NERA
ECONOMIC CONSULTING

# Estimating Long Run Marginal Cost in the National Electricity Market

## A Paper for the AEMC

19 December 2011

## Project Team

NERA Economic Consulting

Adrian Kemp

Martin Chow

Greg Houston

Oakley Greenwood

Greg Thorpe

NERA Economic Consulting
Darling Park Tower 3
201 Sussex Street
Sydney NSW 2000
Tel:  +61 2 8864 6500
Fax: +61 2 8864 6549
www.nera.com

# Contents

# List of Tables



# List of Figures

# 1. Introduction

This report has been prepared for the Australian Energy Market Commission (the Commission) to assist it in its consideration of a rule change that has been proposed by the Major Energy Users (MEU). The rule change proposal seeks to address concerns that, on days of very high demand, large generators are able to cause the wholesale spot price for electricity to increase by more than it should by offering prices that far exceed their costs.

In June 2011, we developed a paper describing the economic concepts of 'competition' and 'market power' in order to develop a framework for assessing the concerns identified in the MEU rule change proposal.[1] Our earlier paper also considered the appropriate market definition for the consideration of the Rule change.

An important conclusion from our earlier paper was that if market prices are significantly and persistently *above long run marginal cost (LRMC)* then this should, given time, prompt new generation investment that restores prices to those levels. In particular we concluded:[2]

> It follows that any assessment of whether a generator has a substantial degree of market power consequently requires:
>
> - a focus on genuine and enduring barriers to entry and expansion, as the fundamental source of the substantial market power, noting that this might also include 'strategic' barriers to entry and expansion; and
> - the undertaking of long-term price cost tests, as evidence of the exercise of market power, eg, comparisons of average spot prices to the LRMC of adding capacity, rather than comparisons of spot prices to SRMC at particular points in time.

This paper describes the practical methodologies that can be used to undertake the long-term price cost tests. Its focus is on:

- explaining the complexities involved in estimating LRMC for the relevant market; and
- explaining the approaches that can be used for calculating average market prices; and
- describing the practicalities of applying a hypothetical monopolist or SSNIP[3] test in order to define markets for generators operating in the National Electricity Market (NEM).

The remainder of the paper is structured as follows:

- **section two** explains the concept of marginal cost, the distinction between LRMC and short run marginal cost (SRMC), and describes the methodologies that can be used to estimate the LRMC for wholesale electricity markets;
- **section three** explains how the methodologies for estimating LRMC for wholesale electricity markets can be practically applied;

---

1 Green, H., Houston, G., and Kemp, A., (2011), 'Potential Generator Market Power in the NEM', *A Report for the AEMC,* NERA Economic Consulting, June.

2 *Ibid*, page 47.

3 SSNIP stands for small but significant non-transitory increase in price. See our earlier report, page 34.

- **section four** sets out the approaches that can be used to estimate average NEM spot prices, the factors that will likely influence observed market prices, and how contract prices can be taken into consideration; and
- **section five** sets out a practical approach to applying the SSNIP test for determining the geographic boundaries for defining one or more relevant markets.

# 2. Methodologies for Calculating LRMC

This chapter provides a brief overview of the concept of marginal cost before considering in greater detail the practical considerations associated with calculating the LRMC in an electricity generation market.

## 2.1. The concept of marginal cost

In our earlier report we describe in detail the concept of marginal cost, and the distinction between SRMC and LRMC.[4] To assist with our consideration of methods for calculating LRMC, in this section we reproduce our summary of these concepts.

Marginal cost is the added cost of producing a specified increment in output or, equally, the cost that is avoided by reducing production by a specified amount. Marginal cost can be estimated in either short run or in long run terms. The fundamental difference between SRMC and LRMC is the time frame under consideration and the implications of this for the extent to which a firm can adjust its production process.

- SRMC is the cost of an incremental change in demand, holding at least one factor of production – generally, capacity – constant; whereas
- LRMC relaxes this constraint and reflects the cost of an incremental change in demand assuming all factors of production can be varied.

An important distinguishing feature of SRMC is that, in the event existing capacity is insufficient to meet all demand, SRMC is represented by whatever price level is necessary to curtail demand to match available supply. It therefore takes account of the costs of shortages faced by customers, in addition to the direct costs of production.

By contrast, the estimation of LRMC accounts for the fact that, in the long run, firms have the option of expanding their capacity in order to meet increased demand. Measuring LRMC therefore involves estimating the costs associated with undertaking a capacity expansion sooner than would otherwise be the case in response to a change in demand.

Both SRMC and LRMC can fluctuate over time and there is no *a priori* reason to expect them to be equivalent at any particular moment. However, there is a strong 'in principle' link between SRMC and LRMC over the long term. In particular, when demand is growing over time, or subject to short term fluctuations, SRMC can be expected to increase to the point at which the expected cost of curtailing demand exceeds the cost of expanding capacity to *meet* that demand, ie, when LRMC < SRMC.

Of course, practicalities mean that the timing and size of capacity expansions will not always be perfect, eg, SRMC may rise above LRMC for a period if the optimal expansion is particularly lumpy, or occurs on slower than the ideal timing. Nonetheless, provided that the concepts are measured over a sufficiently long timeframe, the link between SRMC, LRMC

4 *Ibid*, in particular section 2 and appendix A.

and new investment decisions should mean that, on average, there is no material difference between the value of SRMC and LRMC.

## 2.2. Establishing LRMC in wholesale electricity markets

The key distinction between the concept of SRMC and LRMC is whether productive capacity is treated as fixed or is allowed to vary. In the context of a wholesale electricity market, the LRMC therefore includes the marginal cost of future capital that is required to provide sufficient generation capacity to meet an increase in demand. In other words, because future capital costs will vary according to forecast demand, these future capital costs should properly be included in the marginal costs.

It follows that if average wholesale market prices align with the LRMC over periods sufficiently long to capture capacity expansions, then appropriate market signals are being created about the need to expand supply capacity. By implication, wholesale market prices that are consistent with the LRMC over time will ensure that new generation investments receive sufficient revenues so as to recover both the operating and capital costs of the investment.

Where there is sufficient existing supply capacity, the SRMC should include all costs directly incurred as a consequence of generating electricity, including:

- the generation fuel costs;
- the costs of satisfying any carbon tax obligations; and
- the cost of marginal wear and tear to generation capacity resulting from a change in electricity dispatch.

Relaxing the assumption that existing supply capacity is sufficient to meet demand causes this concept to shift from one focused on the cost of supplying more (which, once at capacity, is no longer possible), to the cost of curtailing demand in order to ensure demand and supply is balanced. The amount that the marginal user would be willing to accept to reduce his or her demand will equal the value of electricity to that user.

In practice at any point in time and provided there is ample capacity to satisfy demand wholesale market prices should reflect the SRMC of the marginal unit of generation needed to satisfy demand. As demand approaches available capacity, wholesale market prices can be expected to increase to signal the risk of shortage and so begin to curtail demand so that supply and demand are balanced. This in turn creates signals for new generation investment. The market price cap[5] in combination with the cumulative price threshold[6] set limits on the extent to which market prices are able to reflect the demand curtailing element of SRMC and so can operate if set too low so as to compromise the incentives for new generation investment.

