The time period for completing each stage will vary between the countries; we expect that the rate of investment and the maturity of the institutions will be the main determinants. For example, since the RSA participates in the IBM, it is possible that it will take a shorter time to achieve its investment targets than the other countries. Also a stable political climate, reduced levels of civil unrest, development of property rights, developments of financial institutions and decreases in the level of corruption, will all influence the rate at which private and foreign investments are made into network expansion. This suggests that the Governments will have to initiate policies that can help them to combat corruption; to improve the social-politics and economics and develop the human capital in the region. Also these initiatives will not be one-offs, but will change as the stages emerges and the economies face different challenges.

*Tariffs.* We assume that the consumers can afford to pay the post-privatisation tariffs; the assessment of the customers' ability to pay is based on their earning capacity. As Newberry (2000) points out, the case for privatisation should be made and very strongly too, on the basis that consumers can afford to pay the post-privatisation tariffs. Therefore, it is important first of all to ensure that the consumers can pay for the service, which if they can, will guarantee that the privatised firms can exist as going concerns to meet their licence conditions. Phased emergence of the regional market will allow a gradual withdrawal of Government subsidy to the sector. The added benefit of the model is that it is possible to ascertain what the consumers can afford before each stage emerges.

*Demand.* From the outset, there will be a low demand for power supply. However, one of the issues in the region is that the Government lacks the funding to expand network capacity. Given the low demand, maximising the benefits of economies of scale will only come from the expansion of the density of supply. Low demand will be a problem during the vertically integrated stage; however, increases in capacity will give more geographic areas accessible to the network.

*Ownership of firms.* Foreign firms are not allowed to own 100% of an indigenous business. We expect that the weak institutional framework will limit entry of foreign firms. However, since these countries can only raise debt from foreign investors, they will be willing to reach a compromise with any prospective investor. This is where the type of variations of the option contract that Bell (2001) proposes comes in. The processes for organising the contracts will help to reveal the private attributes of the investment, thereby giving the investor the necessary parameters with which to conduct an unbiased profitability assessment. It will also give the country the opportunity to ensure that they curtail contract re-negotiations when they are already in the market.

*Revenue collection and taxation.* I assume that there is an well-organised metering service, hence losses from revenue collection will be minimal. Also there are processes for taxation revenue collection. This is important because I think that this is the only process by which 'tariff re-balancing' can be done whenever appropriate.

*Universal Service Obligations & Social Action Plans.* I hope that once the system moves into the market regime, Government can include any of its social goals as part of the licence conditions for the operators. Governments will also set up schemes by which they can manage issues that relate to fuel poverty.

### *4.4.3 The Model*

The model consists of five stages, which can be classified into the vertically integrated and the unbundled systems. The difference between the public corporation and the regulated monopoly is that the managers in the latter have profitability targets; they are more accountable to the shareholders and are encouraged to act in the best interest of the consumers. The market regime starts after the regulated monopoly regime is completed.

Separating generation and supply from the Grid functions are what distinguish the vertically integrated from the market regimes, which start from stage 3. It is not advisable to launch a free and independent bilateral contracting regime straight out of stage 3. Instead, it will be efficiency enhancing for the system to operate a compulsory pool first. But the decision to include or exclude capacity payment should be based on the energy balance of the system, the transmission and distribution capacities of the Grid (constraint boundaries), the physical size of the network and capacity mix (see section 3 part 3.11).

Commodity trading starts in stage 4. The industry participants would have had sufficient time to settle into the idea of a market regime. This may be the best time to introduce independent bilateral contracting and the requirement for individual balancing of injections with offtakes. I believe that at this stage, most of the institutional reform and contractual requirements would have been developed to the extent that reliable information is easily accessible. Also the utilities would be better placed to manage their daily operations. In addition, the finance sector would have matured sufficiently so that it would be possible for participants to use different products to hedge the risk of operating in the spot markets.

The last stage is the regional power pool. This market is driven by well-organised regional contractual and regulatory arrangements. The countries would be well prepared, plus the years of transition from the vertically integrated to this stage, would give the internal markets the expertise to contribute to the challenges and policy initiatives required for the regional organisations to succeed. Our expectation is that by the time this stage begins, the countries would have found ways to curtail some of the internal problems that they have to the extent that there will be confidence and transparency in the regulatory processes. This is also a stage at which regional treaties will emerge.

