ANALYSING REGION BOUNDARY CHANGES IN THE NEM:

WHAT HAVE WE LEARNED?

Darryl Biggar[1]

13 March 2007

1. In the past twelve months, regulatory authorities in Australia, especially the Australian Energy Markets Commission ("AEMC"), have given a great deal of attention to the question of the appropriate region boundaries in the National Electricity Market (the "NEM"), and how best to address the problems arising from network congestion in a regionally-priced market such as the NEM.

2. In the course of these deliberations, progress has been made in understanding the theoretical foundations of the NEM, in understanding the problems arising in the NEM's regionally-priced market, and, to an extent, in how to go about analysing potential region boundary changes.

3. This paper attempts to set out the current state of understanding of the conceptual analysis of region boundary changes in the NEM. This paper is divided into three parts. The first two parts set out the current understanding of the "problems" arising in a regionally-priced market such as the NEM – the so-called "mis-pricing" and "hedging" problems. It is shown how the likely mis-pricing and hedging impacts of a given constraint can be determined simply by examination of the form of the correctly-oriented constraint equation. The third part of this paper shows how to determine the correctly-oriented form of the constraint equations for any given change in the configuration of the administrative pricing regions and interconnectors through a simple transformation of the existing constraint equations.

[1] Darryl Biggar is a consultant to the ACCC/AER. This paper was prepared for the one-day seminar "Regulatory Issues in Telecommunications and Electricity", ACCC, 13 March 2007. The views expressed in this paper are those of the author and do not necessarily reflect the views of the AEMC, the AER or the ACCC.

1. What is the "problem" with the current pricing arrangements in the NEM?

4. All good public policy analysis begins with a clear definition of the "problem" or "market failure" which demands attention.[2] Without a clear definition of the problem it is not possible to correctly identify the options to be considered, or to correctly assess the pros and cons of those options.

5. In my view, a fairly clear picture is emerging as to the nature of the underlying problems in the NEM which are at the core of the "congestion management" and "region boundaries" issues. In my view there are two key problems with the current congestion management arrangements in the NEM. These problems I refer to as the "mis-pricing" problem, on the one hand, and the "hedging" problem, on the other. The following sections look at each of these (related) problems in turn.

6. The problems identified here are not the only problems in the NEM. However, the issues discussed here lie at the heart of the collection of issues known as "congestion management" issues. Many other issues – related to, say, the exercise of market power, incentives for transmission investment, or incentives for network service quality are related, but do not lie at the heart of the congestion management issues. In fact, in my view, the set of congestion management issues can be viewed as those issues which directly arise from these problems.

7. Of course, naming and describing these problems here does not automatically imply that these problems are material or significant. Although some progress has been made, the relative significance of these problems to other problems in the NEM is not yet clear. There is no intention here to imply that policy action to correct these problems is automatically warranted. That conclusion can only be drawn after careful analysis of the pros and cons of various policy actions to address these problems, including the option of remaining with the status quo.

1.1 The "mis-pricing" problem

8. Under the current NEM rules, the NEM Management Company (NEMMCO) operates a centralized spot-market dispatch process. Every five minutes of the day, generators and large loads submit their bids and offers to this dispatch process. Simultaneously, NEMMCO collects information on the state of the transmission network, current levels of demand and makes short-term forecasts of future demand. This information is passed to a central NEMMCO computer, known as the "dispatch engine". This computer, through an optimisation algorithm, determines the amount that each generator and large load should produce or consume – known as their "dispatch targets".

9. These dispatch targets are chosen by the dispatch engine in such a way as to maximise the "value of spot market trading" taking into account the physical losses on the transmission network and subject to the constraints that (a) the total amount of energy injected must equal the total amount withdrawn plus the total "losses" on the network; and (b) the physical limits of the transmission system must not be exceeded.[3]

---

[2] The Productivity Commission's Office of Best Practice Regulation in its Draft Best Practice Regulation Handbook states on page 1-1 that "The problem to be addressed and the related policy objective should be identified as first steps in the policy development process. This should be followed by consideration of a range of options (including no action) for achieving the objective and an analysis of the likely economic, social and environmental consequences".

[3] In most cases, the constraints on the dispatch process are not actually equal to the physical limits of the network but, instead, are such that the physical limits on the power system are not exceeded following an adverse event on the network known as a "credible contingency" (such as the loss of a major transmission link). However, this distinction is not important for our analysis. There are additional constraints on the

10. In this paper I will ignore the effects of losses. Although losses are a real factor in any practical transmission network, their effect is usually relatively small. Ignoring losses greatly simplifies the presentation and the analysis. The extent to which ignoring losses is more than an assumption of convenience is raised as a question below.[4]

11. The "value of spot market trading" is equivalent to the economist's notion of "total surplus" – that is, the area under the demand curve less the area under the supply curve at each location. In effect, the dispatch engine yields dispatch targets which maximise total economic surplus, subject to the physical constraints on the transmission network.

12. This design of the NEMMCO dispatch process guarantees that the resulting dispatch will maximise short-term economic efficiency, given the current state of the transmission network, provided the following conditions hold:

(a) the offer curves of each scheduled generator (or bid curves of each scheduled load) accurately reflect the social marginal cost of the output of each generator (or the marginal value of consumption to that load) at least at the margin;

(b) the actual physical limits of the transmission network are accurately represented in the dispatch process.

13. Joskow and Tirole (2004), highlight a number of additional conditions relating to the handling of those consumers which are insensitive to the spot market price and therefore may need to be rationed under certain circumstances.[5]

14. For the purposes of this paper, it is the first condition above which is the most relevant. Under what conditions will generators have an incentive to submit an offer curve which accurately reflects their social marginal cost (at least at the margin)? It turns out that scheduled generators will have an incentive to submit an offer curve reflecting their social marginal cost if and only if:

(a) the price a generator (or large load) is paid for its output and the amount to which it is dispatched, is a price-quantity combination on that generator's offer curve;[6]

(b) there is adequate or effective competition between generators at all locations in the market;

(c) all external impacts from electricity production (such as greenhouse gases) are internalized.

15. Let's focus on the first of these conditions (we will discuss the issue of market power later). Under what conditions will there be "consistency" between the offer curve of a generator and the quantity for which it is dispatched?

---

dispatch process, such as "ramp rate limits" which reflect the physical limits on the ability of a generator to increase or decrease its output. Again, we will put these to one side. There are also a number of other markets in "ancillary services" which are also co-optimised with the "energy" market. For the purposes of this analysis, we can put these ancillary services markets to one side.

[4] For an analysis which includes a full consideration of the impact of inter-regional and intra-regional losses on dispatch see Biggar (2006b).

[5] For example, Joskow and Tirole (2004) show that the conditions required include "(b) "rationing, if any is orderly and makes efficient use of available generation" and "(d) Consumers who can react fully to the real time price are not rationed. Furthermore, the LSEs serving customers who cannot fully react to the real time price can demand any level of rationing they prefer contingent on the real-time price".

[6] This is sometimes referred to as "consistency between pricing and dispatch".

16. Every generator and large load in the NEM is connected to the transmission network at a point known as its "connection point" or "node". At present in the NEM, connection points are assigned to a notional grouping of connection points known as an administrative pricing "region". One connection point in that region is designated the "regional reference node". Power is allowed to flow between these notional regions on notional "interconnectors". It is useful to keep in mind that there is no necessary connection between the boundaries of administrative pricing regions, or the location of notional interconnectors, and any piece of physical equipment in the real transmission network.

17. Importantly for our purposes, all generators in the same administrative pricing region receive the same price (ignoring losses), known as the regional reference price ("RRP"). This uniform price in each region is equal to the cost of supplying an additional unit of electricity at the "regional reference node".

18. If none of the physical limits on the transmission network are binding, the dispatch process described above will dispatch all generators to the point where the marginal cost of an additional unit of output (as revealed in each generator's offer curve) is equal for all generator's at all locations across the network. In other words the price for electricity is uniform across the network. Furthermore, each generator will be dispatched to an amount on its offer curve corresponding to that uniform price. Provided there is effective competition between generators, generators will have an incentive to submit an offer curve which reflects their true marginal cost (at least at the margin).

19. Now consider what happens when one or more constraints reflecting the physical limits on the transmission network are binding. In this case the dispatch engine must increase the output of some higher cost generators, while reducing the output of some lower cost generators. As a result, the local marginal price for electricity (that is, the marginal cost of an additional unit of output at a location, valued at the generators' offer curves) will vary at different locations across the network.

20. In other words, when any transmission constraint is binding there will arise different locational marginal prices at different geographic locations across the transmission network. Each transmission constraint will, in general, give rise to a different pattern of geographic differentiation of prices across the network – but for each constraint, the pattern of geographic differentiation that will arise is the same each time that constraint binds.

21. Since the pattern of geographic differentiation is the same for a given constraint each time that constraint binds, each constraint can be thought of as dividing up the network into pricing regions – with different connection points in the same "region" if and only if they would receive the same price when that constraint is binding (of course – these "regions" could be very small – just a single connection point).

22. Now, it may be that the boundaries of these hypothetical regions created by a transmission constraint, just happen to coincide with administrative pricing region boundaries. In this case, all connection points are correctly priced when that transmission constraint binds and there is no "mis-pricing" problem. However, in many cases, the boundaries of the pricing regions created by a given transmission constraint will not coincide with the administrative pricing regions.

23. In this case there arises a problem. In this case, a generator will be dispatched to a level of output consistent with its local locational marginal price, but will be paid an amount for its output equal to the regional reference price. Since the regional reference price and the locational

price are different there is said to be an "inconsistency" between pricing and dispatch.[7] Such a generator is said to be "mis-priced".

24. In the case of entities (such as scheduled generators) who receive dispatch instructions from NEMMCO, this inconsistency between pricing and dispatch is not a problem in itself (since the dispatch targets are efficient as long as the bids and offers reflect underlying costs).[8] However the inconsistency between pricing and dispatch has a serious impact on the incentives for such entities to submit a bid or offer which reflect their true costs.

25. Consider first the case where a generator is paid a relatively high (regional) price for its output, but is dispatched to a level on its offer curve consistent with a much lower local price. If the quantity the generator would like to produce at the high regional price is larger than the quantity the generator would like to produce at the lower local price[9], such a generator is said to be "constrained off". A generator that is constrained off has a strong incentive to offer its output at a lower price, in an attempt to increase the amount for which it is dispatched, to bring its dispatch volume back into line with the price it is paid.

26. If there are a number of competing generators that are constrained off, each has an incentive to lower the price at which it offers its output, in an attempt to increase the amount for which it is dispatched. Whether there is just one or several generators which are constrained off, this process continues until either: (a) each constrained off generator is able to increase its dispatch up to the quantity it would like to be dispatched given the price it is paid (in which case some other generators elsewhere in the network must have their output reduced); or (b) each constrained off generator offers its output at the lowest allowed offer price, which is $-1000/MWh.[10]

27. Now consider the case of a generator which is paid a relatively low (regional) price for its output, but is dispatched to a level on its offer curve consistent with a much higher local price. If the quantity the generator would like to produce at the much higher local price is larger than the quantity it would like to produce at the low regional price, the generator is said to be "constrained on". A generator that is constrained on has an incentive to offer a portion of its output at a higher price, in an attempt to reduce the amount for which it is dispatched, to bring its dispatch back into line with the price it is paid.

28. Again, whether there is just one generator or many generators which are constrained on, each has an incentive to raise the offer price for all or a portion of its output to the point where either (a) each constrained-on generator is able to reduce its output to the level it would like to be dispatched given the price it is paid (in which case some other generator(s) elsewhere in the network must have its output increased); or (b) each constrained-on generator offers some or all of its output at the highest allowed offer price, which is $10,000/MWh.

29. Generators which are constrained-on or constrained-off have a strong incentive not only to change their offer curves but also to change whatever other bidding parameters are available to them to increase or decrease the amount for which they are dispatched.