---

[5] The market price cap (MPC) is a limit on the dispatch price, and is currently set at $12.500/MWh.

[6] The cumulative price threshold (CPT) provides a mechanisms for reducing the dispatch price to the administered price cap (which is currently set at $300/MWh) if the sum of the half-hourly wholesale market spot prices over a rolling seven-day period exceeds the threshold. The CPT is currently set at $187,500/MWh.

The LRMC is estimated assuming that productive capacity is no longer fixed and should include:

- the long run marginal operating cost of meeting additional demand, namely:
  - the generation fuel costs;
  - the costs of satisfying any carbon tax obligations; and
  - the cost of marginal wear and tear to generation capacity resulting from a change in electricity dispatch, *and also*
- the long run marginal capital cost associated with providing sufficient capacity to meet the additional electricity demand over a future time horizon.

Future long run marginal operating costs therefore include those additional costs incurred as a result of an increase in demand given existing capacity, ie, marginal operating costs measured by reference to either existing or new capacity. The relevant operating costs are all those with a causal relationship to demand and associated with the generation of electricity.

The long run marginal capital costs include the increase in future capital costs that are caused by an increase in demand above that which can be supplied through existing capacity. The relevant capital costs to be assessed is the difference in the least cost combinations of the future capital costs of generation mix needed to satisfy a change in the future load profile.

## 2.3. Methods of calculating LRMC

In principle, marginal cost is simply the first derivative of the electricity generation cost function with respect to output. However, in practice the value of the LRMC is usually approximated by estimating how long run operating and future capital costs change if expected demand changes.

There are two broad methodologies that are used to estimate the capital cost component of the LRMC for a market, ie:

- the perturbation approach (also known as the ‘Turvey’ approach); and
- the average incremental cost approach (AIC).

These two approaches involve similar steps but differ in the precision with which they measure the effect of changes in demand on capital costs.

### 2.3.1. The perturbation approach

The perturbation approach to estimating how future capital costs vary as a consequence of an increment or decrement of demand can be summarised as follows:

1. forecast average annual and maximum demand as reflected by the anticipated load duration curve over a future time horizon of, say, 20 years;
2. develop a least cost program of generation capacity expansion that ensures that supply can satisfy demand, given the reliability standard;

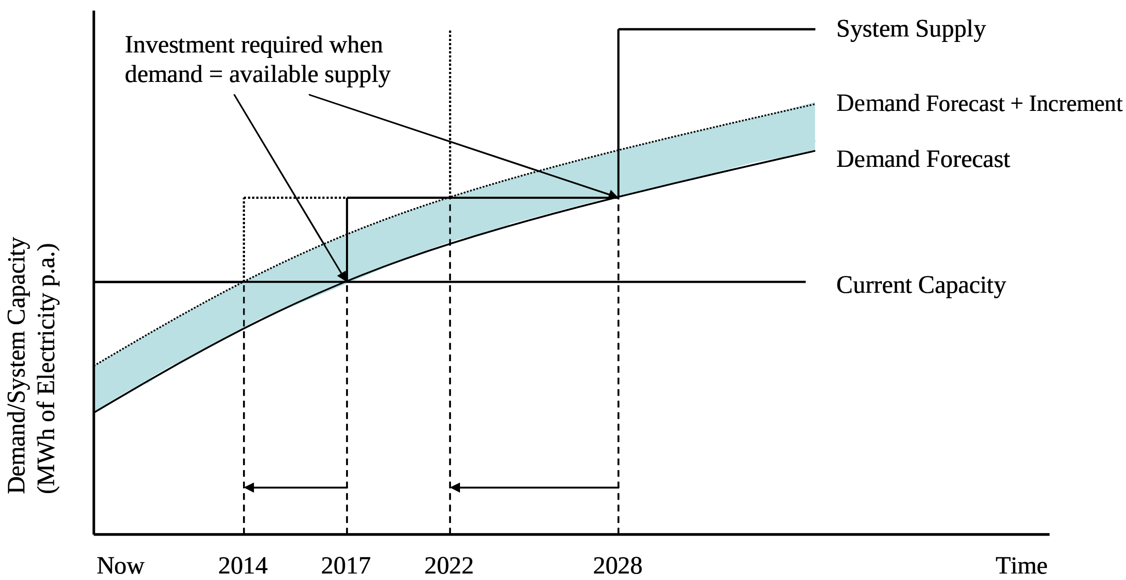
3. increase or decrease forecast average and/or peak demand by a small but permanent amount and recalculate the least cost generation capacity expansions needed to equate demand and supply;[7] and
4. calculate the long run marginal cost (LRMC) as the present value of the change in the least cost capital program plus the change in operating costs, divided by the present value of the revised demand forecast compared to the initial demand forecast.

Algebraically, the perturbation approach to estimating LRMC can be expressed as follows:

$$\text{LRMC} = \frac{\text{PV(revised optimal capex plus opex} - \text{optimal capex plus opex)}}{\text{PV(revised demand} - \text{initial demand)}}$$

The perturbation approach to estimating LRMC is illustrated in Figure 2.1 below.

**Figure 2.1: Perturbation Approach to Estimating LRMC**




By taking the present value of future costs using a discount rate equal to the time value of money, converts the stream of future costs to a single value in today's dollars. A similar calculation is undertaken for the future stream of the change in demand, which acknowledges that the value of consumption in the future is less than the value today. This ensures that the discounted costs (ie the top line) are divided by a time equivalent measure of demand.

In short, the perturbation approach considers the effect of a permanent increment (or decrement) to the base demand forecast upon future capital expenditure for new generation capacity plus the effect on future operating expenditure estimated by reference to either new or existing generation. This acknowledges that the increment of demand can be satisfied by both expanding the output of existing generation capacity and so incurring additional

7 It is important to consider how increments or decrements in both peak and average demand influence the future capacity plan, since these could result in a different combination of generation plant investments to satisfy demand at least cost.

operating costs, and/or by expanding total generation capacity, which incurs both additional capital and operating costs.

The solid stepped line above represents a series of projected increases to system capacity, optimised in terms of their order and timing so as to meet future demand at least present cost. The dashed stepped line represents the same projected increases to capacity, but brought forward as required to meet the forecast demand plus some assumed permanent increment. As noted above, LRMC is calculated as the change in the present value of capital plus operating expenditure required to maintain the supply demand balance divided by the present value of the marginal change in expected demand. The latter is represented by the shaded area in figure 2.1, between the difference between the demand forecast with and without the increment.

The perturbation approach is consistent with the earlier discussion explaining the concept of LRMC because it directly estimates the change in future costs (both the operating costs of existing generation plus the capital and operating costs of an increase in generation capacity needed to supply the increment in demand). Importantly, it estimates LRMC by considering the *difference* in the present value of incremental future system costs as a consequence of a permanent increment of demand.

The main features of the perturbation approach are:

- it closely approximates long run marginal cost because its focus is on the change in costs necessary to respond to a specified change in demand;
- it is forward looking, since it is based on anticipated capital investments necessary to balance supply and demand; and
- it incorporates only those costs, and all costs, that are caused by demand growth above existing capacity.