Government controls energy polices; therefore, we expect that public policy will be relevant in all the stages of the model. However, the policy initiatives at each stage will differ because they will be designed to meet the needs and challenges of each regime. For instance, Trading Acts will not be required during the Public and Regulated Monopoly regime; it will be more appropriate when the market arrangements emerge. But price controls can be applied throughout the 5 stages.

### *4.4.4 Limitations of the model*

We have already shown that the countries in SSA have some inherent problems that make the application or replication of the regulatory reform from Europe and the USA difficult there. Of course, there are a number of reasons why it will be impossible for these countries to follow the model, which this paper recommends; at least not at the rate at which the World

Bank expects them to implement these power pools. Also instead of spending the time pursuing why it may be difficult for these countries to ensure that they follow our model, it seems better to ask what they can do at this stage to improve the situation in their electricity systems. Some of the very generic issues that they need to address include:

- Commercial losses, which include theft of infrastructure and revenue losses are quite high; this seems an important factor for the finance problem that some of these utilities have. One thing that they can do is to franchise out supply services ( see Schmalensee, 1979) which would cover all aspects of metering, bills collection and customer services.

- The regulatory agencies have limited resources with which to operate. They also require quality training in economic regulation, competition policy and pricing issues. The Government should provide funding for training of these staff. The emergence of market regimes will require staff that can deal with challenges such as access pricing, setting tariff, standards of performance and price control.

- There is no gain to be made from retrenching existing staff on the basis that they do not have the expertise. The opportunity cost is actually very high if one considers that increases in the number of unemployed will increase the crime rate in the countries. The best approach would be to place an embargo on employment of unqualified personnel but to begin to recruit better qualified staff and to provide the less qualified with the opportunity to develop technical skills that they can use to support the work in the industry.

- Managers' performance can be improved if they are made more accountable and given some incentives to improve performance. One option can be to use benchmarking to stimulate competition between similar cost and revenue units within utility companies.

If the four generic issues mentioned above are dealt with, many of the countries in SSA will begin to set the pace for the evolution of successful market regimes. Indeed would first of all help the private indigenous entrepreneurs to have confidence in their countries to the extent that they too can invest. We believe once these local investors begin to trust their own investment in these countries, some of the foreign firms within the domestic markets will want to diversify into electricity. Of course, this will set the pace for additional foreign investments. It is also important that the Government manages the entry of foreign firms; it may be best to ensure that one firm, settles in and an assessment of the incumbents made before opening up the market to multiple foreign investors. There may be some merit in thinking that best value discovery can be made when it's all a free entry and exit situation. It may not work that way because an inefficient entry policy can deter incumbents from further investment into network capacity.