---

[7] See the Electricity Rules 3.1.4 (a) 4 and the discussion in CRA (2002).

[8] In the case of entities who do not receive dispatch instructions from NEMMCO but "self-dispatch", such as non-scheduled generators and loads, mis-pricing will directly impact on their output decisions and therefore will have a direct impact on efficiency.

[9] This will be the case if the marginal cost curve of the generator is upward-sloping. If the marginal cost curve of the generator happens to be vertical in the region between the local price and the regional price, and there is effective competition at that location, there will not be "inconsistency between pricing and dispatch" and the generator will not have an incentive to attempt to alter its output in the ways described here.

[10] If two or more generators offer their output at the same price, the "tie-breaking" rules require that the dispatch engine dispatch those generators in an amount proportional to their offer quantity.

"[G]enerators have been known to declare themselves "inflexible"[11], to use their ancillary services bids in a way which prevents their output in the energy market being reduced, or to reduce their ramp rates to slow the rate at which their output can be scaled down.[12] The AER (and its predecessor, NECA) has prosecuted generators for these actions and is considering its options to clarify and strengthen its powers to prevent similar actions in future. Unfortunately, generators face strong continuing commercial incentives to innovate in developing new techniques to prevent their output being reduced (or increased) when they are constrained off (or on, respectively)."[13]

30. Even more seriously, generators which are constrained on or off may simply fail to follow the dispatch targets issued by NEMMCO. Under current NEMMCO procedures, if a generator does not follow its formal dispatch targets for three consecutive dispatch intervals, NEMMCO imposes a constraint in the dispatch process which has the effect of holding constant that generator's output.[14] In the case of a generator which is attempting to resist being constrained on or constrained off, it may be highly desirable to have its output fixed in this way. Yet, failure to follow the dispatch instructions of the central dispatch process puts system security and reliability at risk. The AER notes:

"The requirement for participants to comply with dispatch instructions, as specified in chapter 4 of the Rules, is fundamental to the secure operation of the power system. … Registered Participants must endeavour to comply with dispatch instructions".[15]

31. This problem of mis-pricing, leading to distorted bidding incentives, has also been referred to as the problem of "disorderly bidding" by Charles River Associates (CRA)[16], or the problem of "strategic bidding" in the recent AEMC draft determination[17]. A recent paper from CRA in the US refers to this problem as the lack of "incentive compatibility".[18]

32. The mis-pricing problem is largely unique to the Australian NEM. Although there are a number of other markets overseas which feature uniform zonal or regional pricing (such as the UK, Texas and California), virtually all of these have an explicit system of "constrained up" or "constrained down" payments which ensure that generators which are constrained-on or off are

---

[11] That is, to bid "fixedload". The AER has stated: "It is critical that the inflexibility provisions are used only where abnormal operating conditions exist", AER, October 2006, "The events of 31 October 2005: Investigation Report", page 36.

[12] "The AER is concerned that the practice of rebidding reduced ramp rates for commercial reasons jeopardised system security. This is because the market systems are prevented from being able to quickly adjust power flows to respond to issues that emerge in the market. The AER is aware that other physical bid parameters (including frequency control ancillary service trapeziums) have also been used for commercial reasons, with a detrimental impact on power system security management. AER, October 2006, "The events of 31 October 2005: Investigation Report", page 36.

[13] Biggar (2006c).

[14] See NEMMCO, "Operating Procedure: Dispatch: Document Number SO_OP3705", section 9, "Management of Non-Compliance in Energy Market Dispatch", page 17-18. See also clause 3.8.23(a) of the Rules. Non-conformance in the energy market may increase the FCAS payments under the "Causer pays" policies (set out in various NEMMCO documents).

[15] AER (2006b). The AER has recently issued a Compliance Bulletin making clear its intention to prosecute instances of non-conformance with dispatch targets.

[16] CRA (2004a) notes: Mis-pricing "can result in "disorderly" market behaviour where participants exploit what are in effect loopholes in market rules resulting [in] a less efficient dispatch order. The absence of intra-regional pricing means there is no commercial discipline on this behaviour. To date however, we understand that the limited occasions where this has been observed has not led to material changes in the dispatch order, but may eventually lead to uncertainty about contracting levels and reduced liquidity of contracting, in turn leading to lower levels of contracting, and increased price volatility".

[17] AEMC (2007), page 30-31. The use of the term 'strategic bidding' is potentially unfortunate as it is the same term the AEMC uses to describe the exercise of market power which is, as noted in the text, a different problem.

[18] See CRA (2004b).

paid the correct locational price at the margin (or a price based on their costs where there is an absence of effective competition).

33. One exception, however, is the case of Alberta, Canada. This is a regionally-priced market. However, the system operator can make constrained-up but not constrained-down payments. Unsurprisingly then, the same problem of mis-pricing leading to distorted bidding incentives occurs in Alberta in the case of generators which are constrained off.

> "When real-time congestion does arise, usually because of an unexpected grid outage or other problem, the system operator decides which resources to constrain up and down using the bid-based merit order from the Power Pool Administrator as far as practical. But in the absence of constrained-down payments, generators may try to avoid being constrained down by lowering their bids to zero (the lowest bid allowed in Alberta), making merit-order dispatch largely meaningless if not impossible".[19]

34. Mis-pricing is not only a problem for scheduled generators. In the NEM, large loads also have the facility to bid the amount they would like to purchase into the NEM dispatch engine. At present, this facility is primarily used by hydro generators, who have the capacity to pump water uphill at times of low prices.

35. In principle, such scheduled loads can also be mis-priced with the same implications for distorted bidding incentives. For example, suppose a load wishes to consume 1000 MW if the price is below $10/MWh and nothing otherwise. If that load is mis-priced it could be required to pay a regional reference price of $100/MWh while being dispatched to 1000 MW. In this circumstance, the load could be said to be "constrained on" and has an incentive to lower the price at which it is prepared to consume (or simply to not make an offer at all). Conversely, if the load were mis-priced it might occur that the load is not dispatched even though the regional reference price is $5/MWh. In this case, the load is "constrained off" and has an incentive to raise the price at which it is prepared to consume, in an attempt to increase the amount for which it is dispatched.

36. In fact, the problem of mis-pricing, with the ensuing incentive to distort bids offers can, in principle, affect any dispatchable entity which makes offers to NEMMCO, including "market network service providers" (MNSPs) and FCAS providers.

37. Importantly, even entities which do not directly participate in the spot market are affected by mis-pricing. Even so-called non-scheduled generators and loads (which normally act as "price-takers" in the market), will produce more or less than is efficient when faced with a regional reference price which differs from the local nodal price.

> "The most important short-run inefficiencies cause by [regional pricing] arise from the fact that any value of the [regional price] will differ from the [nodal price] at many/most nodes, which gives those who pay or are paid the [regional price] incentives to consume or produce more or less than they would if they paid or were paid the [nodal price] at their node. Unless these short-run incentives are offset by other short-run payments or charges, dispatchable resources will adopt strategies to influence the schedules they get back from the ISO and/or to let them deviate from their schedules without being caught or penalised. Price-taking loads (and any price-taking generation) who pay the regional price instead of [the nodal price] and simply respond to the [regional price] as they choose will consume (or produce) more or less than the efficient amounts corresponding to the [nodal price]."[20]

---

[19] CRA (2004b), page 15.
[20] CRA (2004b), page 37.

38. One extreme case of mis-pricing is worth special mention. Under the current market rules, NEMMCO has the power to over-ride the price determined in the dispatch process and replace that price with the price ceiling ($10,000/MWh). When this occurs (as occurred in Victoria on 16 January 2007) all the generators in that region are simultaneously "mis-priced" and have an incentive to manipulate their bids (down to $-1000/MWh), bid inflexible, fail to follow dispatch instructions, and so on, in an attempt to maximise the amount for which they are dispatched. Such actions can threaten the ability of NEMMCO to maintain the system in secure operating state.[21]

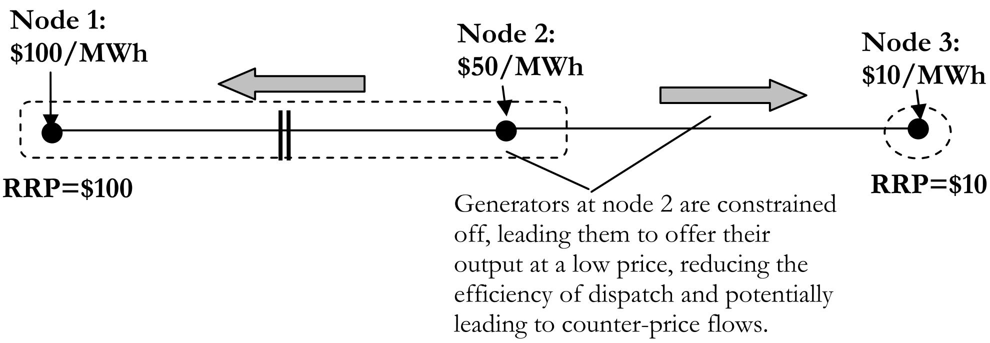
39. As we will see below, mis-pricing is economically inefficient for several reasons. Mis-pricing may (on rare occasions) threaten system security and reliability, reduce the short-term efficiency of dispatch, reduce the firmness of settlement residues and induce inefficient generator and load location decisions. First, however, let's look more closely at the effects of mis-pricing in simple networks.

Mis-pricing in radial networks

40. Mis-pricing can easily be illustrated using very simple network examples.[22]

41. For example, figure 1 below illustrates a simple three-node network with two regions and two transmission links. Nodes 1 and 2 are in one administrative pricing region and node 3 is in another. At node 1 there are a number of generators with a variable cost of $100/MWh, at node 2 the generators have a variable cost of $50/MWh, and at node 3 the generators have a variable cost of $10/MWh.

Figure 1: Generators constrained off in a three-node network



42. In this example, power flows have reached the physical limit on the link between node 1 and node 2 in the direction of node 1. In this simple network, if all nodes were correctly priced, this single binding constraint would have the effect of dividing the network into two pricing regions – with node 1 in one region (with a price of $100/MWh) and nodes 2 and 3 in another (with a price of, say, $10/MWh). But of course, these regions do not correspond with the administrative pricing region boundaries in this network, giving rise to mis-pricing.

43. In particular, the generators at node 2 are paid the price at node 1 ($100/MWh) but are dispatched on their offer curve to a price corresponding to $10/MWh. These generators are

[21] Mis-pricing may also give rise to incentives on generators to distort their bids between the energy and FCAS markets. This is clearest in the case when the VoLL override is applied in the energy market but not in the FCAS market. In this case generators have an incentive to not offer their output in the FCAS market and instead focus on the energy market. It is not yet clear whether this can happen during other episodes of mis-pricing.

[22] Several more examples can be found in the appendix of Biggar (2006a).

therefore mis-priced. Since these generators have a variable cost of $50/MWh, at the regional reference price they would like to produce at their maximum capacity. But, if they offer their output at their true economic cost, these generators, which are dispatched to the point on their offer curve corresponding to the price of $10/MWh, will not be dispatched at all. Since, at an offer price equal to their true cost, these generators are dispatched for less than the amount they would like to produce, these generators are "constrained off".[23]

44. These generators which are constrained off have an incentive to lower their offer in an attempt to be dispatched. If they lower their offer price below $10/MWh they will be dispatched ahead of the generators located at node 3. The additional output of these generators may be sufficient to restore consistency between pricing and dispatch. However, if it is not, these generators will continue to reduce their offer price, all the way down to the allowed price floor, $-1000/MWh.

45. Whether these generators lower their offer to the price floor or not, the reduction in the offer price at node 2 will cause the dispatch engine to dispatch node 2 generation ahead of the generation at node 3. This is inefficient since, in this example, node 3 generation is cheaper than node 2 generation. Furthermore, depending on the pattern of load in the network, the direction of the flow on the link between node 2 and node 3 may be reversed – in fact it may easily arise that the flow between these two regions is "counter-price" (from the high price region to the low price region) giving rise to "negative settlement residues", which are discussed further below.