It should be noted that LRMC estimates based on the perturbation approach may be influenced significantly by the size of the increments or decrements in demand used in the calculation.[8] It is important to analyse the sensitivity of the resulting estimate to variations in the size of the hypothesised increment or decrement.

### 2.3.2. The average incremental cost approach

The average incremental cost (AIC) approach shares many of the same steps as the perturbation approach but involves an important simplification in assessing the effect of changes in demand on future costs. The AIC approach to estimating LRMC can be summarised as follows:

1. forecast average annual and maximum demand over a future time horizon of say 20 years;
2. develop a least cost program of generation capacity expansion that ensure that supply can satisfy demand, given the reliability standard;

[8] Our proposed approach to this matter involves investigating the sensitivity of the estimate to different increment scenarios. We explain this further in section 3.3.3.

3. estimate LRMC as the present value of the expected costs of the optimal strategy divided by the present value of the additional demand supplied (assuming the supply demand balance is maintained).

Under the AIC approach, the estimate of LRMC can be represented as:

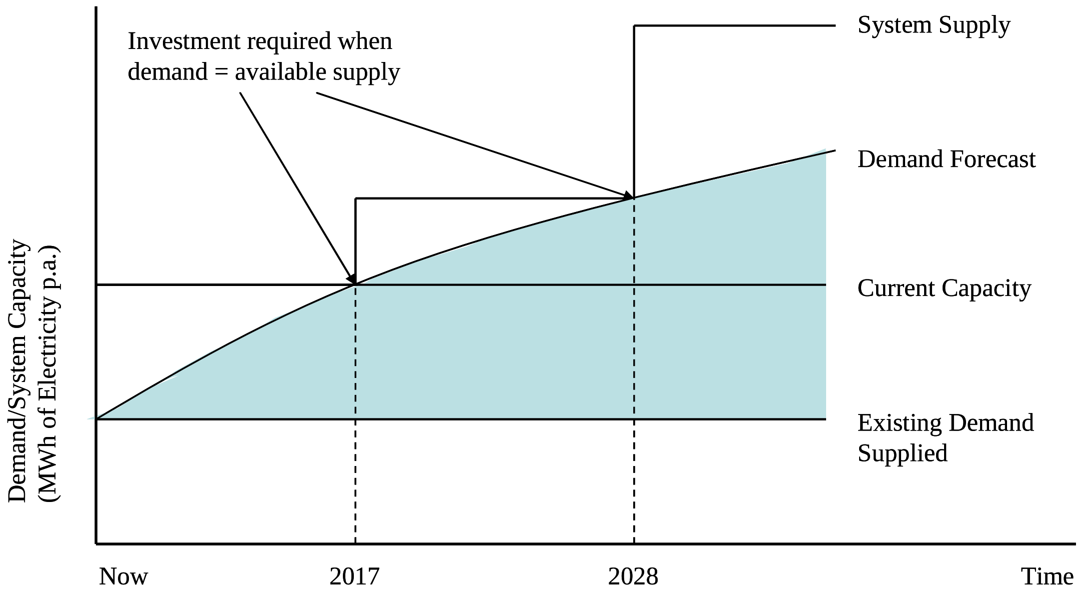
$$\text{LRMC} = \frac{\text{PV(new generation capacity + marginal operating costs)}}{\text{PV(additional demand served)}}$$

New generation capacity refers to the capital costs of new generation investments undertaken to satisfy forecast average annual and maximum demand. Marginal operating costs refer to the additional operating costs of both existing and new generation capacity required to satisfy forecast demand.

Importantly, under the AIC approach the term 'additional demand served' refers to demand over and above that which is *currently being supplied* rather than that which *could be supplied*, ie, above and beyond what can be supplied with existing capacity. This is because in the long-run it is assumed that *all* factors of production, including existing capacity, are able to be varied. This distinction is particularly important when calculating the LRMC of a system that is not currently capacity constrained.

A simplified version of the AIC approach to estimating LRMC is illustrated in Figure 2.2 below.

**Figure 2.2: AIC Approach to Estimating LRMC**




In this example, the LRMC would be calculated as the present value of the expenditure associated with the optimal capital program plus the marginal operating costs divided by the present value of the change in demand supplied, as shown by the shaded area above. Importantly, this measure of demand is *not* the same as the generation capacity associated with the capital projects. This is because the concept of LRMC is seeking to estimate the marginal cost of a change in output caused by a change in demand by customers. It would make no sense to simply divide cost by capacity in the years it is available since that would not represent the marginal cost per unit of future consumption.

The term 'marginal operating cost' also refers to the operating costs of supplying demand over and above that which is *currently* being supplied and not the operating costs of supplying demand over and above that which *could be* supplied. Therefore, marginal operating costs can be divided into the following two broad categories:

- the marginal operating costs of meeting any increases in demand with existing capacity (these costs will be relatively material in the case where the next capacity investment is some time in the future); and
- the marginal operating costs associated with operating new capital to meet the projected increase in demand.

The principal shortcoming of the AIC approach is that it uses average future capital costs to approximate the likely marginal costs associated with a change in demand. Put another way, the AIC does not discriminate across the 'size' of increments to capacity, ie, ignoring the time value of money each unit of new investment is treated equally in their ability to match supply and demand.

## 2.4. Summary

In our opinion the perturbation approach should be preferred over an AIC approach because it most closely aligns with the principles underpinning the concept of LRMC.

The AIC approach essentially calculates the amount needed to ensure that the total incremental costs of new generation capital expenditure and both new and existing generation operating costs are recovered to satisfied future demand. In contrast the perturbation approach focuses on how future costs change as a consequence of a permanent change in demand, and so is closer to the marginal cost concept.

In practice the AIC approach would be expected to generate a smoother estimate of the LRMC over time. As a consequence the resultant AIC estimate would likely be lower than an estimate using the perturbation approach when there is insufficient existing generation capacity to satisfy a permanent increment in demand over the near future, and vice versa.

# 3. Application to Wholesale Electricity Markets

In this section we explain how the methodologies described in section 2 can be practically applied to estimate LRMC for the supply of wholesale electricity. We focus initially on explaining the complexities of determining the least cost programme of generation investments, before describing the modelling tools that can be used to estimate LRMC consistent with the methodologies outlined in section two.

## 3.1. Developing a least cost programme of generation investments

A common element in all of the methodologies for estimating LRMC is the development of a least cost programme of generation investments to satisfy forecast demand. However, the methodologies differ on the way these costs are translated to estimate LRMC.

For electricity wholesale markets, the least cost combination of generation to satisfy forecast demand involves a combination of generation types. This is because:

- electricity demand varies considerably during both a day and over a year; and
- some electricity generation technologies have high upfront construction costs with lower operating costs, while others have lower upfront costs and higher operating costs.

This means that the *average cost* of generation as a function of the hours of generator running during a year[9] can vary considerably between generation types. As a consequence, the least cost combination of generation plants capable of supplying any given demand requires a consideration of *both* the costs of the electricity generation types capable of meeting the relevant demand, and the particular load profile of that market.

