PANEL SESSION 2

Invited paper 5

Evaluation, market power and transition pricing in the Australian electricity market

John Salerian
Industry Commission

5.1 Introduction

Australia is in the process of establishing a national electricity market. The objective is to develop a market that operates as close as possible to the concept of economic efficiency by creating competition in those components that are contestable. However, the technical nature of electricity production and transmission still requires a significant amount of regulation to create and run an electricity market that produces market outcomes consistent with economic efficiency.

The way in which the market is regulated can affect the degree to which the operation of the market is economically efficient. In addition to short term efficiency considerations, the operation of the market has important consequences for the way in which augmentation of both transmission and generation will take place. One important issue is whether the design of the electricity market will lead to an efficient system of transmission and generation in the long run.

Recently, there has been an increase in research relating to regulation of the industry and pricing issues, particularly on the pricing of electricity transmission in networks. For example, in 1996 there were several theoretical articles published in the Journal of Regulatory Economics (see Wu et al. 1996; Chao and Peck 1996; and Bushnell and Peck 1996).

A theme emerging from these articles is that some of the principles underlying proposals on the regulation of electricity markets are so-called ‘folk theorems’ — that is, they are commonly accepted assertions about the economic principles that in fact do not apply to the electricity market. For a discussion of these, see Wu et al. (1996).

Wu et al. (1996) claim that these assertions arise because the economic principles being applied are borrowed from other applications of economics, such as transport economics. However, the technology of electricity production and transmission is such that principles from other applications are inappropriate for the electricity market.

This paper, in part, aims to illustrate the economic principles embodied in an economically efficient electricity market through the use of a mathematical programming model of a hypothetical market. This model is an extension of a model previously developed as part of the Industry Commission’s study of the South Australian electricity industry. For that study, the Commission developed a mathematical programming model of the electricity industry in South Australia to evaluate market power issues (IC 1996a). Although the model included dynamic transmission losses, it did not include the externality associated with electricity flows in a network that arise because of Kirchoff's laws (see Chao and Peck 1996 for a discussion). It is this part of the electricity technology that distinguishes the economic principles required for electricity markets from those applicable in other markets.

The aim of this paper is to present an economic model of an electricity market that includes the unique characteristics that arise because of the nature of the technology used in electricity markets. This is achieved by revising the methodology used to represent the transmission. In the model presented here, the power flow equations and variables for a network are included, based on the methodology outlined in Chao and Peck (1996).

The model is then solved for a hypothetical electricity market involving 12 generators distributed around a network consisting of four nodes. Demand for electricity takes place at two of the nodes. The model represents an annual market consisting of 34 time periods, that is, the 8760 hours in the year have been allocated to 34 time periods (load blocks).

The results are then used to illustrate some of the economic issues that arise in electricity markets. The model is also used to illustrate how the addition of a transmission link can affect all transmission links, nodal prices and merit order dispatch of power stations.

The structure of the paper is as follows. The mathematical programming methodology is outlined in section 5.2. To set the scene, section 5.3 presents the structure of the hypothetical electrical market used in the model. Section 5.4 describes the formal model used here. Section 5.5 outlines the methodology relating to the demand for electricity. Section 5.6 outlines the methodology for the transmission of electricity and section 5.7 describes the methodology used for the production of electricity. In section 5.8 the two scenarios to be simulated are described. The results for the simulations are described in section 5.9. Section 5.10 is the conclusion and suggestions for further research.

5.2 Mathematical programming methodology

Samuelson (1952) showed that it was possible to construct a maximisation problem that guarantees fulfilment of the conditions of perfectly competitive equilibria among spatially separated markets. This provided the opportunity to use mathematical programming to simulate market behaviour. Later, Takayama and Judge (1971) significantly extended the applicability of the technique by showing that the competitive and monopoly models could be formulated as quadratic programming models.

Takayama and Judge (1971) also showed that two alternative formulations, the quantity formulation (primal) and the price formulation (the purified dual of the primal) could be used. Takayama and Woodland (1970) proved the equivalence between these two formulations. Takayama and Judge (1971) also showed that the quantity and price formulations could be combined to form another maximisation problem where both quantity and price are explicit variables in the model — this is the ‘general’ formulation referred to by MacAulay (1992) and is sometimes referred to as the ‘self-dual’ and ‘primal-dual’ formulations. Takayama and Judge (1971) also refer to it as the net social revenue formulation.