46. As a second example, let's retain exactly the same physical network and generation capability as in the network of figure 1, but now let's consider changing the administrative pricing region boundaries of this network, so that now nodes 2 and 3 are in the same region, as illustrated in figure 2 below.

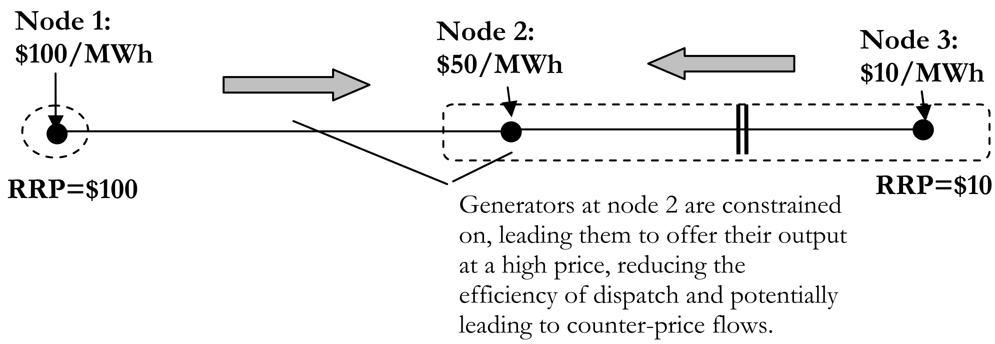
47. Now let's suppose that there is a constraint between node 2 and node 3, in the direction of node 2. In an efficiently priced network, this transmission constraint would divide the network into two pricing regions – with nodes 1 and 2 in one region and node 3 in the other. Again we see that these pricing regions do not coincide with the administrative region boundaries defined in this network. In particular, node 2 is again mis-priced. In this case however, generators at node 2 are "constrained on" – they are dispatched to a price corresponding to $100/MWh but are paid only $10/MWh for their output.

48. Since these generators have a variable cost of $50/MWh they would prefer to not be dispatched when they are paid only $10/MWh. They will therefore seek to bid in such a way as to reduce the amount for which they are dispatched. They can do this by raising their offer price above $100/MWh (the price of generation at node 1) or up to the price ceiling of $10,000/MWh.

49. As before, distorting bidding in this way is inefficient since any reduction in the output of node 2 generators must be offset by an increase in higher-cost generation at node 1. In addition, depending on the location of loads in this network, it may be that the increase in output of node 1 generation causes flows between node 1 and node 2 to be counter-price, leading to negative settlement residues.

[23] In the terminology of CRA (2004a), the generators at node 2 are "gatekeeping" generators.

Figure 2: Generators constrained on in a three-node radial network



The economic harm from mis-pricing

50. Mis-pricing is economically inefficient for several reasons. As already noted, mis-pricing reduces the short-term efficiency of dispatch, reduces the ability of affected generators to hedge their risk, reduces the firmness of settlement residues, and induces inefficient generator location decisions. This is explained further below:

1. First, mis-pricing may, in certain (rare) circumstances, **threaten system security and reliability**. When generators distort their bids or fail to follow dispatch instructions from NEMMCO, circumstances may arise where some customers cannot be served, or the network is not able to withstand shocks known as "contingencies". However, these circumstances are rare.[24] Where system security or reliability is threatened NEMMCO has the power to direct a generator to increase or reduce its output. This power (backed up with the promise of compensation for generators who are directed in this way) is usually adequate.

2. Generator mis-pricing and the associated distorted bidding incentives will, in general, **reduce the short-term economic efficiency of dispatch**. Whenever any dispatchable entity (such as a generator or a large load) is able, by distorting its bid above or below its true short-run marginal cost, to increase or decrease the amount for which it is dispatched, the overall efficiency of dispatch is lowered – that is, more expensive generation must be dispatched to meet the total load on the system while less expensive generation remains available without violating the physical constraints on the network.

   In the same way, whenever a price-taking entity (such as a non-scheduled generator or load) is paid a price different from the local nodal price, causing that generator or load to produce or consume more than is efficient at that location, the overall efficiency of dispatch is lowered.

   The magnitude of the reduction in the efficiency of dispatch depends on both the amount of the change in output brought about by the mis-pricing and the cost difference (or difference in the marginal value of consumption) between the generators whose output is reduced and the generators whose output is increased. If the change in dispatch brought about by mis-pricing is small (because (a) the mis-priced generator is small, (b) the magnitude of the mis-pricing is small, or (c) there are no other generators who can substitute for the output of the mis-priced generator), or if the change in dispatch brought about by mis-pricing is made up by other generation with a similar variable cost, the dispatch efficiency impact of the mis-pricing will likely be small. Conversely, if the change in dispatch brought about by the mis-pricing is large, or if the change in dispatch

[24] One possible example is when NEMMCO imposed the "VoLL override" on 16 January 2007.

is made up by other generation with a very different variable cost, the dispatch efficiency impact of the mis-pricing will likely be large.

3. Mis-pricing **increases dispatch risk** for mis-priced generators and (therefore) reduces such generator's optimal level of hedge cover. A mis-priced generator faces no price risk in contracting with its regional reference node, however it may not be able to be certain about the quantity for which it will be dispatched. The magnitude of this "dispatch risk" depends on a number of factors such as: the number of other mis-priced generators with which it competes and the investment in and opportunity for discovering new bidding schemes which increase or decrease the amount which the mis-priced generator is dispatched. If a generator cannot be sure that it will not be the loser in any given bidding war, it will need to build in a "buffer" into its hedging decisions to ensure it is not put in a position where it is unable to cover its hedging obligations through its physical generation.[25]

4. In addition, as we will discuss further below, generator mis-pricing and the associated distorted bidding incentives, **reduce the "firmness" of what are known as "inter-regional settlement residues"**, reducing the efficiency of the forward contract or "hedge" market. Furthermore, mis-pricing will, in some cases, give rise to negative settlement residues, which, as discussed below, give rise to their own problems for the market. This lack of efficiency in the hedge market has flow-on consequences for the prices of hedge contracts and the hedging decisions of generators, which may, in turn, affect incentives to exercise market power.

5. Finally, and possibly most importantly, generator mis-pricing **distorts generator and load location and expansion decisions**. Generators and large loads may choose to locate at a point on the network which exacerbates congestion and may choose to forego opportunities to invest in locations which alleviate congestion.

51. This last point is important and deserves careful further elaboration. This distortion to location decisions was, for example, a primary concern of the Energy Reform Implementation Group who write:

> "Even more concerning in terms of the longer term efficient development of the market is that unpriced congestion can create perverse locational incentives on investments in new generators. In some circumstances, this can lead to strategically located generators using anomalies between optimised dispatch based on bids and regional settlements to effectively gain preferential access to the transmission system. … In the longer term these outcomes are likely to reduce the confidence of investors in the integrity of NEM pricing outcomes and contribute to ill timed or inappropriately located investments, all of which will add costs to consumers.
>
> ERIG is aware of one proposed generation investment where the proposed location would have induced network congestion. The proposed generator would hold a strategic advantage as a result. In this case a relatively small change in the proposed location would have removed the risk of such congestion, but the incentives on the proponent were the opposite.
>
> In its 2005 Annual Planning Report, TransGrid states:
>
>> "At times when NSW is a heavy importer of power from the south the line ratings within the Snowy system and immediately north of Snowy may impose a limitation. The development of any generation in the south west of NSW (such as gas turbines in the Wagga area) will impose greater competition for the

[25] In this case, the generator would need to make what is known as "unfunded difference payments".

limited power transfer capability to the north. Such generation does not provide an effective increase to the net NSW generation at times where the transmission system is limiting." [26]

Despite this, a major new power station is proposed for Wagga.

… [T]he weakness of locational signals to generators needs to be addressed. Aligning the commercial incentives of generators with the capacity of the existing network removes the generator incentive to use network congestion to extract commercial advantage, and contributes to a more efficient use of network infrastructure. … Failure to improve locational pricing signals for generators will directly undermine any long term allocative efficiency benefits through inefficient investment location and will lock in those inefficiencies for the life of the assets".[27]

52. Generator and load location or expansion decisions will, of course, depend on a large number of factors, such as availability of fuels and access to key inputs, such as water. However, it seems clear that, at the margin, generator location decisions will be affected by the spot (and "forward" or "contract") prices a generator is likely to receive for its output at any given location.

53. For example, it seems clear that, other things equal, both scheduled and non-scheduled generators have a reduced incentive to locate or expand in locations which are "constrained on" or "under-priced" (that is, the correct locational price exceeds the regional reference price). (However, as we will see below, in these locations there are other arrangements in the NEM designed to induce generators to invest in these regions). Conversely, large loads will have an inefficiently small incentive to avoid locating in regions which are periodically "constrained on".

54. On the other hand, on the surface it appears that other things equal, scheduled and non-scheduled generators will also have an *over-incentive* to invest in locations which are "constrained off" or "over-priced" (that is, where the correct locational price is below the regional reference price). Indeed, some generators (including non-scheduled generators, such as wind-farms) may be induced to locate in constrained-off locations, even when it is inefficient to do so.

55. Paradoxically, this "over-incentive" to invest in constrained-off regions may have a significant "chilling effect" on the incentives for efficient investment or expansion of generators in constrained-off locations.

56. To see this, consider the position of a low-cost scheduled generator choosing whether or not to locate at a node which is remote from the regional reference node. Such a generator faces the risk that another generator will locate nearby, leading to congestion on the transmission link to the regional reference node. However, in a market with locational pricing there is something of a deterrent on further generation investment – any prospective entrant locating nearby would recognise the risk of transmission congestion and the impact on the local price. Such investment would presumably only go ahead if it could displace some or all of the incumbent generation capacity. In a market with nodal pricing, an efficient low-cost incumbent has a degree of assurance that no new entry will locate nearby. As long as such a generator can maintain its efficiency, it can be assured that no rivals will locate nearby, causing congestion to the regional reference node.

---

[26] Transgrid, NSW Annual Planning Report 2005, page 99.

[27] ERIG (2006), page 116-118. Wambo Power Ventures, the developer of the power station at Wagga has criticised the use of this extract from the TransGrid APR in a letter to the AEMC on 5 January 2007. Access to two alternative sources of gas, together with localised load growth and reliability concerns, are claimed to have played a greater role in the locational decisions than the projected impact of transmission constraints on the output of the station and its access to the NSW RRP. Nevertheless, Wambo Power Ventures have strongly opposed the creation of a "southern NSW" region in the NEM.

57. Now consider the position of a low-cost scheduled generator in a market with regional pricing. In this case, a prospective entrant could locate nearby and potentially displace some of the capacity of the incumbent – even if the entrant is higher cost than the incumbent. If the new entrant is another scheduled generator, when a constraint arises to the regional reference node, both generators will have an incentive to offer their output at $-1000/MWh. The dispatch engine, seeing both generators offer their output at the same price will dispatch each generator for the same volume. In other words, the new entrant can capture 50% of the scarce transmission capacity even though it has higher costs. The problem is even worse when the new entrant generator is a non-scheduled generator such as a wind-farm. In this case the new entrant can capture as much as 100% of the scarce transmission capacity, no matter what its costs.

58. In other words, in a market with regional pricing even an efficient low-cost entrant cannot be assured that further entry will not occur. This threat of further entry potentially has a significant deterrent or "chilling effect" on efficient low-cost generation investment remote from the regional reference node.