To explain these concepts, consider the following simplified example. Assume that there are three plant types capable of servicing an increment in demand in an electricity wholesale market. These plant types have the cost characteristics set out in Table 3.1.

**Table 3.1: Example Plant Type Cost Characteristics**

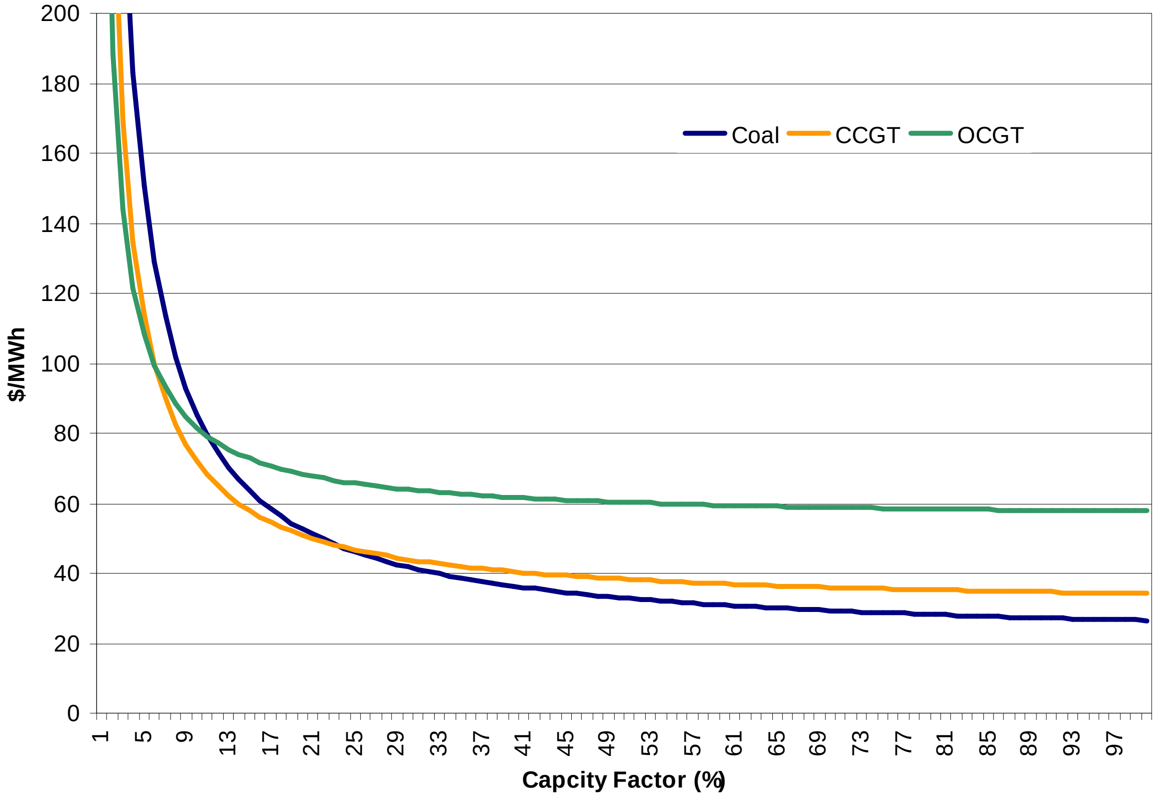
| Plant Type | Variable Operating Costs[10] ($/MWh) | Fixed annualised capital costs ($/kW) |
|---|---|---|
| Thermal Coal | 20 | 2,000 |
| Combined Cycle Gas Turbine | 30 | 1,100 |
| Open Cycle Gas Turbine | 55 | 700 |

9 This is usually expressed as the capacity factor, which is a percentage calculated by dividing the hours of running by the total hours in a year (ie, 8760 hours).

10 This includes both plant variable operating and maintenance costs, plus fuel related costs.

The average generation costs as a function of the hours of generation can be represented in Figure 3.1.

**Figure 3.1: Average Generation Costs as a Function of Capacity Factor**

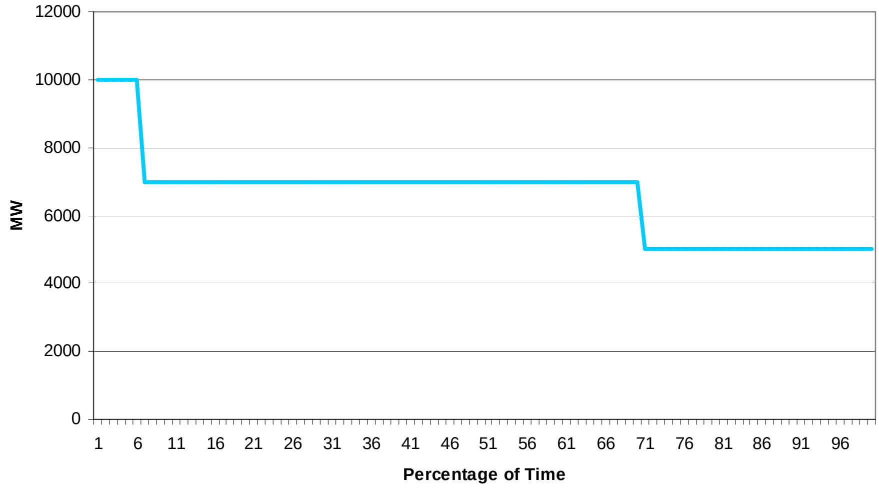



In this example, the average costs by capacity factor of the generation type mean that

- coal has the lowest average cost at capacity factors greater than 23 per cent;
- CCGT between 6 and 23 per cent; and
- OCGT for capacity factors less than 6 per cent.

Now consider a simple annual load duration curve (showing the percentage of time total demand is above each megawatt level over a year) as represented in Figure 3.2.

**Figure 3.2: Example Load Duration Curve**




The least cost combination of generation that could serve this load profile given the assumed prices is therefore:

- 5000 MW of coal;
- 2000 MW of CCGT; and
- 3000 MW of OCGT.

This example demonstrates that, given the load profile of demand, the combination of generation that will satisfy changes in load over time will most likely reflect a combination of generation types. To determine the generation investment profile therefore involves considering the least cost additional generation capacity needed to satisfy forecast increases in the load profile of demand over a specified time horizon. Importantly, any anticipated changes in load profile (even in the absence of growth in anytime average demand), can have a significant impact on the least cost combination of generation to satisfy demand, and so have an impact on associated costs.

The next section describes the modelling tools that can be employed to undertake this analysis.

## 3.2. Modelling tools to generate least cost investment profiles

There are a number of modelling tools that can be employed to estimate the least cost programme of generation investment needed to satisfy forecast electricity demand, as an important input for subsequently estimating the LRMC of meeting forward demand. They differ in terms of the extent to which market specific factors are taken into account in the analysis, namely:

- using market models to estimate the change in costs associated with shifts to a forward looking generation investment profile to satisfy an increment of demand;

- using a simplified model to estimate the least cost combination of new generation investment required to satisfy an increment in demand; and
- using estimates of the costs for investing in specific generation types, combined with a demand profile to calculate the combination of investment needed to satisfy an increment in demand.

The remainder of this section explains how each tool can be applied to estimate the LRMC for the generation market.