The general formulation has wider applicability. For example, it applies where interdependent demand functions do not satisfy the integrability condition (that is there is no unique solution to their integration) or where policy requires constraints on both prices and quantities.

In this particular study the quantity formulation has advantages over the general formulation. First, it reduces the number of variables and equations, which is important when dealing with large scale models. Second, it is easier to explain the technique and develop and implement the model using the model generating software, GAMS.[1] This is important when the time to complete the study is short. For similar reasons, the mathematical programming used here has advantages over other related techniques used to compute economic equilibria, such as non-linear complimentary programming and computable general equilibrium models.

In electricity markets, costs and demand conditions vary by time and from place to place. For example, electricity demand can be met by generation from a range of technologies (gas and coal), a range of plant sizes or imported via transmission lines from interstate. Therefore, spatial–temporal models, which include elements of networks, are particularly useful to capture this complexity.

Theoretical developments and application of this methodology to study pricing and deregulation in spatial energy markets increased during the 1980s. Some examples include Salerian (1992), Kolstad (1989), Provenzano (1989), Uri (1989), Hobbs and Schuler (1985), Sohl (1985) and Uri (1983).

In the mathematical programming model developed here, the supply of electricity is represented by mathematical programming models of power stations and transmission, rather than as supply functions. Mathematical programming has been widely applied to electricity supply, primarily to

1 For more information on the GAMS computer software package see: Brooke, Kendrick and Meeraus (1992); Meeraus (1983); and Bisschop and Meeraus (1982).

evaluate the least cost options to meet pre-determined demand. That is, demand is exogenously specified. Examples include Scherer (1977) and Turvey and Anderson (1977). Two Australian applications of note are the Australian Bureau of Agriculture and Resource Economic's (ABARE) version of the MENSA model (Dalziell, Noble and Ofei-Mensah 1993) and the Commonwealth Scientific and Industrial Research Organisation's earlier version of the MENSA model (Stocks and Musgrove 1984).

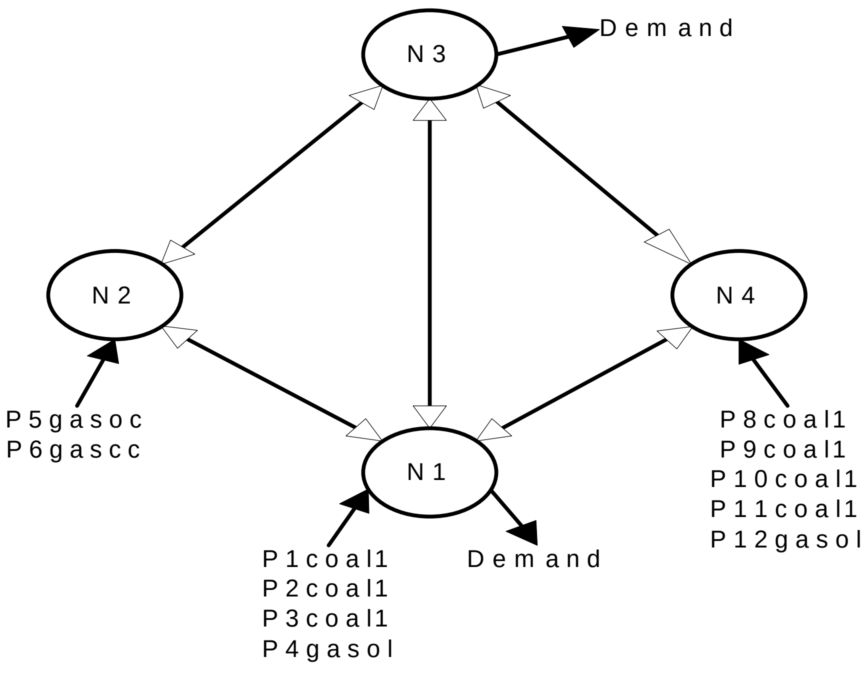
An advantage of the model presented here is that the quantity demanded (and price) is endogenous to the model.

5.3 Structure of an electricity market

A hypothetical network is used in this study. It is made up of four interconnected nodes — N1, N2, N3 and N4 (see figure 5.1).

Demand for electricity takes place at nodes N1 and N3. Generation can take place at nodes N1, N2 and N4.