59. Larry Ruff, in a paper for CRA (2004b) observes that all the long-lived examples of regionally-priced wholesale electricity markets overseas, make use of some other transmission pricing mechanism to improve generator location investment decisions:

> "All of the [regionally priced] markets that have survived for some years require generators and loads to pay grid access charges that are higher where their operations increase congestion. Some such mechanism is needed to provide the long-term locational decisions that a [regionally priced] market, with or without [constrained on/off payments] does not provide".[28]

60. There may also be a sixth impact on mis-pricing – on transmission investment decisions. Mis-pricing, and the resulting economic harms set out here, may create pressure to "build out" constraints which could be more efficiently handled through more accurate price signals. Although it may be possible to alleviate some transmission constraints with a relatively small investment, large upgrades to the transmission network routinely cost in the hundreds of millions of dollars. As a result, it is not economically efficient to "build out" all constraints. Put another way, it is likely that there will always remain an "efficient level of congestion" on any efficiently-run transmission network.

61. However, the economic harm from mis-pricing noted above, may create pressure for the "building out" of transmission constraints even when it is not economically efficient to do so. CRA (2004b) note that one of the lessons from the history of zonal pricing systems (such as the NEM) is that "poor congestion management increases pressure to spend too much on the network".[29] Interestingly, as noted earlier, the only other example of a regionally-priced transmission network without constrained-down payments occurs in Alberta, Canada.[30] The Government of Alberta has explicitly adopted a policy of building out transmission constraints – irrespective of costs. In fact, Alberta has explicitly adopted a policy of pursuing a "congestion-free" network, irrespective of the economic cost:

> "The open access transmission structure in Alberta consists of an implicit system of injection and withdrawal rights for generators and loads. There are no explicit transmission rights. Given this structure, the transmission system must be relatively congestion free or the underlying market model will not function effectively. The ISO must therefore proactively plan transmission development to achieve this result of "congestion-free" transmission. The ISO will be required to ensure that the transmission system internal to Alberta is appropriately reinforced so that under normal operating

---

[28] CRA (2004b), page 20.
[29] CRA (2004b), page 14.
[30] As noted earlier, unlike the NEM, Alberta does have constrained-up payments.

> conditions (i.e. all transmission facilities in service) all in-merit generation can be dispatched and virtually all economic wholesale transactions may be realized without congestion."[31]

62. The Energy Reform Implementation Group recently noted that the adoption of a similar policy in the context of the Australian NEM would not be in accordance with the NEM Objective:

> "The ability of a generator to influence network congestion can be addressed by increasing network capacity. However, ERIG notes that the costs involved in increasing the network capacity to alleviate all congestion are likely to be prohibitively large. Further, ERIG notes that alleviating all congestion would remove the only current signal for the efficient location of generation. A regime which guaranteed the removal of congestion in the future, coupled with a transmission pricing regime which does not charge generators for that service, would bias the choice of technology towards remotely located generation, ahead of local generation options or other alternatives such as demand side response. Such a solution would substantially decrease the long term allocative efficiency potential of the electricity sector. Accordingly, ERIG concludes that this is not an efficient solution and may not align with the NEM objective with respect to serving the long-term interests of consumers."[32]

Existing mechanisms for mitigating the impact of mis-pricing

63. It is worth noting that the NEM already has some mechanisms in place for mitigating some of the worst economic impacts of mis-pricing, at least for constrained-on generation. As already noted, when a generator is "constrained on", that generator has an incentive to offer its output at a high price or, in other ways pretend to be unavailable. Despite the puzzling prohibition in the Rules on compensation for "constrained on" generators[33], there currently exist mechanisms which ensure that such generators are, at least, partially compensated for locating in regions which are likely to be constrained on, and for producing at those times.

64. In some areas of the NEM, such as far-North Queensland, transmission constraints require that, on occasion, high-cost local generators are producing to meet local demand, even though the regional reference price is relatively low. In these circumstances, the failure to correctly price the additional output required may result in a shortfall in supply – either in the short-term as local generators refuse to produce, or in the longer term as new generation refuses to locate in constrained-on locations. In the absence of any further intervention there would be a strong likelihood of load shedding. Reliability for customers in far-North Queensland would suffer.

65. The Electricity Rules allow NEMMCO to intervene in the market in these circumstances so as to maintain the reliability of supply. Specifically, where distorted bidding of the kind noted above threatens to result in the loss of supply to some customers, NEMMCO has the power to issue an order (known as a "direction") to a generator requiring it to produce. The compensation for this output is determined by an independent expert.[34]

66. In addition, transmission networks ("Transmission Network Service Providers" or "TNSPs") have the power to offset the inefficient outcomes of mis-pricing above by signing contracts with new or existing generators, under which the generators agree to produce when

---

[31] Alberta (2003), page 8.
[32] ERIG (2006), page 115-116.
[33] Section 3.9.7(b) of the Rules states: "A Scheduled Generator that is constrained-on in accordance with clause 3.9.7(a) is not entitled to receive from NEMMCO any compensation due to its dispatch price being less than its dispatch offer price".
[34] The requirements governing such compensation are set out in sections 3.12.11 and 3.12.11A of the Rules.

they are required to do so for system reliability reasons and, in exchange, the generators receive a payment for their output. Such contracts between TNSPs and generators are known (somewhat confusingly) as "network support agreements". Powerlink, in particular, is known to have network support agreements in place with generators in far-North Queensland.[35]

67. Both directions to constrained on generators and network support agreements can be viewed as a response to the "mispricing" problem – both "interventions" would not be required with more accurate pricing arrangements in the NEM. Both interventions involve partially correcting the price-signals paid to these generators – in both instances by increasing the payment to "constrained on" generators, at least at the margin.[36]

68. These mechanisms are only a partial and limited solution to the problems arising from mis-pricing. Directions and network support agreements are only used when the mis-pricing problem threatens the reliability of supply. Directions and network support agreements will not be used when generators are constrained on, and yet there is no threat to the reliability of supply. In addition, constraints which cause generators to be constrained off, with the associated distorted bidding incentives, will not normally give rise to system reliability issues and therefore, will also not normally be corrected through directions or network support agreements.[37]

Mis-pricing in meshed networks

69. Both of the simple network examples above (in figure 1 and figure 2) featured simple "linear" or "radial" networks – that is, networks without loop-flows. A radial or linear network is a network in which there is only one path between any two points on the network.

70. In a simple radial or linear network there is a simple rule governing when two locations of the network should receive a different locational marginal price:

> In a linear or radial network, two geographic locations on the network should receive a different price if and only if there is a binding transmission constraint on at least one element of the (unique) transmission network path between those two points.

71. In contrast, a network with loop-flow or a "meshed" network, is a network with two or more paths between any two locations on the network. As we will see, the scope for mis-pricing is significantly larger in a network with loop-flow. In fact, as we will see, in a meshed network two nodes on the network should receive a different price if and only there is a binding constraint on *at least one of the paths* between those two nodes.

---

[35] For details of the Powerlink NSA, see Appendix 3 of AEMC, *Congestion Management Review: Issues Paper*, AEMC, Sydney, 3 March 2006. CRA note: "The Australian NEM may be the only UMP market in the world operating with neither constrained-up payments nor constrained-down payments. NEMMCO can and does constrain generation up and pay constrained-up payments in emergency situations, but the philosophy is to use longer-term solutions. Where network upgrades are too costly or are delayed, a regional (state) grid owner can use network support contracts that pay a generator to run just enough to relieve congestion, with the costs of such contracts recovered through region-wide (state-wide) tariffs; the few such contracts are used to support isolated loads at the end of long lines.". CRA (2004b), page 20.

[36] The use of the term "interventions" here is not intended to be pejorative – I make no judgement that these mechanisms are better or worse than alternatives, such as increasing the number of regions.

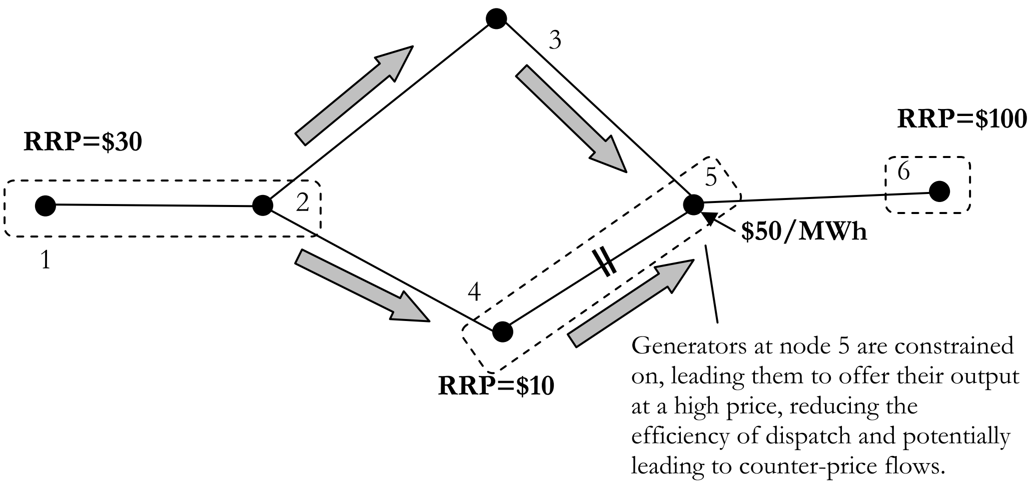
[37] It is useful, as an aside, to consider which generators gain or lose from the existing regional pricing arrangements. At first glance it might appear that generators which are constrained off ("over-priced") gain while generators which are constrained on ("under-priced") lose. However, if we take into account directions and network support agreements we can see that generators which are constrained on may be receiving adequate compensation for their output through these other mechanisms. As a result, although there is some loss in efficiency from regional pricing, it is far from clear that incumbent generators are worse off than they would be under a larger number of regions.

72. As we will see, in a meshed network, it is highly unlikely that the efficient pricing regions arising from any given binding constraint will coincide with any pre-determined broad administrative region boundaries.

73. These conclusions are illustrated below, using the network set out in figure 3 (which is designed to broadly reflect the features of the NEM in the vicinity of the Snowy region). When there is a constraint anywhere on a loop in the transmission network, the physics of AC power flows requires that the locational price for electrical power vary at *all* locations around that loop. The price is the lowest at the node which is directly connected to the "upstream" side of the constrained line, and highest at the node which is directly connected to the "downstream" side of the constrained line. The prices increase around the loop from the lowest-price node to the highest-price node.[38]

74. In the example of figure 3 below, there is a binding constraint between node 4 and node 5 in the direction of node 5. Therefore, applying the principles noted above, the lowest-priced node on the loop is node 4 and the highest-priced node is node 5. Furthermore, the nodal electricity price must increase clockwise around the loop from node 4 to node 5. As we can see, in this example, the nodal price at node 4 is $10/MWh, increasing to $30/MWh at nodes 1 and 2, and increasing to $100/MWh at nodes 5 and 6. In other words, this one constraint (between nodes 4 and 5) divides this network into four pricing regions – node 4, nodes 1 and 2, node 3 and nodes 5 and 6.

Figure 3: Generators constrained on in a meshed network



75. As this example clearly shows, in a meshed network there is no connection between the physical location of a constraint and the boundaries of the efficient pricing regions. In general, there will be as many pricing regions as there are nodes on the loop – some of which may be hundreds or thousands of miles from the physical location of the constraint.

76. Importantly, this nodal pricing outcome arises regardless of the direction of flows on the other (unconstrained) links in the network. In figure 3 we noted that the nodal price for electricity must increase from node 4 to node 2. If, as illustrated in figure 3, the flow between node 2 and node 4 happens to be in the direction of node 4, these flows will appear to be "counter-price" and may give rise to negative settlement residues (depending on the definition of the administrative pricing regions as discussed later).

[38] This is known as the "spring washer effect".

77. Earlier we noted that the single constraint in figure 3 can be thought of as having the effect of dividing this network into 4 pricing regions. However, these pricing regions do not coincide with the administrative pricing boundaries defined in figure 3 above. As before, this implies that some generators are mis-priced.