### 3.2.1. NEM market modelling

The market modelling approach involves the use of a market model to develop a least cost combination of generation investment to satisfy a forecast of future average annual demand and maximum demand, given assumptions about the load profile. This approach allows for:

- a consideration of generation costs over time by analysing the opportunity to trade off augmentation by technology A, which has lower costs over time, even though technology B might have an initial lower capital cost;
- explicit consideration of regional interconnections and so takes into account the implications of generation investments in other regions on the need for new generation capacity in the region of interest, given interconnector constraints;
- minimum generation investment capacities to be explicitly taken into account;
- the NEM reliability standard to be explicitly taken into account in the investment profile; and
- explicit consideration of policies that influence generation investment decision making, including the introduction of a price on carbon and the large-scale renewable energy target.

Having developed the future investment profile, a new profile is developed in light of a hypothesised increment being applied to the forecast future annual average and maximum demand assumptions. This gives rise to a second profile of generation investment to reflect the change in demand.

The long run marginal cost in the base year (ie, the year that is being investigated) is then calculated as the present value of the difference in the costs of satisfying the generation investment profile divided by the present value of the increment in demand. Algebraically:

$$\mathrm{LRMC_{year}} = \frac{\sum_{t=1}^{T} \frac{\mathrm{Gen2_t} - Gen1_t}{(1+\mathrm{i})^t} + \sum_{t=1}^{T} \frac{\mathrm{Opex2_t} - Opex1_t}{(1+\mathrm{i})^t}}{\sum_{t=1}^{T} \frac{Increment_{\mathrm{t}}}{(1+\mathrm{i})^t}}$$

Where:

- $\mathrm{LRMC_{year}}$ is the long run marginal cost for the relevant year;

- $Gen1_t$ is the optimal generation new investment costs to satisfy forecast demand, for each year t;
- $Gen2_t$ is the optimal generation new investment costs to satisfy forecast demand following a permanent increment or decrement in demand, for each year t;
- $Opex1_t$ is the additional operating costs of the optimal programme of generation investment, for each year t;
- $Opex2_t$ is the additional operating costs of the optimal programme of generation investment following a permanent increment or decrement in demand, for each year t
- Increment is the permanent increment in demand that is applied to each year t;
- i is the discount rate; and
- T is the total number of forecast years.

This detailed market modelling approach amounts to the application of a perturbation approach to the estimation of LRMC of meeting the relevant demand, taking into account all of the factors that are likely to influence investment decision making.[11] Importantly, it allows for market rules that may affect the pattern of investment (such as the large scale renewable energy target) to be explicitly incorporated into the determination of the least cost generation investment profile. It follows that the resultant LRMC estimates will have had these influences taken into account in the estimation process.

This approach has the advantage of taking into account current generation capacity relative to forecast demand and so allows for the estimate of LRMC to be lower during periods where there is existing excess capacity, and higher when capacity is close to being constrained. This approach is likely to resemble most closely the price signals created in energy only wholesale electricity markets for new generation investment.

### 3.2.2. A simplified modelling approach

An alternative, simpler approach to approximating LRMC that does not require full market modelling involves using information on new entrant technology costs to calculate the least cost combination of generation capacity to satisfy a load duration curve for a given year. This approach bears many of the characteristics of the average incremental cost approach methodology described in section 2, with additional simplifying assumptions that existing capacity is already optimal, and that future demand grows at a constant rate and with no change to the load profile.

This approach involves:

- calculating the least cost combination of new entrant generation capacity to satisfy electricity demand within a given year;

---

[11] Market models can also be used to determine LRMC using an average incremental cost methodology, as described in section 2. However, given that a perturbation approach is superior to an average incremental cost approach, if market modelling was to be used the proposed perturbation approach is to be preferred.

- applying an increment to the electricity demand and recalculating the least cost combination of new entrant generation capacity to satisfy the new load profile within a given year; and
- estimating the LRMC as the difference in the costs of new entrant generation divided by the increment in electricity demand.

Algebraically this can be expressed as:

$$\text{LRMC}_{year} = \frac{Gen1_{year} - Gen2_{year}}{Increment}$$

Where:

- Gen1year is the total cost of generation to satisfy demand in the relevant year;
- Gen2year is the total cost of generation to satisfy demand in the relevant year following an increment in demand;
- Increment is the permanent increment in demand that is applied to the load profile; and
- year, is the year for which the LRMC is being estimated.

This approach is computationally simpler to the detailed market modelling approach, but has a number of deficiencies, namely:

- it presumes that generation investment is completely divisible so that demand can be optimally satisfied;
- it approximates an optimal, existing investment profile by assuming that new entrant generation has been constructed and is available to satisfy known demand within the period with certainty; and
- it does not take into account expected future growth in demand and the particular way in which changes in demand relative to existing capacity may influence LRMC.

In addition, applying this approach to a NEM region means that the contribution of supply from an interconnecting region is not explicitly taken into account in the estimate of the total cost of generation.

This approach most closely resembles an average incremental cost approach to estimating LRMC. As a consequence it tends to smooth out the estimates of LRMC for any year, because it does not explicitly take into account the future profile of generation investment needed to satisfy the growth in demand relative to existing generation capacity.

### 3.2.3. Simplified calculations to determine expected market prices

The final methodology is often misconstrued as an approach for estimating the LRMC when, in fact it is better described as a calculation of what might be a reasonable level of prices for an existing generation portfolio given assumptions about the underlying costs of the relevant generation technologies.

It involves:

- calculating the annual generation costs for each generation type currently operating in the market, including both variable and annualised capital costs, based on new entrant or present day costs;
- using actual dispatch of existing generation and so its the average cost given the implied capacity factor to determine the total cost that would need to be recovered to fund those generators given new entrant costs; and
- summing the total cost across all existing generators and dividing by the total load in that year.

This approach purports to determine what would have been the prices necessary to ensure that existing generators were able to recover their own costs. However, because this approach does not seek to calculate how future returns are affected by the balance between demand and available capacity, and changing costs in the market it cannot be properly described as an estimate of the LRMC.

## 3.3. Modelling assumption and data considerations

Estimating the relevant historic LRMC requires assumptions to be made about new entrant costs, and the profile of load. In addition, consideration needs to be given to the relevant size of the increment to use to estimate the associated change in future capital and operating costs.

### 3.3.1. New entrant costs

There are two approaches to estimating new entrant costs for the time horizon of the study. The first involves the use of forecast new entrant costs *that were made at the time for which the estimate is to be made.* In other words, the first approach would involve the use of forecast new entrant costs for each year in the future from the year under consideration.

The second approach involves using actual historic construction and operating costs of new generation plants for each future year, from the year being investigated. The difficulty with this second approach is obtaining consistent data on actual construction costs of new generation investments since these are not readily published.

In our opinion, the most relevant approach is to use forecasts of new entrant costs that were made at the time, since these are likely to reflect the best market information at the time about future generation construction costs. This approach also has the advantage of ensuring that consistent estimates of new entrant costs are used for each year being investigated so that relativities between different technologies will be internally consistent.