Figure 5.1: Structure of the hypothetical electricity market



There are two nodes that only generate electricity — N2 and N4 — that is, they have only have generators that supply electricity to the network. Whether or not the node supplies electricity to the network depends on the types of generators that are situated at the node, the demand period and other constraints on the network. For example, there are five generators at node N4 — four coal, one gas — that represent ‘base load plants’ and ‘intermediate load plants’. This suggests that this node will supply electricity in the base period, and consequently in all periods. There are two generators at node N2 — representing gas open-cycle and gas-combined cycle generation — that are ‘peaking and intermediate plants’. This suggests that N2 will supply electricity in peak periods if there are no constraints on the network, but may supply off-peak if other constraints on the network exist. The latter case is examined in more detail in section 5.9.

There is one node that both generates and demands electricity — N1. This node can be thought of as a city load centre with some generation ‘close’ to it. Like node N4, node N1 has a combination of base load and intermediate load plants — three are coal fired and one is gas fired. Since N4 also has demand at the node, whether or not it supplies electricity to the network again depends on the types of power stations at the node, the demand period and other constraints on the network and also the amount of demand at the node.

There is one node that only demands electricity — N3. This node can be thought of as regional load centre with no generation nearby and is always an importer of electricity.

5.4 Mathematical model

This section formally describes the model used in this study. As mentioned in the previous section, the model presented here simulates the economically efficient market equilibrium. Wu et al. (1996) refer to this as economic dispatch.

The model used here has some non linear variables in both the objective function and constraints. It also has some variables that can have negative values.

Notation

For convenience, the notation used here to present the model is based on the GAMS computer modelling language. The notation used to present the model is divided into sets and subsets, parameters and variables.

Sets and subsets

b = set of time periods (load blocks)

n	=	set of nodes in the network
p	=	set of generating plants
d(n)	=	set of nodes where there is demand for electricity
gr(n)	=	set of nodes where generators are located
ng(n,p)	=	set of generators at each node
lk(n,np)	=	set of nodes that are directly linked in the network (node links)
nt(n)	=	set of nodes for which voltage phase angle variables exist

Parameters

Demand

a(n,b)	=	intercept of the inverse linear demand function for each node in each time period
w(n,b)	=	slope of the inverse linear demand function for each node in each time period

Generation

plantcost(p)	=	annualised fixed costs ($ million/GW) of each generator
opcost(p,b)	=	operating (fuel) costs ($ million/GW) — adjusted for the duration of each load block
genmax(p)	=	maximum allowable capacity of plants (MW)
pdata(p,‘avail’)	=	proportion of installed capacity available for use

Transmission

number(n,np)	=	number of transmission lines between nodes
v(n,np)	=	transmission line voltage (kV)
r(n,np)	=	transmission line resistance (ohms)
x(n,np)	=	transmission line inductance (ohms)
gridcap(n,np)	=	transmission capacity (MW)

Variables

QS(n,b)	=	amount of electricity supplied by each node in each time period (GW)
QD(n,b)	=	amount of electricity demanded by each node in each time period (PWh)
THETA(n,b)	=	voltage phase angle at each node in each time period (radians) — these are free variables that can have negative values
QGO(n,b)	=	output of each plant in each time period (GW)
QGC(p)	=	installed capacity of each plant (GW)
QP(n,np,b)	=	quantity of power flow between each node in each time period (GW) — these are free variables that can have negative values
NSW	=	net social welfare ($ million)

Equations

Objective function ($million)

```
NSW = sum((n,b)$d(n),(alpha(n,b)*QD(n,b)+0.5*ibeta(n,b)*(QD(n,b)**2)))
  - sum((p,b),opcost(p,b)*QGO(p,b))
  - sum((p),plantcost(p)*QGC(p))
```

The objective function maximises net social welfare (measured as consumer plus producer surplus — the area under the demand curve minus the sum of the variable costs). The first right hand side term in the equation is the area under the demand curve (integral of the demand curve). The second and third terms are the variable operating costs of the power stations and the annualised fixed costs of each GW of generating capacity for power stations.

Installed capacity balance (GW)

```
QGO(p,b) =l= QGC(P)*pdata(p,'avail')
```

This equation ensures that the output of each power station in any time period is less than its available installed capacity.