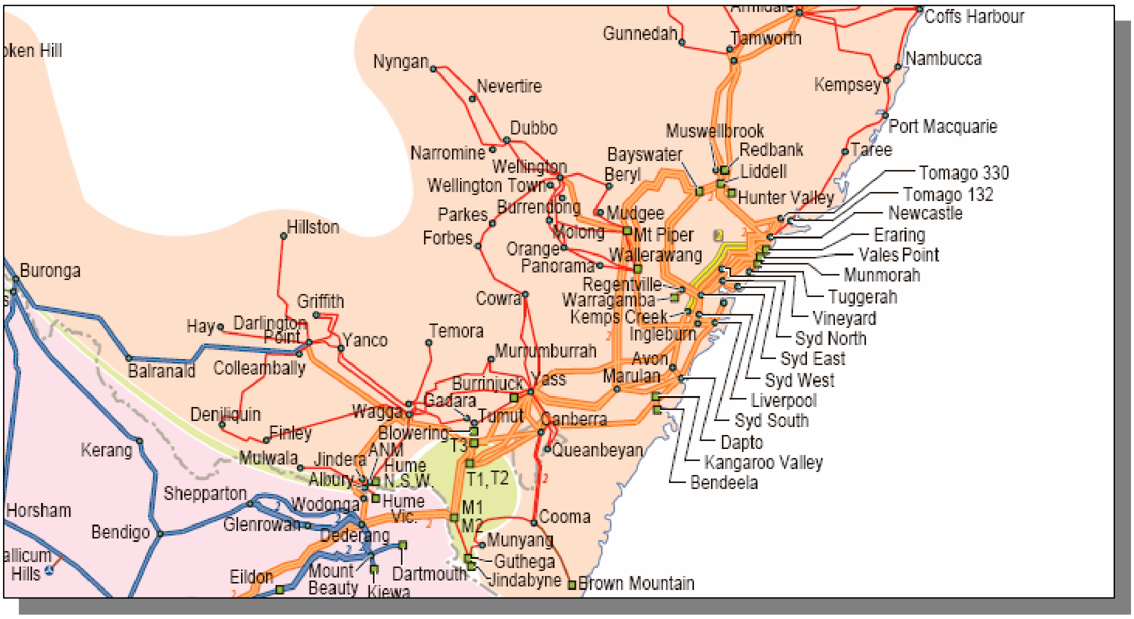
78. In particular, generators located at node 5 in figure 3 are mis-priced. In this case, they are constrained on – that is, they are dispatched corresponding to a price of $100/MWh but are only paid a price of $10/MWh. These generators respond by offering their output at a high price, or otherwise pretending to be unavailable. As before, if these generators are successful at reducing their output, the efficiency of dispatch is reduced. Furthermore, if there is load at node 5, the flow on the link from node 5 to node 6 may be in the direction of node 5 – that is, counter-price – giving rise to negative settlement residues.

79. The important point to recognise here is that, as emphasised above, a single binding constraint on a transmission loop will give rise to different prices at every node around that loop. In fact, in a highly meshed network, a single transmission constraint can give rise to a different price at *every single node* in the network. It is highly unlikely that such narrow pricing regions will coincide with the boundaries of any broad administrative pricing boundaries.

80. This pricing effect of constraints in a meshed network is not merely a theoretical curiosity. Most real-world transmission networks are, in fact, significantly meshed over the bulk of the network. Loops in transmission networks are common – indeed, they are the norm. It is relatively unusual for two points on a transmission network to be connected by a single transmission path.

81. There are good reasons for building loops and multiple paths in transmission networks. Allowing for more than one path between two points on the network allows for a higher degree of reliability – that is, the continued operation of the network in the event of the failure of any one of its elements. Electricity transmission networks are sometimes informally referred to as an electricity *grid* to reflect the fact that there is typically more than one path for the electricity to flow between any two points. The following diagram sketches the transmission network in the NSW region. As can be seen, the network features many loops. Even if we focus on the 330 kV network alone (which does not make electrical sense) there remain a large number of loops.

Figure 4: Extract of the NEMMCO system diagram for the NSW region



(In this figure, the 500 kV network is shown in yellow, the 330 kV network in orange, the 220 kV network in blue, and the 132 kV network in red.)

82. Even if the NEM is not highly meshed by international standards, many parts of the NEM (such as in NSW) remain very highly meshed. The number of locations in the NEM which are served by just one transmission link is relatively small. If the NEM is relatively meshed, it follows that there is no necessary link between the physical location of a constraint on the transmission network and the boundaries of the pricing regions created by that constraint.

83. Nevertheless, there is a common misconception in the NEM that "region boundaries should be aligned with the location of material pinch-points in the network". The view seems to be held that region boundaries should intersect transmission links which are regularly binding, whereas transmission constraints which never bind, or only bind occasionally, should be located within regions.

84. This notion is incorrect. In a meshed network, such as the NEM, the presence of a single binding constraint can give rise to separate prices at a large number of locations. Some of those separately-priced locations may be located hundreds or thousands of miles from the physical constraint. This is discussed further in box 1.

85. The notion that region boundaries should be aligned with material pinch points on the transmission network is, at best, highly misleading. In a meshed network there is no necessary connection between the desirable administrative pricing regions and the physical location of binding network constraints. Rather than aligning region boundaries with the location of transmission constraints, administrative region boundaries should be aligned with the boundaries of the pricing regions that arise from all material constraints. Put another way:

> Two locations on the transmission network should be placed in separate regions if they would be separately priced by a material binding constraint.

Box 1: Should region boundaries be chosen to align with major "pinch points" on the network?

There is a view in the NEM that the geographic location of region boundaries should intersect those transmission links which are regularly constrained. This notion dates back at least as far as the Parer report (2003), which stated:

> "Regional boundaries should be located at natural 'pinch points' in the network".[39]

Snowy Hydro may also have had the same concept in mind when it noted that one of the benefits of placing a region boundary between Murray and Tumut in the Snowy region would be that:

> "By adjusting regions to align with significant pinch points of congestion in the transmission network, pricing signals are transparent and reflect the actual physical situation."[40]

The AEMC, in its recent draft decision on region boundaries for the Snowy region explicitly observes that, under the original design of the NEM:

> "Regional boundaries were intended to reflect major 'pinch points' of transmission congestion, in order to promote efficiency in dispatch, trading and investment".[41]

The notion that the geographic boundaries of pricing regions should be chosen so as to bisect points of material congestion in the transmission network is an intuitive one. Intuitively, it seems natural that points of material congestion should separate the transmission network into regions of separate prices. However, this intuition is potentially highly misleading. Reliance on this intuition may easily give rise to policy prescriptions which are incorrect.

In fact, as discussed in the text, in a meshed network, a single transmission constraint can give rise to the need for separate pricing at *every* node in the network. In a meshed network, two nodes should receive a different price (that is, should be in different regions), if there is a transmission constraint on *any* path between those two nodes.

By way of example, in the NSW region of the NEM there is a substantial network loop which includes all the large generating units, known as the "Western Ring". A constraint at any location on that loop gives rise to separate pricing at all the generating units on that loop. In the NEM there are many hundreds of constraints which give rise to separate pricing of all the generators on the "Western Ring".

In a meshed network the administrative region boundaries should not be aligned with the physical binding constraints on the network. Rather, in a meshed network the administrative region boundaries should be aligned with the boundaries of the hypothetical pricing region boundaries that would arise in a nodally priced market, when a material constraint is binding.

How serious is the mis-pricing problem?

86. Just how serious is this mis-pricing problem? This is a question which has many dimensions. Earlier we noted that mis-pricing has implications for short-term efficiency of dispatch, for hedging, and for longer-term generator location decisions. Assessing the impact of

---

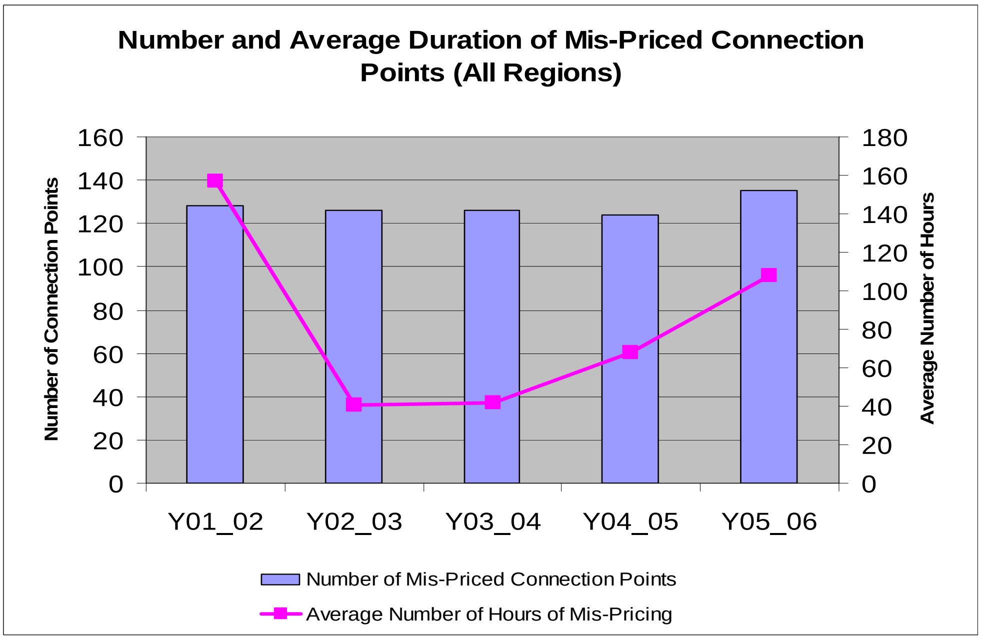
[39] Warwick Parer (chair), "Towards a Truly National and Efficient Energy Market", Council of Australian Governments, Energy Market Review, Final Report, December 2002.

[40] Snowy Hydro proposal, letter to the AEMC 11 November 2005, page 6.

[41] The AEMC goes on to observe that, depending on the proposal adopted, region boundaries may need to change in the future to correspond to new points of congestion on the transmission network: "For example, in the Split Region Option, the precise location of the boundary between the Tumut region and the NSW region may need to change to reflect the most pressing points of congestion". AEMC (2007), page 64.

mis-pricing can therefore involve looking at the impact on short-term dispatch efficiency, the hedging market, or on generator location decisions.[42]

87. First, we might ask the following: just what is the incidence of mis-pricing? How many connection points are mis-priced and how often? One study has been carried out by Biggar (2006c). This work was subsequently reproduced and extended by NEMMCO. The following diagram summarises, at a high level, the results found by NEMMCO. As this diagram shows, across the NEM there are around 120-130 connection points which are mis-priced (out of around 284 generator connection points in total) and the average number of hours of mis-pricing per connection point has dropped from around 160 hours per connection point in 2001/02 to a base of around 40 hours per connection point in 2003/04, before climbing to more than 100 hours in 2005/06.[43]



88. This analysis only shows the frequency and duration of mis-pricing but says nothing about its economic impact. To assess the economic impact we need to look further at the implications for the short-term efficiency of dispatch, as well as the impact on the hedge market and on generator location decisions.

89. Assessing the short-term economic efficiency impacts of mis-pricing is somewhat difficult. The economic impact of mis-pricing arises from the distorted bidding incentives. Therefore, any assessment of the impact of mis-pricing must determine how each mis-priced

---

[42] ERIG (2006), page 115 noted that some of its submissions commented on the significance of the mis-pricing problem. "A direct consequence of the NEM regional design is that intra-regional congestion is not currently priced. Many stakeholder submissions indicated that the current regime for providing locational signals for generation investment is flawed. The Energy Networks Association (ENA) considers that the approach adopted by ERIG should include the introduction of some level of cost reflective, locational pricing for all generators".