That said we also believe there is merit in considering the sensitivity of the estimates of LRMC to changes in the assumptions underpinning the new entrant costs. This can include assumptions for the discount rate, assumed plant construction costs, and fuel prices.

### 3.3.2. Electricity demand

To develop a future profile of generation requires consideration of future energy demands. As with new entrant costs, there are two demand assumptions that could be used, namely:

- electricity demand forecasts that were made at the time we are investigating, as published by the AEMO (or the National Electricity Market Management Company (NEMMCO), as applicable); and
- actual electricity demand over the relevant period, combined with the most up to date electricity forecasts for any future years.

In our opinion, as with new entrant costs it is appropriate to use forecasts of electricity demand that were made at the time for which the estimate is to be made. This is because investment decisions, which influence observed market prices, are based in part on demand forecasts applicable at the time rather than actual demands. In practice this would mean, for example, using forecasts published prior to 2008 to estimate the LRMC applicable to 2008.

The AEMO publishes demand forecasts at both the 10 per cent and 50 per cent probability of exceedance (10POE and 50 POE respectively). In market modelling exercises we recognise that capacity expansion and compliance with the NEM unserved energy standard is primarily linked to the 10POE forecast and so the level of investment in peaking plant, but the 50POE forecast is the dominant factor in assessing energy dispatch and returns to other plant. Accordingly we combine analysis of prices and revenues from both the 10POE and 50POE demands to assess profitable entry of plant at different levels of annual utilisation and compliance with the reliability standard.

### 3.3.3. Size of the load increment

Finally, estimating LRMC requires the consideration of a new generation profile following a permanent increment or decrement in electricity demand. The size of the increment can itself affect the estimate of the LRMC, and this influence can vary depending on the balance between existing generation capacity and demand in the year in question.

The size of the increment should be sufficiently large so as to influence the generation profile. We anticipate that a permanent increase in demand of between 1 and 5 per cent would influence the profile of generation investment and so would be sufficient for estimating the LRMC. That said it will be important to understand the sensitivity of LRMC estimates to differences in the size of the increment, at least within this range.

### 3.3.4. Choice of discount rate

The choice of discount rate can affect the estimates of LRMC directly through the discounting of incremental capital and operating expenditures, and if a market modelling approach is used, by affecting the estimates of new investment costs that would need to be recovered through market prices in each year. The influence of the discount rate on the LRMC will vary according to the anticipate profile of investment.

As a matter of principle, the discount rate applied to the LRMC calculation should be the same used in any modelling undertaken to generate the profile of generation investments required to satisfy a forecast demand profile. To determine the appropriate discount rate will therefore require consideration of the appropriate weighted average cost of capital (WACC) to apply to electricity generation investments in the NEM.

As with the other modelling assumptions, it will be important to understand the sensitivity of LRMC estimates to changes in the WACC.

## 3.4. Factors influencing estimates of the LRMC

It is important to acknowledge that the two principal modelling approaches discussed above seek to estimate the LRMC by understanding how future demand affects forward looking capital and operating costs. It therefore represents those costs that might be avoided or incurred as a consequence of a marginal change in demand.

However, no single modelling approach or scenario can reflect all market circumstances that bear on the estimation of LRMC, for example, because of:

- the extent that market factors may or may not be taken into account in the modelling eg, minimum reserve levels, minimum capacity of generation investment etc; and
- market uncertainties, including in forecast demand and the cost of new entrant generation.

As a consequence of these uncertainties there is likely to be merit in applying multiple methodologies and conducting sensitivities around key assumptions to ascertain the likely range of LRMC. This approach will improve the confidence in a LRMC estimate.

## 3.5. Summary

The cost characteristics of the optimal portfolio of electricity generation to satisfy a particular load profile mean that the least cost approach to satisfying future growth in demand is generally a combination of generation types. As a consequence, modelling techniques need to be employed in order to obtain an estimate of LRMC for meeting a specified market demand.

In these circumstances, the extent of complexity implied by market modelling means that there may be merit in undertaking a simplified modelling approach based on an average incremental cost methodology. This provides an opportunity to obtain rough estimates of the LRMC as a starting point for the analysis of market power.

In our opinion, a perturbation approach to estimating LRMC should be preferred given it most closely aligns with the concept of LRMC. However, a perturbation approach will need to make use of complex market modelling, which can be time consuming when considering multiple years. This may mean that a combination of approaches is used, with less sophisticated methods employed to develop an understanding of the possible range of LRMC, combined with fewer more intensive analyses using market modelling for periods where a more robust estimate is required.

# 4. Estimating Average Market Prices

This section describes the different possible approaches to estimating average market prices, before describing the factors that can influence observed market prices.

## 4.1. Calculation of NEM spot prices

The AEMO has the responsibility for balancing electricity supply and demand through a centrally coordinated real time dispatch process. The process for determining the NEM spot price involves:

- registered scheduled generators submitting bids for particular quantities of electricity two days ahead of the supply requirement for each of the 48 trading intervals in that day - the bids specify the quantities that each generator is willing to supply at particular prices that it nominates;
- the AEMO publishing a pre-dispatch schedule that sets out supply and projected demand for all trading intervals (30 minute periods) over the following (approximately) two days;
- in principle the bids are ordered from lowest to highest (known as the bid stack), and are used by the AEMO to dispatch generators every five minutes (more precisely the AEMO uses a mathematical optimisation algorithm to identify the lowest cost means to satisfy demand across the NEM from scheduled generators in the different parts of the NEM, and to supply operating reserves, while remaining within the capacity of the transmission network);
- a dispatch price (ie, the five minute price) is calculated for each pricing region of the NEM by reference to the bid submitted by the most expensive generator (which may be in another region) that is required to be dispatched into production to satisfy demand in that region;
- the spot price is calculated every half hour by averaging the six dispatch prices that occurred during that interval; and
- dispatch prices and spot prices in different regions will differ only by the effect of transmission losses unless transmission interconnections between price regions are scheduled to operate at their safe operating limits.

The National Electricity Rules allow for three types of bids, each of which are subject to a floor price of -$1,000 and a ceiling price of $12,500 per megawatt hour.[12] These bid types are:

- **daily bids** which are submitted the day before supply is required and are incorporated into the forecasts that are prepared prior to dispatch;

---

12 This is also known as the market price cap and is the price set in the market if any customer load is curtailed in the event of a supply shortfall. The market price cap is reviewed on a periodic basis and is determined by the Reliability Panel. The market price cap is set at a high enough level to facilitate entry of very low capacity factor peaking suppliers and also to encourage participants to manage the resultant financial risk by entering into financial contracts. The price of these contracts is then intended to act as a signal for new investment.

- **re-bids** can be submitted up to approximately five minutes before dispatch and allow the generator to alter the quantity of electricity it will supply within a particular price band, however generators are not permitted to change the prices of the bands of capacity offered in the daily bid; and
- **default bids** are those that stand when no daily bid has been made.

## 4.2. Methodology for calculating average market prices

There are two principal approaches for averaging NEM spot prices, namely:

- time weighted prices found by taking an arithmetic average of the spot prices over the relevant time period; or
- volume weighted prices found by taking a weighted average of the spot prices over the relevant time period, weighted by the volumes dispatched.