Maximum generation capacity constraint (GW)

```
QGC(p) =l= pdata(p,'units')*pdata(p,'genmax')/1000
```

The maximum generation capacity constraint ensures that the amount of capacity installed is no greater than the maximum allowed capacity specified for each plant.

Node supply (GW)

```
QS(n,b) =l= sum(p$(ng(n,p)),QGO(p,b))
```

The supply (injection) of electricity at a node at a point in time must be less than equal to the output of generators operating at the node at that time.

Real power flow equation (MW)

```
QP(n,np,b)*1000 =e=  g(n,np)*(v(n,np)**2-v(n,np)*v(np,n))
  + y(n,np)*v(n,np)*v(np,n)*THETA(n,b)$nsa(n)
  - y(n,np)*v(n,np)*v(np,n)*THETA(np,b)$nsa(np)
  + 0.5*g(n,np)*v(n,np)*v(np,n)*(THETA(n,b)$nsa(n))**2
  - 0.5*g(n,np)*v(n,np)*v(np,n)*2*THETA(n,b)$nsa(n)
    *THETA(np,b)$nsa(np)
  + 0.5*G(n,np)*v(n,np)*v(np,n)*(THETA(np,b)$nsa(np))**2
```

The amount of real power flow is related to the physical characteristics of the transmission line between the two nodes — its voltage, impedance and resistance — and the difference between the voltage phase angle at the node at each end (see Chao and Peck 1996 for a description). The real power flow equation uses an approximation of Kirchoff’s laws to determine the amount of real power that flows between two nodes, including transmission losses. The approximation is valid for small differences in voltage phase angles, which is the case under normal operating conditions. The approximation formula also ensures that the loss equation is convex.

The power flow variable is a free variable and there is one variable at each end of the link between two nodes. By convention, a negative value means power is being delivered to the node via the link. A positive value means that power is being exported from the node via the link. The sum of the two power flow variables on each link is the total transmission loss over the link.

The variable, theta, is the voltage phase angle and is a free variable. The number of voltage phase angle variables is equal the number of nodes less one.

Supply and demand balance (GW)

```
QD(n,b)$d(n)*(1000/hours(b)) =e= QS(n,b)$gr(n)
  - sum(np$lk(n,np),number(n,np)*QP(n,np,b))
```

The supply and demand balance states that the quantity of electricity demanded at the node in any time period must be equal to the sum of the quantity of electricity supplied at the node and the net power flow imported or exported at the node.

Transmission capacity constraint (GW)

```
number(n,np)*QP(n,np,b) =l= gridcap(n,np)/1000
```

The transmission capacity constraint ensures that the total amount of power transmitted between the nodes in any time period is no greater than the maximum (thermal) capacity along that link.

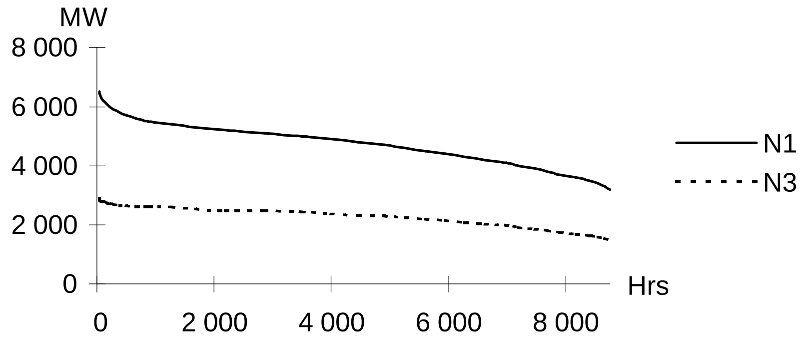
The full model description in the GAMS programming software is in Appendix A.

5.5 Demand

For each node, the load duration curve provides a useful description of demand across the year (see figure 5.2).

With a single node, the load duration curve is obtained by arranging the hourly loads at that node during the year into descending order (Turvey and Anderson 1977; Scherer 1977). However, in a network with more than one demand node, there is an additional complexity introduced because the demands at each node must be for the same points in time.