[43] See Biggar (2006c). The study estimated that it would take considerably more regions (around 70) to eliminate all (or virtually all) mis-pricing in the NEM. This observation is consistent with the comment by CRA from 2002 in which they note that: "Clause 3.1.4(b) [of the Electricity Code] requiring 'consistency between central dispatch and pricing' is, in the limit, incompatible with the concept of a regional market, as implemented in the NEM. In order to be consistent with intra-regional constraints, which must be recognised in physical dispatch, pricing would have to vary within each region, which cannot happen in this market design. Taken as an absolute, this clause could be interpreted as precluding anything other than a full nodal market". CRA (2002), page 31.

generator *would have* bid, *but for* the mis-pricing. This requires some assumptions about the bidding behaviour of generators.

90. Work along these lines has been carried out by IES (2004) and IES (2006). In both cases these studies focused on quantifying the economic efficiency benefits that would arise from improving the price signals in the Queensland region of the NEM. The 2004 report concluded:

> "on the basis of our estimates of dispatch costs for 2005/06, a change in the market rules from regional pricing to nodal pricing yields as much benefit to the market as the amount of transmission investment that would be required to eliminate half the dispatch costs due to intra-regional transmission constraints in QLD"[44]

91. The 2006 study by IES focused on the "dynamic efficiency" benefits from correcting the mis-pricing – particularly the impact on generator location and investment decisions. That study found a benefit to the market with a present value of more than $200 million. [45]

Mis-pricing and market power

92. We have seen that generator mis-pricing in the NEM may lead generators to distort their bids, leading to inefficient dispatch, contracting and location decisions. But how does this problem relate to the problem of generator market power? Is the distorted bidding arising from mis-pricing just a special case of the distorted bidding that arises from market power? Would enhancing competition reduce the problem of generator mis-pricing? Can mis-pricing reduce the incidence of market power?

93. It is well known that market power may induce generators to distort their bids. Normally, generator market power would induce a generator to offer its output at a price *above* its true marginal cost. However, in the presence of transmission constraints it is possible that market power may also induce a generator to offer its output at a price below its true marginal cost.[46]

94. We have seen above that mis-pricing also gives rise to distortion in generator bids. Is it the case, then, that the distortion due to mis-pricing is just a special case of the distortion arising from market power?

95. This suggestion has been made recently by the AEMC, and earlier by CRA. In the recent draft determination the AEMC refers to the distorted bidding that arises from mis-pricing as "strategic bidding" – the same phrase it uses to refer to the exercise of market power. The AEMC notes:

> "The Commission understands that such strategic bidding may not only be compatible with a workably competitive market, but that measures to directly curb strategic bidding may have unintentional and detrimental implications …"[47]

96. In a 2004 paper, CRA recommended that any concerns about what it referred to as "disorderly bidding" should be addressed by referral to the relevant competition authorities, suggesting that disorderly bidding is primarily a competition problem.[48]

97. However, the notion that the distorted bidding arising from mis-pricing is a special case of the exercise of market power is incorrect. In fact, the distortion in generator bids brought about by mis-pricing is completely unrelated to the distortion brought about by market power. If,

---

[44] IES (2004), page 33.
[45] To my knowledge these studies have not yet been subject to critical review and critique.
[46] See, for example, the analysis by Joskow and Tirole (2000).
[47] AEMC (2007), page 29.
[48] CRA (2004), page 8.

at the prevailing regional price, a group of mis-priced generators would like to increase or decrease their output by, say, 1000 MW, they will take whatever bidding actions they can to achieve this outcome – whether there is just one mis-priced generator who wants to change its output by 1000 MW, or one hundred mis-priced generators, each seeking to change its output by 10 MW. "Competition" per se has no impact on the economic impact of the mis-pricing problem.

98. The same points were made to the MCE in 2004 in response to the CRA report:

> "The question is whether such 'disorderly' bidding is properly regarded as an exercise of market power. Market power generally refers to an ability to 'give less and take more'. However, … a generator will have incentives to engage in this sort of behaviour whether one or 100 generators are located in the same sub-region or area – in other words, an increase in the number of competitors would not relieve the problem; in fact, it would worsen it. Therefore, it is questionable whether this behaviour constitutes an exercise of market power, at least in its commonly used meaning.
>
> The Group believes that disorderly bidding may lead to dispatch outcomes that are inconsistent with economic efficiency. However, it is incorrect for this bidding to be labelled an exercise of market power."[49]

99. The above quote suggests that an increase in the number of mis-priced generators might make the mis-pricing problem worse. Is there some link between the number of mis-priced generators (which might be a proxy for competition) and the economic impact of the mis-pricing problem?

100. There are two issues which are worth highlighting. First, as noted above, for a given volume of mis-pricing, it doesn't matter whether that volume is offered by a single generator or many small generators. That generator or generators will seek to distort their bids to bring their dispatch back into line with the price they are paid – or, if they are not successful, will shift their bid to the allowed offer-price ceiling ($10,000) or floor ($-1000). Holding constant the total volume of mis-pricing, an increase in the number of mis-priced generators has no impact on the economic outcome.

101. Of course, if increasing the number of mis-priced generators also increases the total amount those generators would like to shift their output, this will make the mis-pricing problem worse. This is discussed further box 2 below.

102. There is a second economic impact of mis-pricing. Since all the mis-priced generators distort their bids to the same level, the dispatch engine no longer has accurate information on the relative cost of the mis-priced generators – it may dispatch them all equally even though some of the mis-priced generators are much more expensive than others. An increase in the number of mis-priced generators – if that increases the dispersion in the underlying costs of production – will therefore reduce the efficiency of the outcome.

103. Broadly speaking however, there is no immediate link between generator competition and the extent of the mis-pricing problem. Mis-priced generators are essentially price-takers, whether there is just one, or many affected generators.

---

[49] Group submission (2004), page 7-8.

Box 2: Does Promoting Competition Promote Efficiency?

In most markets, taking policy steps to enhance competition will improve market outcomes – that is, it will lower prices, increase output, improve product quality, improve the quantity and quality of investment and so on. Does this apply in the context of a regionally-priced electricity market?

It is often asserted that policies that would (at least in other markets) normally be expected to enhance competition should be pursued. For example, The AEMC notes that Snowy Hydro claim that its proposal to abolish the Snowy region would promote the NEM Objective by:

> "reducing cost to customers by improving the incentives on Tumut to increase available generation and allowing Tumut generation to compete on a equal footing with "western ring" generators".[50]

The AEMC draft determination similarly notes:

> "The Snowy Hydro proposal allows Tumut generation to compete, as it were, on a 'level playing field' with these generators, who already have incentives to bid below cost under these circumstances"[51].

In "normal" markets, we would expect that allowing generators to compete on an "equal footing" would enhance competition. But, can we be sure that putting generators on an "equal footing" in a market which is periodically mis-priced will enhance competition? Even more importantly, can we be sure that an observed increase in output and reduction in the market price is a sign of a more efficient market?

In fact, if some generators are currently mis-priced, a change in the region boundary so as to place some other generators "on an equal footing" will increase the number of mis-priced generators. Normally this would increase the total amount that the mis-priced generators would like to shift their output – increasing the size of the distortion to the overall market.

Suppose that a change in a region boundary caused generators which are currently correctly priced to be constrained off. We would expect that these generators would have an incentive to offer their output at a low price. It is likely that the output of these generators will thereby increase, displacing either generators which were already constrained off (with ambiguous welfare consequences) or displacing generators elsewhere in the market. To the extent that the increased output of these newly-mis-priced generators displaces generators elsewhere in the market, it must lower prices elsewhere. In consequence, it appears that output has increased, and prices have gone down. At the same time, higher-cost generation from these mis-priced generators is displacing lower-cost generation elsewhere in the market, reducing the efficiency of dispatch.

Although it might appear that an increase in output and a reduction in price is a sign of increasing economic efficiency, this interpretation would, in this case, be incorrect. Instead, the increase in mis-pricing (to the extent that it displaces correctly-priced generation elsewhere) leads to a reduction in the efficiency of dispatch.

Competition is a means to an end. It is not an end in itself. Competition is the means to the end of efficient outcomes. Although in most markets, promoting competition will promote efficiency, in markets where there are market failures, competition may come into conflict with the achievement of efficiency objectives. This is known as the theory of the second best – in a market with at least one market failure, correcting a second market failure will not necessarily improve overall outcomes.

In the National Electricity Market, it may be possible to engineer a change in region boundaries which increases output and lowers prices. However, to in every case interpret those changes as evidence of greater "competition", synonymous with greater economic efficiency, would be incorrect. Unfortunately, in the National Electricity Market it is not possible to use the "promotion of competition" as a short-hand or a rule-of-thumb, equivalent to "the promotion of efficiency".

---

[50] AEMC (2007), page 2-3.
[51] AEMC (2007), page 34.

104. The presence of mis-pricing can, under certain circumstances, limit the exercise of market power. A generator which is constrained on or off will typically have little or no influence on the price it receives. Such a generator has little or no market power. In this case it is theoretically possible that improving the price signals on such a generator may allow that generator to exercise any market power it might have.[52]

105. At the same time, however, it is important to note that where a generator is constrained on and its output is required for reliability purposes, that generator is again able to exercise its market power – but this time through the compensation it seeks in, say, a network support agreement. Indeed, in this case, the regional-pricing arrangements accentuate such a generator's market power by eliminating the incentive for any local load to respond to the local high price signal at such times by curtailing its demand.[53]

106. In addition, it is widely accepted that a generator's incentive to exercise market power depends strongly on its financial contracting position. As we have seen, generator mis-pricing may affect a generator's hedging decision in two ways – either directly, by increasing the dispatch risk that a constrained-on or constrained-off generator faces, or indirectly, by affecting the effectiveness of the inter-regional settlement residues as a hedging device. It is possible that correcting the mis-pricing problem could have a material impact on hedge positions of some generators, materially reducing market power.

107. Nevertheless, the possibility remains that improving the price signals on generators may increase their ability to exercise market power. This may, in turn, have an impact on the hedge market – in particular, by reducing liquidity in the hedge market. The impact of market power on hedge market liquidity is an issue which deserves future research.

Mis-pricing and constraint equations

108. Importantly, it turns out that for a given configuration of administrative pricing regions and interconnectors, it is very easy to determine precisely which nodes will be mis-priced (and whether they are constrained on or off) simply by observing the form of the "correctly oriented constraint equations" for that configuration.

109. To see this, it is necessary to introduce the concept of "constraint equations". As we noted earlier, in the NEM, the market prices and the output targets of all scheduled generators and loads are determined by a central computer system known as the dispatch engine. The dispatch engine computes the combination of output targets (the "dispatch") which maximises short-term economic welfare subject to two sets of constraints:

(a) the "energy balance equations" which require that the energy injected into each administrative pricing region plus the flow in on any notional interconnector must equal the total load in that region plus the flow out on any notional interconnector; and

(b) the "constraint equations" which reflect the physical limits on the elements of the real, physical transmission network.

[52] This assumes that a generator is only present at a single location on the network. When a generator is located at more than one location, including the regional reference node, it may be able to use any market power it has at times of binding transmission constraints to raise the price it receives at *all* its other locations, enhancing its incentive to exercise market power. In this case, greater geographic differentiation of prices may reduce the incentive to exercise market power.
[53] See Harvey and Hogan (2000)

110. Each constraint equation is represented in the dispatch engine in the form "left-hand-side" "less than or equal to" [54] "right-hand-side". In the current formulation of the NEM dispatch engine, these constraint equations are all required to be *linear*. In fact, the left-hand-side of the constraint equation is required to be of the form:

(a) for each generator[55], a simple fixed number (known as the coefficient) times the output of that generator[56] plus

(b) for each interconnector, a simple fixed number times the flow on that interconnector.

111. The sum of these left-hand-side terms must be less than or equal to a number computed by the dispatch engine which is known as the constraint "right hand side". In mathematical notation, all the constraint equations in the NEM dispatch engine can be written in the form:

$$\alpha_1 Q_1 + \alpha_2 Q_2 + ... \alpha_N Q_N + \beta_{A \to B} F_{A \to B} + \beta_{B \to C} F_{B \to C} + ... \leq RHS$$

Where:

α_i is the coefficient in the constraint equation on generator *i*, Q_i is the output of generator *i*, $\beta_{A \to B}$ is the coefficient in the constraint equation on the interconnector from region A to region B and $F_{A \to B}$ is the flow on the interconnector from A to B and so on; RHS is a number representing the physical limit in the network.