### 4.5 Conclusion & Further Research

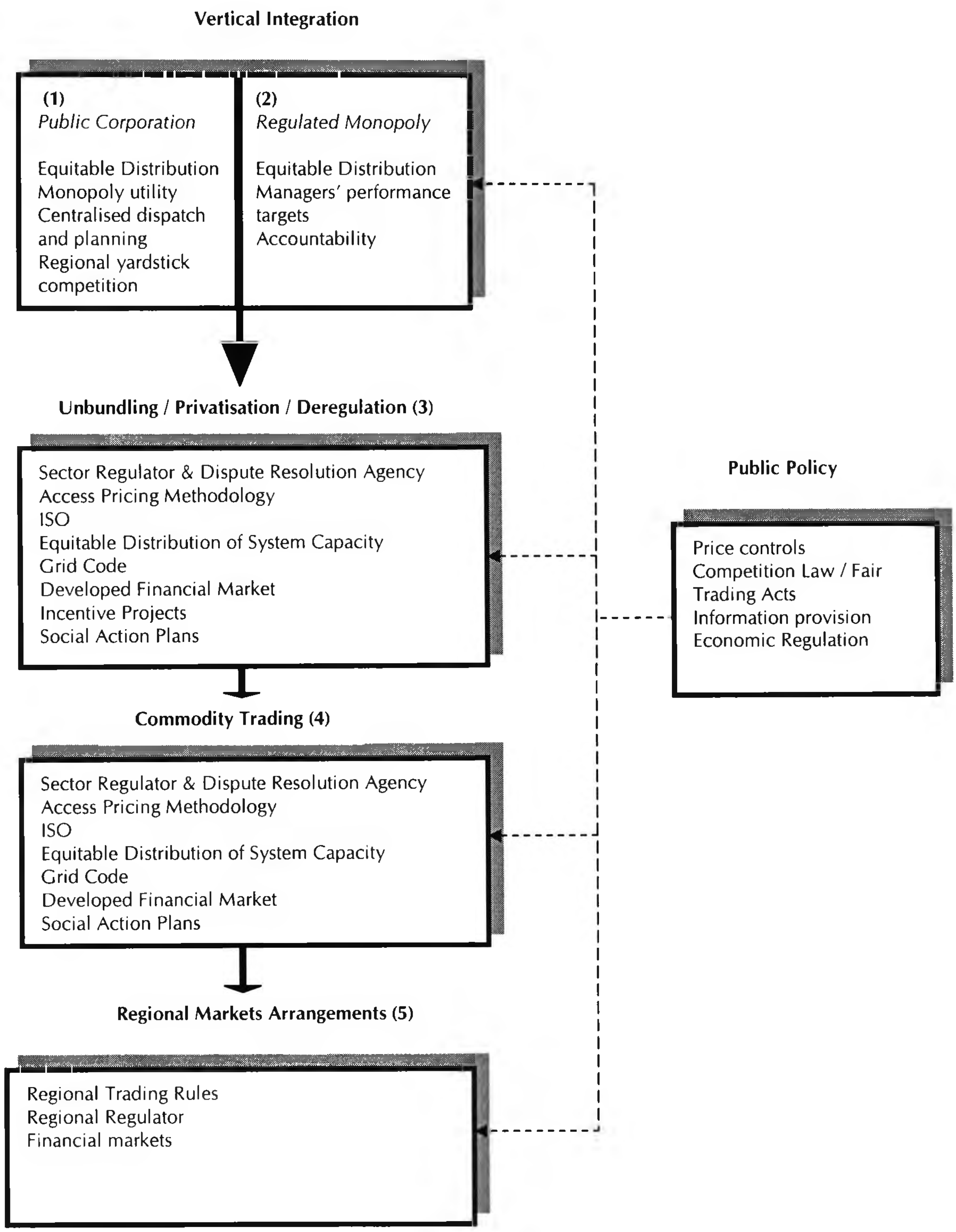
This study has shown that SSA countries will not earn significant benefits from implementing regional power pools in the East, West and Central Africa. It has introduced a model, which they can follow to introduce these pools, and if they consider, it may enhance the development of efficient competition within the regimes.

I have already mentioned the apparent influence of the acclaimed success of the deregulation of telecommunications on the Heads of Government in SSA. The empirical knowledge that most people have about these successes are limited to those published by the World Bank. Many of these studies on the efficiency gains that developing countries make from deregulating utility networks focus on telecommunications; and they provide insights into how deregulation increased the number of customers that have access to telephones. There is little said as to whether these have actually improved the social welfare of the African population. For example, how many hospitals, schools, roads, public transportation services have improved since 2000, when the Government auctioned the telephone bandwidth. How much foreign investment have the foreign firms brought into the country since then? What is the trend of the exchange rate since then; are there breaks in this trend correlated with the expatriation of earnings by the foreign firms?

At the present time, I find it worrying that the success in the domestic markets is only measured and continues to be correlated by the number of people that carry mobile phones. We believe that the answers to the questions that we highlighted in the last paragraph will give a different dimension to the idea that deregulation of telecommunications in the region is a great success. I expect to investigate these questions.

**Figure 4.2**

**Deregulation Matrix**



# Section 5

# Conclusion

This section discusses the inferences that are made in the whole thesis and which directly relate to the research questions that are set out in the introductory section 1, sub-section 1.6.

*Section 1.* This section used a historical case study approach to review the competition and regulation issues that arose in the evolution of the reformed British electricity industry. It found that contractual and regulatory structures must be in place and sufficiently mature before any country can embark on restructuring of its electricity system. The inference from the study is that competition policy will fail in any country that deregulates its electricity markets when these structures do not exist.

*Section 2.* This section used data exploration to describe, and ordinary least squares (OLS) regression plus a maximum likelihood (ML) technique in the Kalman Filter, to estimate the parameters in the models of the components of the PSP between January 1994 and December 2000.