Figure 5.2: Load duration curve for nodes N1 and N3



To ensure that demand at each node are coincident in time, the following procedure is used. First, the load duration curve for one node (N1) was derived in the manner described above. Second, the load duration curve for the second demand node was determined using the chronological order of loads from the load duration curve in N1. This method was chosen because of its practical convenience. However, the method introduces some averaging issues into load duration curves of nodes other than the base node, N1. This means that the share of the implied load duration curve for N3 may differ from that of its actual load duration curve. It would be useful to further investigate the effects of any bias and consider alternative methods for determining load blocks.

Thirty four demand periods are defined by dividing the load duration curve for the first node into 100 MW blocks. This creates load blocks of unequal duration, measured in hours. Each of the 34 load blocks is assumed to have an independent linear demand function that relates the amount of electricity demanded to its price.

In this study, the demand function is:

$$\text{Price} = a + w * \text{Quantity}$$

This means that the quantity (and price) of electricity is endogenous. Any other effects on demand are assumed exogenous and are implicit in the constant term of the demand function.

The parameters of each demand function, a and w**,** are calibrated using given prices and quantities and an assumed own-price elasticity of demand, E, by:

$$a = \text{PRICE}\ (1 - 1/E);\ \text{and}$$

$$w = 1/E * P/Q$$

Here the price-elasticity of demand is assumed to be -0.3.

5.6 Transmission

The presentation here is based on that of Chao and Peck (1996). The real power flow in a network, based on Kirchoff's laws, is given by:

$$Q_{ij} = G_{ij}V_{ij}^2 - G_{ij}V_iV_j\cos(A_i-A_j)+Y_{ij}V_iV_j\sin(A_i-A_j)$$

where Q is the power flow from node i to node j and G, V and Y are parameters relating to resistance, voltage and admittance. A is the voltage phase angle at each node. The voltage phase angle and the power flow can be negatively valued. When the power flow for Q_{ij} is negative, the power flow is from node j to node i. The transmission loss along the line is given by $Q_{ij}+Q_{ji}$.

Under normal operating conditions, the real power flow equations can be approximated by the following convex function:

$$Q_{ij} = G_{ij}(V_{ij}^2 - V_iV_j) + Y_{ij}V_iV_j(A_i-A_j) + G_{ij}V_iV_j(A_i-A_j)^2$$

In this power flow equation, the marginal transmission losses are higher than the average transmission losses and the losses increase with power flow. This creates a rent on the transmission of electricity, so that the value of electricity exported (sold) out of the network exceeds the value of electricity imported (purchased) into the network. This rent represents the income earned by the whole of the transmission network.

There will be an additional rent earned by the network if one or more of the transmission links is constraining, that is at it maximum transmission capacity. The marginal value or unit price of any transmission constraint is given by the shadow price or value of the Lagrangian value associated with the constraint on the transmission capacity. These rents have been termed the 'merchandising surplus' by Wu et al. (1996).

Each of the nodes are connected by links made up of a number of transmission lines. The assumed technical properties of the transmission lines and the assumed maximum transmission capacities along the link are described in table 5.1.

Table 5.1: Characteristics of the network

			Transmission lines			
Link	*Link capacity (MW)*	*Distance (km)*	*No.*	*Voltage (kV)*	*Resistance (/km)*	*Impedance (/km)*
N1⇔N2	3 000	150	6	330	0.03	0.3
N1⇔N3	1 500	375	4	500	0.025	0.25
N1⇔N4	1 500	150	2	330	0.03	0.3
N2⇔N3	3 000	450	5	330	0.03	0.3
N3⇔N4	2 000	400	2	330	0.03	0.3

The distance between nodes plays an important role in determining the overall characteristics of the line. For example although the same type of line links nodes N2 and N3 and nodes N2 and N1, the overall characteristics of the line vary significantly. In particular, the total resistance along a line between N2 and N3 is three times as great as the total resistance between N2 and N1 because the distance between N1 and N3 is three times as great as the distance between N2 and N1.

5.7 Production model

The amount of electricity supplied in each period by a node with generators is the sum of the amount of electricity generated by all plants at that node in that period.

The amount generated at the node is limited by the technical capabilities, operating capacities and costs of the plants at the node (see table 5.2). A plant's output in any period can be no greater than its installed capacity, which in turn can be no greater than the maximum available capacity of the plant.[2] For example, although the maximum capacity that exists at node N1 in any period is 4021 MW (reflecting the maximum capacities of the plants at the node) the maximum supply from the node may in fact be limited to be a smaller amount if the installed capacities at any plant are less than the maximum (reflecting the cost of installed capacity). The different cost structures of plants also place constraints on the order and time that plants are loaded. This 'merit order' means that base plants supply electricity in off-peak periods, while base and intermediate plants supply electricity in intermediate periods, and base,

[2] With maintenance and other outages, the output is usually less than the available capacity. However, in this case availability is assumed to be 100 per cent.

intermediate and peak load plants all operate to supply electricity in peak periods.