112. For example, the following are examples of hypothetical constraint equations:

- In a network with three regions "VIC", "SA" and "SNY" and interconnectors, VIC-SA, and VIC-SNY, there might be a constraint equation of the form:

  $$0.3F_{VIC \to SA} - 0.8F_{VIC \to SNY} \leq 1000$$

- Similarly, in a network with generators "BAY, "ERA", "MTP", "WW" and an interconnector SNY-NSW, there might be a constraint equation of the form:

  $$0.3Q_{BAY} + 0.5Q_{ERA} + 0.2Q_{MTP} - 0.3Q_{WW} + F_{SNY \to NSW} \leq 500$$

113. There are infinitely many ways of representing the same physical limit in the network in the form of a constraint equation. However, for any given configuration of region boundaries and interconnectors there is a unique way to represent the constraint equation so as to yield the correct pricing outcomes. When a constraint equation has been formulated in such a way as to yield the correct pricing outcomes, it is said to be "correctly oriented".[57]

114. It turns out that a constraint equation is "correctly oriented" if both:

---

[54] In fact, the NEM dispatch engine also allows constraints to be expressed in the form of "equals" or "greater than" constraints – but these can be expressed as a "less than" constraint without loss of generality.

[55] More strictly, for each "connection point". But since there is a virtual one-to-one correspondence between generators and connection points, the difference here is small.

[56] More strictly, the "net injection" at that connection point which is equal to any generator production less local load at that connection point.

[57] More information on constraint orientation can be found in CRA (2003).

(a) the constraint equation is formulated in such a way that, in all regions, the coefficient on the net injection of power at the regional reference node for that region is zero;

(b) the sum of the coefficients on the interconnectors around any loop that exists between the notional interconnectors is zero. [58]

115. Importantly for our present purposes, when a constraint equation has been correctly oriented for that configuration, we can, by observing the form of the constraint equation, make some observations about the mis-pricing and hedging problems that will arise when that constraint equation binds.

116. In particular, inspection of the form of the correctly-oriented constraint equation allows us to predict, when that constraint equation binds:

(a) which connection points which will be mis-priced, and whether the generators at those connection points will be "constrained on" or "constrained off";

(b) whether or not the inter-regional settlement residues will be "firm" and, to an extent, whether or not there is a risk of negative settlement residues due to loop flow.

117. Let's focus here on the use of constraint equations to determine which generators will be mis-priced when a given constraint binds.

118. As part of the dispatch process, in addition to computing the prices, dispatch and flows, the dispatch engine also computes, for each constraint equation an associated "constraint marginal value". The constraint marginal value is zero unless the constraint is binding (that is, affecting dispatch), in which case the constraint marginal value is positive.

119. As mentioned earlier, under the current NEM arrangements, the NEM dispatch engine determines a single price for electricity in each region (the regional reference price, which is the same as the locational marginal price at the regional reference node). However, it is possible to work out the locational marginal price at every node within a region using information on the constraint marginal value of each binding constraint and the coefficient of that node in the corresponding binding constraint.

120. Specifically, suppose that the constraint marginal value on the *n*th constraint equation is λ^n, and the coefficient on node i in the *n*th constraint equation is α_i^n. The locational marginal price at node *i* is then related to the regional reference price for that region as follows:[59]

$$P_i = RRP - \sum_n \lambda^n \alpha_i^n$$

121. Since the constraint marginal value is zero when the constraint is not binding and positive otherwise we can deduce the following:

(a) A node or "connection point" in the network is mis-priced if and only if there is a constraint binding which has a non-zero coefficient on that connection point.

[58] The introduction of loop paths around notional interconnectors imposes a special requirement for constraints to be correctly oriented. See Biggar, "Orienting Constraints in a Network with Interconnector Loops", 24 November 2006.

[59] The same equation can be found in Biggar (2006b), page 19, Biggar (2006c), page 4, Snowy CSP/CSC trial (chapter 8, part 8A of the Rules, subsection (j)), CRA (2005), page 4.

(b) A generator at that connection point is "constrained off" when a given constraint equation is binding if the coefficient on that connection point in that constraint equation is positive. Similarly a generator at that connection point is "constrained on" when a constraint equation is binding if the coefficient on that connection point in that constraint equation is negative.

122. For example, consider again the hypothetical constraint equation set out above: $0.3Q_{BAY} + 0.5Q_{ERA} + 0.2Q_{MTP} - 0.3Q_{WW} + F_{SNY \to NSW} \leq 500$.

123. When this constraint equation binds, the generators at BAY, ERA and MTP will all be constrained off and the generators at WW will be constrained on. In contrast, when the following constraint equation binds: $0.3F_{VIC \to SA} - 0.8F_{VIC \to SNY} \leq 1000$, no generators will be mis-priced since there are no generator or connection points terms in this constraint equation with a non-zero coefficient.

124. As another example, consider again the network configuration set out in figure 1. In that network it turns out that the correctly-oriented constraint equation for the network limit between nodes 1 and 2 is as follows:

$$Q_2 - F_{A \to B} \leq RHS$$

Where Q_2 is the net injection of power at node 2, $F_{A \to B}$ is the flow on the interconnector from region A (nodes 1 and 2) to region B (node 3), and RHS is the physical limit on the flow between node 1 and node 2 in the direction of node 1.

125. In this constraint equation, since the coefficient on Q_2 is positive and non-zero it follows that the generator at node 2 will be mis-priced and will be constrained off, as shown in figure 1.

126. In exactly the same way, in the network of figure 2, the correctly oriented constraint equation for the network limit between nodes 2 and 3 is $-Q_2 - F_{A \to B} \leq RHS$. Now, since the coefficient on Q_2 is non-zero and negative, it follows that the generator at node 2 will be mis-priced and will be constrained on, as shown in figure 2.

127. Clearly, it is very easy to determine which connection points will be mis-priced when a given physical limit is binding simply by observing which connection points have non-zero coefficients in the corresponding correctly-oriented constraint equation.

128. One consequence of this observation is that, if we have information on the frequency with which certain constraints bind, we can use that information to determine the frequency and duration of mis-pricing at different connection points in the NEM. This observation was the basis of the analysis in Biggar (2006c) as discussed earlier.

129. In practice, just how many constraint equations in the NEM constraint library have generator terms on the left-hand-side? There are tens of thousands of constraints in the NEMMCO constraint library, many of which are out-of-date and no longer used. Many of the others have never been binding. Let's restrict attention to just those constraint equations which

bound at least once during the 2005/06 financial year. There are 747 such constraint equations.[60] Of these 443 (59.3%) have at least one generator term on the left-hand-side.

130. These constraints accounted for 45% of the total constraint-minutes binding[61]. In other words, during 2005/06 around half the time that any constraint was binding, the binding constraint gave rise to mis-pricing at least one generator connection point.

131. A number of these constraint equations have a large number of generator terms on the left-hand-side. In fact, 109 (14.6%) have more than *ten* generator terms on the left-hand-side. Some of the more significant of these constraint equations are set out in the table below:

Constraint Equation ID	**Description**	**Binding for (hours)**	**Number of Generator Terms**
Q^NIL_1CS	QLD Central-South transfer limit	5.4	35
Q_CS_1700 and Q_CS_1900	QLD Central-South transfer limit	109.1	35
N>>N-81_1T	Outage on 81 line, avoid 82 line overload	77.2	23
N>>N-NIL_1N	No outages; avoid 82 overload on 81 line trip	52.6	21
N>>N-NIL_1T	No outages; avoid 82 overload	18.8	21
V>V1NIL	Avoid overload on Hazelwood 500/220 kV 2,3,4 transformer	14.0	18
V>>H_NIL_2_R	Prevent overloading on South Morang 500/330 kV transformer	80.3	14
V::H_NILQF_BL_R	Prevent transient instability in the event of Hazelwood to South Morang line trip	32.8	13
T>T_PMSH_220_1	Outage on Palmerston to Sheffield line; avoid Georgetown to Sheffield overload.	22.6	11

Summary of the mis-pricing problem

132. In brief, this section makes the following key points:

- A binding constraint can be thought of as dividing up the network into hypothetical regions of uniform nodal prices. Mis-pricing arises when the boundaries of these hypothetical regions do not align with the administrative pricing region boundaries. Mis-priced generators may find that when they offer their output at a price that reflects their true cost, they are dispatched to a quantity which is above or below the quantity they

[60] Excluding constraints which start with "@" (generator fixed load), "#" (discretionary constraints) and "F" (FCAS constraints). Constraints with different "version number" and "effective date" are counted as separate constraints even if they have the same constraint id.

[61] If there are 2 constraints binding for ten minutes, that counts as 20 constraint-minutes.

would like to produce at the prevailing regional reference price. These generators have an incentive to distort their bids, to use other mechanisms, or to simply fail to follow dispatch instructions, in order to increase or decrease the amount for which they are dispatched.

- Mis-pricing reduces the short-term efficiency of dispatch, imposes dispatch risk on mis-priced generators, reduces the firmness of settlement residues (described later) and gives rise to inefficient generator location decisions.

- In a meshed network, a single binding constraint on a loop will give rise to a large number of hypothetical pricing regions. There is no necessary relationship between the boundaries of these regions and the physical location of the constrained link. In a meshed network administrative region boundaries there is no necessary relationship between the boundaries of the administrative price regions and the physical location of a constraint.

- The exercise of market power will also give rise incentives on generators to distort their offers away from the true cost. However the distortion due to the exercise of market power is not the same as the distortion due to mis-pricing. The distortion due to mis-pricing is not reduced by increasing the number of generators affected and, in fact, will generally be made worse by increasing the number of mis-priced generators. Placing generators on an "equal footing" or allowing generators to compete on "a level playing field" will not necessarily improve economic welfare.

- It is straightforward to identify which generators will be mis-priced when a given constraint binds merely by examination of the correctly-oriented form of the corresponding constraint equation. During 2005/06 around of half the time that a constraint was binding, at least one generator connection point was mis-priced.

1.2 The "hedging" problem

133. In the previous section we saw that the mis-pricing of certain generators, which arises under the current regional pricing arrangements in the NEM, leads to incentives for distorted bidding and thereby distorted pricing and reduced short-term efficiency of dispatch.

134. But spot market outcomes are only one factor affecting the decisions of electricity producers. A large proportion of the electricity traded in the spot market is covered by financial risk-management or "hedge" contracts. In the longer term, hedge contract prices potentially have a larger impact on generator investment and location decisions than wholesale spot market prices. An efficient electricity market requires not only an efficient wholesale spot market but also an efficient "forward" or hedge market.

135. What does it mean for the hedge market to be efficient? We might say that a hedge market is efficient if the prices of the hedge contracts accurately reflect the expected future spot prices.

136. Under what conditions will a hedge market be efficient? This is a question on which further research is required. Even putting aside the impact of transmission constraints, hedge market efficiency will depend on many factors, such as the level of transactions cost, bid-ask spreads, and the level of competition on both the supply and demand side of the market. Hedge market efficiency may also depend on the availability of special hedging products such as insurance against generator outages.

137. However, for our purposes we are interested in the impact of transmission constraints on hedge market efficiency. As we have seen, in the presence of transmission constraints, price differences arise between generators at different locations. The presence of price differences between locations in an electricity market gives rise to a flow of funds to the system operator, known as "settlement residues", "congestion rents", "constraint rentals", or "merchandising surplus".

138. If the locational marginal price at node *i* is P_i and the net injection of power (that is the power produced less the power consumed) at node i is z_i , the merchandising surplus or settlement residues are equal to $-\sum_i P_i z_i$.