The data exploration is an extended analysis of the earlier study by Wolak and Patrick (2001). It found that: demand, prices, spikes and volatility, were higher during the peak periods such as in the winter months, peak load regime and table A. In the winter season, it found lower: demand, prices, spikes and volatility, during the summer months, baseload regime and table B. The patterns of price variable gross demand and the Generators' availability during the spring and autumn lies between what occurs in the winter and summer months.

Capacity payment was the most volatile component of the PSP in all the sub-samples examined.

The decrease in PSP after 1998 was due only to the significant reductions in SMP. In contrast, the dramatic increase in LOLP increased CP. I conjectured that the increases in Uplift resulted from the higher costs that NGC might have spent to resolve constraints, start-up, availability payments, and increased Uplift. These suggest that the Generators changed their behaviour after 1998; in particular that they shifted their capacity manipulation from SMP towards Uplift and CP.

The weather seasonality, which prices exhibit, confirms that usage of power for heating is a reliable basis for separating electricity data between periods (see Granger et al, 1979). It therefore supports the approach of investigating weather and time dependency of prices, which this thesis carried out (quarterly, half-hourly, table indicated and load regime). Electricity systems that wish to enhance viability of the Generators' off-peak production, particularly at night time, when the retail consumers are more likely to use storage heating facilities, could do so by using two-part time of day pricing rules.

The price pattern that this thesis finds is consistent with the earlier findings by Fehr and Harbord (1993) and Wolak and Patrick (2001). The results from this present study confirm that the pattern of prices, capacity and demand plus their variability can be used as a generalised expectation of the price and capacity trend in an electricity market that has a similar capacity mix and structure as the England and Wales' pool.

The variance decomposition of SMP and CP provides an insight into the unobserved trends in these variables; it reveals properties of these variables that were not possible to ascertain in the data exploration. The results are consistent with those from the data exploration. There is also a structural break in the trend in SMP after 1998. This suggests that the policy to trade outside the pool, combined with the harmonisation of the trading arrangements in gas and electricity, might have placed a downward pressure on the Generators' bidding strategies after 1998. In contrast, there is no clear structural break in the path of the CP; instead it exhibits a violent short-run spike during week 52 in 1999; thereafter, it increases at a constantly increasing rate until the end of the 2000.

The DGES based his arguments of the Generators' earning excessive rents on the deviation between his estimated one-year avoidable costs and the pool prices. Wolfram (1999) uses the same mark-up approach to investigate the profits that National Power (NP) and PowerGens

(PG) made in the pool. Most of the cost estimation analyses are not reliable because the variables that they use to calculate the Generators true costs are usually incomplete. It therefore makes many of these studies unreliable for policy purposes. This study is the first to use the deviation between gross demand and declared availability to investigate spikes in SMP; and it finds no justification for the irregular SMPs. Therefore, it confirms that on average, spikes in electricity markets may not always reflect market fundamentals.

The estimation of the relationship between plant margin and SMP, CP and Uplift and declared availability is the first result of such an investigation on the pool variables. It confirms that on average, once a system exceeds the notional plant margin, prices may not be responsive to the excess capacity on the network. It raises the question about appropriate policy initiatives for timely resolution of constraint on networks.

This thesis reveals that, subject to a system capacity situation, paying Generators to make their plants available may distort price mechanisms. This is because Generators would always use capacity to manipulate prices. Capacity payments can be included in the price rule; it can only approximate price mechanism if it is robust and flexible to the extent that the Regulator can use it to encourage appropriate investments in network capacity.

The whole empirical analysis shows that price mechanisms can be inefficient in electricity markets. Therefore, this thesis concludes that policy advisors should expect that no matter the rule, which they choose for setting prices, the market outcome might still be inefficient some of the time and the aggregate costs of generation may be socially inefficient.

*Section 4.* The main problems in sub-Saharan Africa (SSA), which include the endemic corruption, the weak civil society; the lack of property rights, unstable socio-politics and poor economic indices, are barriers to foreign investment in the region. These countries need to build stronger civil societies in which institutions can develop and economic and political competition may be sustained (see Johnston, 1997).

## 5.2 Summary of Contributions

This is a 'shopping list' of my contributions to the entire research and knowledge in the literature on electricity industry restructuring and price behaviour in generation segment. My initial objective was to contribute to public policies by providing a better understanding of the possible outcomes in electricity markets. This appears to me to be an original contribution to the topic area and I am confident that I have sufficiently covered the aim for embarking on the study.