These principles are readily applied in mathematical programming models (see Munasinghe 1990). The model used in this study is based on that of Turvey and Anderson (1977) and is similar to others, such as ABARE's MENSA model (Dalziell, Noble and Ofei-Mensah 1993) and the Industry Commission's model assessing the potential for market power in electricity supply in South Australia (IC 1996b).

Table 5.2: Plant data

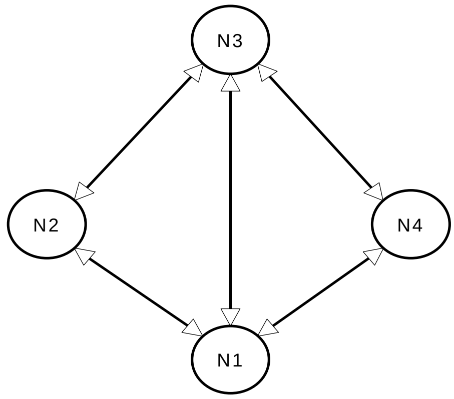
Node	*Power station*	*Availability*	*Maximum allowable capacity (MW)*	*Fuel cost ($ per MWH)*	*Capacity ($m per MW)*	*Life (years)*
Node 1	P1coal1	1	1 270	14.8	1.2	30
	P2coal1	1	861	14.3	1.4	30
	P3coal1	1	2 540	14.3	1.3	30
	P4gasol	1	500	18.5	1.2	30
Node 2	P5gasoc	1	unlimited	32	0.5	30
	P6gascc	1	unlimited	23.5	0.85	25
Node 3	P7dist	1	unlimited	150	0.3	30
Node 4	P8coal1	1	1 268	12.1	1.45	30
	P9coal1	1	960	13.0	1.3	30
	P10coal1	1	1 268	12.2	1.4	30
	P11coal1	1	890	14.8	1.25	30
	P12gasol	1	400	18.5	0.8	30

5.8 Scenarios modelled

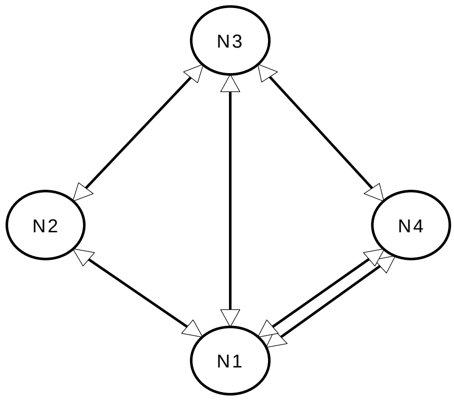
As the market develops through time extra generation capacity could be installed or existing capacity removed and new demand centres could develop or existing ones could contract. While these changes in themselves affect market outcomes, they are not modelled in this study. Rather, the study concentrates on how changes to the network itself - through the augmentation or creation of additional links between nodes - affects the real power flow between nodes and the market outcomes. In order to do this, two scenarios are considered (see figure 5.2) and their outcomes compared.

Figure 5.2: Time scenarios

Scenario 1



Scenario 2



An additional line is added between nodes N1 and N4, increasing its transmission capacity.

5.9 Results and policy discussion

Once the physical properties of electricity flow are considered, changes in one area of the network will affect all lines and nodes. For example when transmission along one link in the network is constrained, it not only affects the power supply to the nodes that it joins but may also affect every other link in the network, every other node and consequently all generators and demand throughout the network. Similarly, augmenting a link between two nodes or creating a new link between nodes that are not currently connected may have far reaching, and perhaps unexpected consequences elsewhere in the network. It may be the case that augmenting an existing link (or building a new one) may be beneficial to the nodes that either link connects and profitable for the owner of the link, but will decrease the overall welfare in the entire network if there are large adverse effects in other parts of the network which offset the benefits at the link. Or it could be that potential links that increase the overall welfare of all participants in the network are not built because of the adverse effects on certain existing participants.