The arithmetic average involves summing the regional reference prices over the relevant time period, and dividing by the number of time periods. Algebraically:

$$\overline{RRP}_{arithmetic} = \frac{\sum_{i=1}^{n} RRP_i}{n}$$

Where:

- $\overline{RRP}_{arithmetic}$ is the arithmetic mean of the regional reference node price;
- $RRP_i$ is the regional reference node price in settlement period i; and
- $n$ is the number of settlement periods that are being averaged.

The weighted averaging of spot prices involves summing over each trading period the volume of energy dispatched multiplied by the regional reference price, and then dividing by the sum of dispatch over the relevant period. Algebraically:

$$\overline{RRP}_{volume} = \frac{\sum_{i=1}^{n} l_i RRP_i}{\sum_{i=1}^{n} l_i}$$

Where:

- $\overline{RRP}_{volume}$ is the load weighted mean of the regional reference node price;
- $l_i$ is the load dispatched in settlement period $i$;
- $RRP_i$ is the regional reference node price in settlement period $i$; and
- $n$ is the number of settlement periods that are being averaged.

The weighted average and the arithmetic average are equal where the weights (in this case, load) are equal in each period, reflecting that all data points contribute equally to the average. It follows that a weighted average is used in circumstances where it is appropriate for some data points to contribute more to the average than other data points.

In our opinion, a volume weighted average of the regional reference price is the more appropriate averaging methodology for the purpose of comparing market prices with estimates of LRMC. This is because the ultimate question to be addressed is the scope for generators to exercise market power and so to raise prices above the level implied by effective competition.

## 4.3. Consideration of contract prices

We discussed in our earlier report that it is conceivable there may be periods during which substantial market power is being exercised, but which the average spot price is unaffected. As a consequence, to the extent that information is available, it might be useful to gather data on electricity contract prices as an additional consideration.

This would allow for a comparison of the total market revenues from contract and spot market payments. If contract prices are simply a reflection of expected spot prices then over time the net revenue from contract and spot would differ from spot only by a risk premium. However, if market power is being exercised in the contract market the difference would be greater.

Contracts are purchased either Over-The-Counter (OTC) or via the Sydney Futures Exchange (SFE). Since 2002-03 the volumes traded via the SFE has increased considerably, such that in 2008-09 volumes traded through the SFE represented 153 per cent of NEM demand, while OTC volumes were estimated to be 105 per cent of NEM demand. This suggests that SFE contract prices are likely to be broadly indicative of NEM contract prices, although in the absence of information on OTC contracts or direct bilateral contracts this can only be taken as a strong hypotheses. We note that, if the two trading venues gave rise to systematically different prices, arbitrage opportunities could be expected to close any such gap.

We also note that the National Power Index published by d-cyphaTrade provides an indication of base electricity futures listed on the SFE. As a consequence, it provides an indication of possible movements in generational contract prices in the NEM.

Wholesale market contracts protect retailers from the effects of unexpectedly high prices, for example in an unusually hot summer season, and similarly insure generators from unusually low prices. But high spot prices that are inconsistent with prevailing conditions may then flow through to subsequent contract prices creating a lag effect that would need careful consideration over a number of years. It should also be recognised that while the spot price is primarily determined by generator decisions for a given market condition, contract prices and volumes are clearly affected by decisions of retailers and customers about the level of risk management they are prepared or able to buy. Contract analysis is therefore more complex than spot analysis.

To undertake an analysis that accounts for contracts, we would need information on contract prices, the percentage of output generators have under contract, and the timeframe for contracting (most likely three to four years at a minimum). In addition, we would need to consider the competitive level for the risk premium between spot and contract prices.

While it might be possible to develop a market price that incorporates both spot prices and contract prices under assumptions of an appropriate hedging strategy, we are not proposing to undertake such an analysis initially, because of the complexity and range of assumptions that would be required. However, we will undertake a limited filtering test of the outcomes from the spot analysis against readily available contract prices to ascertain if the observed differences are competitively plausible. That is we will test if there is a case for further examination.

## 4.4. Factors influencing observed market prices

Observed market prices reflect the outworking of the price setting process in order to match electricity supply with demand in each five minute dispatch period. These prices may be influenced by a number of short term factors that can lead to average annual prices deviating from estimates of LRMC. Additionally, the market modelling methods that can be used to estimate LRMC are themselves necessarily a simplification of the factors influencing market prices, which can cause the resultant estimates deviating from average annual prices.

The factors that influence observed market prices include (amongst others):

- unanticipated changes in load or supply due to generation or network failures or unexpected extended periods of drought, or hot or cold weather;
- localised dispatch requirements to satisfy loads given network constraints; and
- operational factors for generators.

The remainder of this section explains each of these influences in greater detail.

### 4.4.1. Unanticipated changes in load or supply

The market price setting arrangements are geared to ensuring that supply and demand is met in each and every five minute period. However, there are a number of circumstances that can result in five minute prices fluctuating significantly, including:

- a sudden transmission network outage;
- unanticipated periods of hot weather, which can result in prices remaining higher than anticipated to ensure the short-term balancing of supply and demand; and
- generator failure resulting in plants being unavailable to supply the market.

These circumstances can have a significant influence on spot prices quite separate from the underlying, longer term investment considerations.

Consider for example a circumstance where the dispatch price for one five minute period is $40/MWh and an unanticipated generator breakdown occurs. The unexpected nature of the breakdown can result in the next five minute dispatch price spiking to, say $5000/MWh. This might therefore result in a 30 minute settlement price being say, $1000/MWh. Assuming that the fundamental cost characteristics of available generation plants for that 30 minute period given the merit order means that the new settlement price should have been say $50/MWh, this highlights how generator breakdowns (and, equally, transmission outages) can have a significant influence on spot prices.

### 4.4.2. Localised dispatch requirements

The five minute generation dispatch orders provided by the AEMO are based on a combination of the bid prices and quantities, and network constraints. In operating the network the AEMO needs to ensure that demand and supply is matched in all parts of the network. The geographic dispersion of both generators and loads, and the associated inter- and intra-regional network constraints means that, at times, the lowest cost combination of plant will require the AEMO to limit the dispatch of the lowest cost generators upstream of a constraint and so dispatch higher price generators downstream or closer to customer load.

In addition, the AEMO must amend network constraint parameters to reflect periods when parts of the network are out of service for maintenance purposes. In scheduling network maintenance, the aim is to optimise the resultant impact on market prices (and there are incentives in the transmission regulatory arrangements in this regard). However, unexpected additional network or generation outages can compound the resultant impact on spot prices.

Dispatch prices, and therefore spot prices in the NEM reflect the marginal price for supply to the regional reference node, which can be thought of as the electrical centre of gravity of the region, inclusive of the effect of network and other operating limitations. A generator that is electrically constrained from the reference node in its region due to network limitations can therefore be exposed to spot prices higher than the price it has offered to the market but at which it is unable to be dispatched.