*Section 1.* Provides a better understanding of the regulatory oversight that emerging markets need to put into their electricity industry reform for them to succeed.

*Section 2.* Extends the earlier data exploratory work that Wolak and Patrick (2001) carried out. It provides a full account of the trend of prices, gross demand and the Generators' availabilities throughout the life of the pool regime.

It is the first study to:

- Use the three component variances decomposition in basic structural modelling (BSM), to investigate the unobserved properties in the pool prices variables (SMP and CP).
- Use the deviation between gross demand and the Generators declared availability to analyse the spikes in SMP.
- Estimate a time of day (TOD) system marginal price.
- Examine the relationship between plant margin, prices and Generators Declared Availability.

*Section 3.* There are many studies that report the development of the electricity privatisation in sub-Saharan Africa (SSA). Many of these studies have not explicitly considered the inherent factors in the region that make it impossible for the countries to attract the desired levels of foreign investment. This thesis contributes to the literature on privatisation in SSA by critically evaluating the factors that inhibit the development of contestable regimes in the power pools that the Heads of Government in the East, West and Central Africa expect to set up within the next five years. It also provides a stage path, which the countries can follow to introduce competition-related regimes.

## 5.3 Future Research

The essays uphold capacity and price manipulation as well as spikes being generic problems in the England and Wales' pool. The Regulator sought but failed to secure a legal basis to instil good behaviour on the Generators in 2000. However, the Competition Commission upheld no inclusion of market rule as a licence condition of Generators' licences. This raises some issues regarding (1) how best to secure lower prices and (2) how to ensure that the capacity manipulation does not jeopardise the security and safety of the transmission system. During the consultation of the NETA, the British Government thought that trading outside the pool, and which is carried out in a cost-targeted regime, would provide a long-term solution to the problems that plagued the pool. The DGES was certain that such a regime would reduce balancing costs. Further research into the following may provide evidence-based results of the benefits under NETA.

- If the pool did not close in 2001; or if the implementation date for the new electricity trading arrangements (NETA) were substantially delayed beyond 27 March 2001, was there scope that the CCs decision would have emboldened competitors to continue to act anti-competitively?

- NETA was designed so that systematic price differentials would not exist between the spot, the imbalance and the balancing mechanism (BM) (see OFFER, May 1998; OFFER, November 1997; OFFER, 1999). Given that Generators are contractually obliged to adhere strictly to their licence conditions, are there economic benefits of having market abuse conditions—or market rules—as licence conditions in residual pools?

- England and Wales implemented its new electricity trading arrangements (NETA) on 27 March 2001. One of the main features of the regime was the removal of capacity payments and the introduction of a penal imbalance *cash-out* regime. Generators that have any deviations between their day-ahead *final physical notification* (FPN) and metered positions at gate closure, earn system buy price (SBP), for spills onto the system. Or they are charged the system sell price (SSP), when they receive energy top-ups. Given that the objective for the design of NETA was to target imbalance costs directly at the Generators responsible for them, it appears that a comparative study on balancing costs between NETA and the pool will validate the effect of the changes in the trading arrangements on balancing costs.

The broad implication of the study of the factors that may inhibit the success of competition policy regimes in sub-Saharan Africa is whether these countries have actually made any significant efficiency gains from restructuring the telecommunications industry. If there is reliable data, it appears that it is important to quantify the real social benefit that these countries have made from restructuring and deregulation telecommunication networks.

- Section 3 is an empirical case study of the England and Wales' pool regime. Further research can be carried out on the relationship between plant margin and price in vertically integrated and publicly owned electricity utilities in one of the countries in SSA. The expectation of such a study can be that prices will increase with increases in the reserve margin because of the level of inefficiency with which public corporations deliver goods and services.

- Some policy advisors to the Heads of Government in SSA are convinced that competition policy is the only option that they can use to improve the density of supply, overall efficiency and to attract foreign and private equity into electricity network capacity building. They regularly cite the success which some of the countries in the region claim to have made from the restructuring and deregulation of the telecommunications industry. The important question is whether telecommunications deregulation has actually improved the social welfare of the population. It will be useful to conduct a comparative analysis of social welfare before and after privatisation in these regions. Some of the variables that could be considered in such a study would include foreign investment, social amenities and macroeconomic indicators such as the exchange rate and the level of employment within the sector.

I plan to explore these issues further.