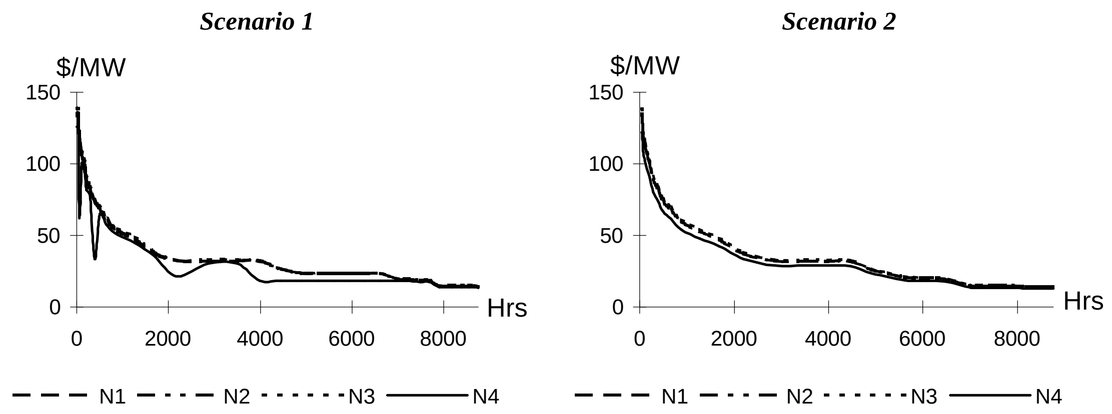
Scenario comparisons

Some of the effects proposed above can be illustrated by comparing the market outcomes between Scenario 1 and Scenario 2 over all periods (see figure 5.3). These effects are examined in more detail through a comparison of the power flows, node prices and merchandising surplus at three distinct time periods - a base load period, an intermediate load period and peak load period (see figures 5.4 and 5.5). It is also interesting to compare the merit order of plant dispatch at various nodes under the different scenario to see the effect of the linkage on individual plants at nodes (see figure 5.6).

Nodal prices

The model is ‘price endogenous’ meaning that prices are chosen as part of the solution. As expected the prices are higher in peak load periods and lower in base load period (see figure 5.3).

Figure 5.3: Nodal prices — all nodes, various scenarios



An interesting feature occurs at the very peak demand periods. Here the competitive market solution involves rationing supply so that demand is not satisfied rather than dispatching further high cost plants to satisfy very high demands. Consequently at the very peak demands prices increase even further.

Without transmission constraints (Scenario 2), prices are approximately the same in all periods. In contrast, when there are transmission constraints that limit the supply of electricity from node N4 (Scenario 1), the nodal price at the node varies significantly from those seen at other nodes.

Power flows and nodal prices

The power flows and nodal prices for peak load, intermediate load and base load periods are illustrated in figures 5.4 and 5.5.

In Scenario 1, the market outcomes at peak load and base load are as expected — power flows from nodes with low nodal prices to nodes with high nodal prices. At peak load, all plants are operating and node N2 (with the peak load plants) is supplying electricity into the network. Although power flows through the link N1⇔N2⇔N3 in the base load period, node N2 does not supply any electricity to the network. Rather the flows reflect the link's use as an alternative route for electricity supply, reflecting the operation of Kirchoff's laws.

An interesting outcome occurs at the intermediate load. Here transmission between nodes N1 and N4 is constrained by the maximum grid capacity along the link. The amount of electricity that node N4 supplies directly to node N1 and indirectly to the system through node N1 is limited. Equally importantly, the constraint between nodes N1 and N4 also constrains the amount of electricity that node N4 supplies directly to node N3 because of the effects of Kirchoff's law. As a result of the combination of these factors, node N2 now supplies electricity to nodes N1 and N3 — and does so even though N3's nodal price is less than N2's costs. Similarly node N1 sells to node N3 for a demand price less than its costs. In this case nodes N1 and N2 supply electricity below costs in order to alleviate bottlenecks elsewhere and are compensated for doing so by other nodes in the network.

As expected, in Scenario 2 the direct effect of augmenting the link between nodes N1 and N4 allows more power to flow along that link in all periods — which also removes the constraint on transmission experienced in the intermediate period in Scenario 1. Now power always flows from nodes with low nodal prices to nodes with high nodal prices.