### 4.4.3. Operational factors for generators

For coal and CCGT plants it is not technically possible simply to start up or shut down the plant in short periods. When combined with scheduled maintenance on turbines this means that not all generation capacity will be available at all times. As a consequence there can be circumstances where a simplistic assessment of the theoretical least cost combination of generation cannot be dispatched due to either the availability of plants, or the need to incur start up costs. This can impact on the resultant observed market prices. On the other hand generators only receive the spot price and must ensure they submit prices that cover their start-up and other such costs from the spot price for energy for each start up in the periods in which they expect to be dispatched. This complicates the preparation and review of prices in bids by generators.

## 4.5. Summary

In our opinion a volume weighted average of market spot prices is the appropriate methodology for calculating average market prices for the purposes of analysing generator market power. That said, in undertaking an analysis of these spot prices it will be important to consider the many factors that can influence average spot prices in addition to the potential exercise of market power.

In addition, a filtering test of the outcomes inclusive of contracting will also be compared to estimates of LRMC, to provide further insights on possible generator market power.

# 5. Practical Application of the SSNIP Test

Our earlier paper described the SSNIP test in the following terms:[13]

> '…defining the boundaries of a market can be interpreted as establishing the smallest area of product, functional and geographic space within which a hypothetical profit maximising monopolist could successfully impose a small but significant and non-transitory increase in price (a 'SSNIP'). The establishment of market boundaries should start by considering the product, geographic and functional areas of supply by the firm whose conduct is in question. One then asks whether a hypothetical monopolist could profitably impose a SSNIP on those products, usually of between 5 and 10 per cent above the price level that would apply under conditions of workable competition, and assuming that the price of all other products remain constant.'

Importantly, the SSNIP test should not be used mechanically to determine the boundaries of the relevant market. Rather, it provides information to inform the decision on the relevant market given the surrounding circumstances of the matter being considered.

This section describes a methodology by which the SSNIP framework might be applied to aid in the determination of the geographic boundaries of one or more relevant markets for the purposes of considering the MEU's rule change proposal.

## 5.1. Applying the SSNIP Framework

Application of the SSNIP test for the purpose of determining the boundaries of relevant markets for the analysis of competition issues involves the postulation of a 5 or 10 per cent price imposed by a hypothetical monopolist over the narrowest possible formulation of the market. That formulation involves specifications as to the product itself, the function level of the supply chain of which it is produced, the geographic area over which it is bought and sold, and the time period. If such a price rise were to bring increased profits to the hypothetical monopolist, then the boundaries of the market have been correctly defined. This is because the ability to increase profits indicates that there are no sufficiently close substitutes to cause the price rise to be 'defeated' through either:

- customers switching to alternatives; or
- suppliers from outside the boundaries of the hypothesised market switching their production towards sales to customers within that postulated market.

To be clear, the 'defeat' of a SSNIP involves the hypothetical monopolist losing a sufficient quantity of sales so that the price rise is not profitable. If the SSNIP is unprofitable, then the substitutes that defeated it (whether in the form of a slightly different product, produced in the context of a greater degree of vertical integration, or from a wider geographic area) should be included in the market. This process is then repeated until all dimensions of the relevant market are established.

---

13 Page 34, Green, H., Houston, G., and Kemp, A., (2011), 'Potential Generator Market Power in the NEM', *A Report for the AEMC*, NERA Economic Consulting, June.

Some aspects of the test warrant further emphasis, because they may have an important bearing on how the test is conducted, ie:

- the starting point for the hypothesised price rise is the *competitive price level*, in order to avoid the ‘cellophane fallacy’ of including dimensions of the market that are only close substitutes because the process of testing for substitutes commenced with a monopoly price;[14]
- since the test focuses on the change in profits, the effect of lost sales must take into account the gross margin that would otherwise have been earned on those sales – this means that the loss of relatively low margin sales will not have much effect on profits, while the vice versa applies for the loss of high margin sales; and
- the postulation of a 10 as opposed to a 5 per cent price rise does not necessarily have a material effect on the outcome of the exercise, since a greater price rise will give rise to more lost sales, but will also deliver a larger profit gain on the remaining units sold, ie, a 10 per cent SSNIP is a ‘more exacting’ test, but rather an exercise in testing for any non-linearity’s in the field of potential substitutes.

Of course, the additional challenge in this particular instance is to develop a methodology that is capable of practical implementation in a context in which a new price is established every 30 minutes. We set out below a series of steps for applying the SSNIP framework in the context of wholesale electricity markets that, taken together, are computationally tractable, and in line with the general principles described above. The key steps are as follows:

- begin with a single NEM region (the narrowest practicable definition of the geographic market[15]) and assume that a single firm owns all of the scheduled generation units located in that region, ie, a hypothetical monopolist operates in the NEM region;
- calculate the optimal dispatch for each half hour in the one or more years in which the SSNIP is to be applied, assuming that plant is offered to the market at either its short run marginal costs (the base case) or its short run marginal cost plus 5 and 10 per cent and taking account of changes in interconnector flows as a consequence of changes in “costs” of the hypothetical monopolist;
- calculate the market price applying under each of these dispatch scenarios for each of the three short run marginal cost scenarios described above (ie, the base case short run marginal cost, the base case plus 5 and 10 per cent);
- calculate the gross margin for the hypothetical monopolist under each of these market price/dispatch scenarios, based on energy sent out from the monopolist’s plants; and
- compare the gross margin under the two SSNIP scenarios with that applying under the base case to determine whether the SSNIPs are profitable.

---

[14] Recall that a real monopolist will already be pricing up to the point where substitution effects prevent it from increasing its profits by raising prices further – see page 45, Ibid.

[15] Strictly speaking, infra-regional constraints might mean that a market is narrower than a single NEM region, but we do not consider that possibility here.

The approach set out above abstracts from outturn market prices, and so is protected from the finding of an inappropriately wide geographic market boundary in the event that prices already reflect a degree of market power.

The approach carries the implicit assumption that demand does not respond to an increase in the price of between 5 and 10 per cent. To the extent this assumption has an effect on the results it will over-estimate the profits of the hypothetical monopolist's price increase. That said, since electricity demand is generally considered to be price inelastic especially for relatively small price changes, the effect of this simplifying assumption is expected to be minimal.

Finally, as indicated, we believe there is merit in applying this methodology for both a 5 per cent and a 10 per cent increase in short run marginal costs for the relevant region's generators. This is because there may well be step changes in the dispatch from the region's generators associated with larger changes in costs for one or more particular plants. This effect could have a big influence on the results and so should be explicitly considered.

## 5.2. Data and calculation considerations

The approach to applying the SSNIP test described above requires information on:

- existing plant capacities, thermal efficiencies, and fuel prices for the region being investigated and any interconnected region;
- the transmission losses associated with supply from an interconnected region;
- the average transmission interconnector constraints for every 30 minute period; and
- observed loads for the region being investigated and any interconnected region.

These data are readily available from the AEMO.

Finally, the approach set out above can be applied for a number of historical years to determine whether the SSNIP is profitable under a number of different demand profile scenarios.

NERA
ECONOMIC CONSULTING

NERA Economic Consulting
Darling Park Tower 3
201 Sussex Street
Sydney NSW 2000
Tel: +61 2 8864 6500
Fax: +61 2 8864 6549
www.nera.com

NERA Australia Pty Ltd, ABN 34 092 959 665