139. Under what conditions will the hedge market be efficient in the presence of transmission network constraints? We might conjecture that hedge market efficiency depends on both:

(a) access, by market participants, to the total settlement residues or merchandising surplus in a manner which facilitates hedging; and

(b) access to financial instruments which allow for hedging against transmission network outages.

140. Confirming that both of these conditions are required for hedge market efficiency is a question for future research. The following box demonstrates that market participants need access to the total residues in order to hedge the risks of inter-regional trading.

Box 3: Access to the total settlement residues is a precondition for efficient hedging

To see that access to the total settlement residues is required for hedge market efficiency, consider the following. Suppose that the generators at node *i*, produce a (time-varying) quantity Q_i^S, incurring a cost $c_i(Q_i^S)$. Putting aside all hedging arrangements, these generators sell their output at the wholesale spot price P_i and therefore receive a (time-varying) profit $\pi_i^G = P_i Q_i^S - c_i(Q_i^S)$. Now consider the retailers located at node *i*. These retailers sell a (time-varying) quantity Q_i^D, receiving a revenue from the end-customers $R_i(Q_i^D)$. Putting aside all hedging arrangements, these retailers purchase all their electricity requirements at the wholesale spot price P_i and therefore receive a (time-varying) profit $\pi_i^R = R_i(Q_i^D) - P_i Q_i^D$.

The total (time-varying) profit of the industry as a whole (generators plus retailers) is therefore:

$$\sum_i \pi_i^R + \sum_i \pi_i^G = \sum_i R_i(Q_i^D) - \sum_i c_i(Q_i^S) + \sum_i P_i(Q_i^S - Q_i^D) \qquad \ldots(1)$$

Now, in the absence of any transmission constraints and losses, the price paid by retailers is the same as the price paid to generators. Since the quantity of power produced must equal the total quantity of power consumed, the last term in the expression above disappears and we have that the total profit in the industry is simply

$$\sum_i \pi_i^R + \sum_i \pi_i^G = \sum_i R_i(Q_i^D) - \sum_i c_i(Q_i^S) \qquad \ldots(2)$$

The variability in this total profit represents the total amount of risk which cannot be hedged by contracts within the industry alone (it may still be possible to hedge some of this risk through contracts with parties outside the industry – such as financial institutions, insurance companies, input suppliers or downstream customers).

Now consider the impact of transmission congestion. The question we need to ask is the following: to what hedge instrument do market participants need access, in a market with transmission constraints, in order to be able to obtain the same hedging outcomes as in a market with no transmission constraints? Comparing equations (1) and (2) above we see that market participants need access to a hedging instrument with a payout equal to $\sum_i P_i(Q_i^S - Q_i^D)$.

We can define the net injection of power at node i as $z_i = Q_i^S - Q_i^D$. Then we can see that in order to fully hedge the impact of the transmission constraints, participants need access to an instrument with a payout $\sum_i P_i z_i$ - just equal to the total settlement residues.

141. As conjectured above, it is important not just that the market participants have access to the total settlement residues – but that they have access to the total residues in a form which is suitable for hedging purposes. The second key problem with the current NEM arrangements – that I have called the "hedging" problem - arises from the fact that the total residues are not at present made available to market participants in a form that can be used for hedging.

142. A hedging instrument is "firm" if a market participant can, by purchasing a fixed share of that instrument in advance, obtain a perfect hedge for a transaction which involves purchasing a fixed quantity of electricity in the forward market in one region and selling the same fixed quantity of electricity in the forward market in another region.

143. If a market participant purchases, say, 100 MW of electricity in the forward market in region B and sells 100 MW of electricity in the forward market in region A, that participant requires a hedging instrument with a payment stream equal to the difference in the spot price between those regions times 100 MW. Therefore, a hedging instrument is firm if it yields a payoff equal to the spot price difference between those two regions multiplied by a fixed number. If such an instrument were available, the market participant could calculate the precise share of the instrument that it needed to purchase in order to obtain the payoff required for a perfect hedge.

144. Under the current NEM arrangements there is a tool for hedging the risk of inter-regional trading – in the form of payment streams known as the "inter-regional settlement residues" or IRSRs. The inter-regional settlement residues for an interconnector are equal to the flow on that interconnector multiplied by the price difference between the regional reference price in the importing region less the regional reference price in the exporting region.

145. However, as we will see, under the current NEM arrangements the IRSRs are not (always) an effective instrument for hedging inter-regional trading risk. Under the current arrangements the IRSRs are not a "firm" hedging instrument, in the sense defined above.

146. The reason is that, under the current NEM arrangements, although the total residues are "firm"[62], some of these total residues may be paid directly to (or from) certain mis-priced generators while some other parts of the total residues may be spread over two or more of the inter-regional settlement residue funds. As we will see, dispersing the total residues in this way makes it difficult or impossible for market participants to obtain the "firm" hedging instruments they require.

Box 4: Inter-Regional Trading or Inter-Regional Hedging?

In its recent draft determination, the AEMC makes frequent reference to the impact of various proposals on inter-regional hedging. Indeed, one of the AEMC's decision criteria is "the likely effect of the proposal on inter-regional trading and risk-management"[63].

Intuitively, it seems desirable that a given policy should seek to "reduce the risks of inter-regional trading". There is, however, a potential pitfall in understanding here which it is useful to highlight, so it can be avoided.

There is no need for access to the inter-regional settlement residues (or any form of residues) in order to perfectly hedge a transaction which involves the purchase or sale of forward contracts in another region. For example, consider the case of a generator in region A who wishes to sell a 100 MW swap contract in region B. That generator can obtain a perfect hedge for that transaction (i.e., not face any risk at all) by hedging that transaction with a portfolio which involves buying an equivalent swap in region B and selling a swap in region A.

In other words, any market participant can engage in riskless "inter-regional trading" of hedge contracts with or without the inter-regional settlement residues. The value of a firm inter-regional hedge (such as "firm inter-regional settlement residues) is therefore not that it reduces the risks of "inter-regional trading" of hedge contracts. What then is the value of a "firm" inter-regional hedge?

Different regions in the NEM will, at different times, have an "imbalance" in the supply or demand of hedge contracts. Regions which are net exporters will tend to have a surfeit of generators wishing to offer swap contracts. This will tend to drive down hedge prices in those regions, relative to the expected future spot price. Conversely, regions which are net importers will tend to have a shortage of hedge contracts, driving up hedge prices in those regions relative to the expected future spot price.

Deviations between hedge contracts and the expected future spot prices create arbitrage opportunities. Speculators (who may be the existing market participants) may then enter this market, buying hedge

[62] This is discussed further in box 5.

[63] AEMC (2007), page 8.

contracts in regions where the price is depressed and selling them regions where the price is at a premium. In carrying out this activity, speculators would normally be taking on some risk, and would normally be required to be compensated for that risk. We would expect that, as a result, the difference in the prices of these hedge contracts from their underlying expected future spot prices would not be completely eliminated.

However, if there is a "firm" hedge available, these speculators can carry out the activity of arbitraging differences in hedge contracts across regions without incurring any risk. As a result, the hedge contract prices would be expected to be perfectly aligned with the underlying expected future spot prices.

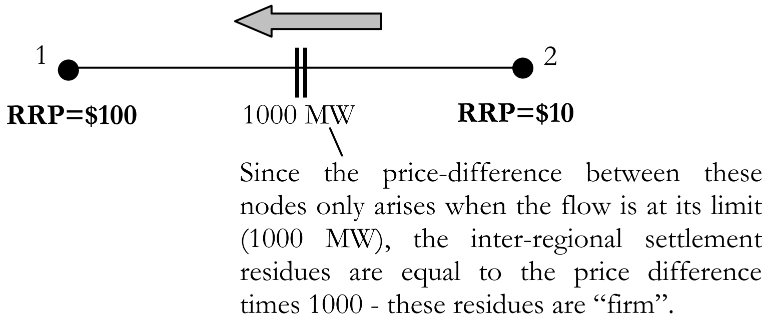
In summary, the benefits of improving the firmness of a hedging instrument is not that it reduces the risks of selling or buying a hedge contract in another region – thereby increasing the volume of such trade. Rather, as we have seen, such transactions can already be perfectly hedged. The benefit of improving the firmness of a hedging instrument is that it reduces the risks associated with arbitrage of differences in hedge prices across different regions – thereby more perfectly aligning the prices of those hedge contracts with their underlying expected future spot prices. The phrase "reduce the risks of inter-regional trading" should be used with care.

Lack of firmness in radial networks

147. The lack of firmness problem can be made clearer using simple network examples. Consider first the following simple two-node, one link network. Let's suppose that the transmission link has a fixed power flow limit of exactly 1000 MW.

148. In this simple network the inter-regional settlement residues (which are equal to the total residues) are firm. The reason is as follows: In this simple network, the prices at the two nodes are the same until the power flow limit on the transmission link is binding, at which point price differences can arise between the two nodes. Therefore, the inter-regional settlement residues (which are equal to the price-difference times the flow) are exactly equal to the price-difference times a fixed number (in this case 1000).

Figure 5: IRSRs in a simple two-node one-link network



149. Since the inter-regional settlement residues are precisely equal to the price-difference multiplied by 1000, a market participant seeking to hedge a transaction involving purchasing a 100 MW swap in region B and selling a 100 MW swap in region A can purchase a share of these residues (in fact, to obtain a 100 MW hedge, the market participant need purchase precisely one tenth of these residues), to obtain the payout that it requires to precisely hedge the risk that it faces from its transaction. Under the assumptions set out in box 5, in this simple network the inter-regional settlement residues are "firm".

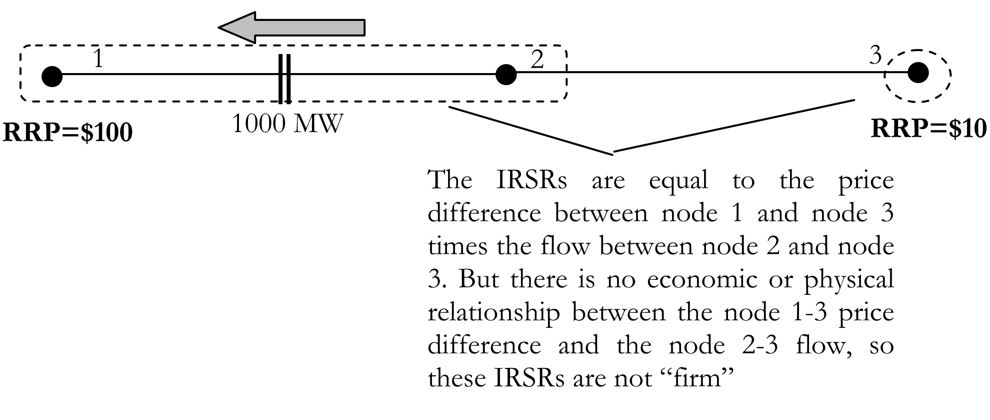
150. Now consider the slightly more complex network set out below (the same network as figure 1 earlier). In this network, the flow between node 1 and node 2 has reached its limit, which

gives rise to a price difference between the two regions. In this network the inter-regional settlement residues are equal to the price difference between these regions (the difference in price between node 1 and node 3) times the flow on the interconnector between the regions (between node 2 and node 3).

151. However, there is a clear problem – when a price difference arises between the regions (due to a constraint between node 1 and node 2) the flow between node 2 and node 3 could be at any level and in any direction (it could even be "counter-price"). There is no physical or economic relationship between the price difference between the regions and the flow on the interconnector.

152. As before, let's consider the case of a market participant who wishes to hedge a transaction consisting of the purchase of a 100 MW swap at node 3 and the sale of a 100 MW swap at node 1. What share of the inter-regional settlement residues should this market participant purchase? It is impossible to say. The required share of the IRSRs depends on the flow on the link between node 2 and node 3 at times of price difference between the regions. But, as we have seen, there is no physical or economic relationship between the price difference between the regions and the flow on the link between node 2 and node 3. [64]

Figure 6: Inter-regional settlement residues in a simple three-node network



Negative settlement residues

153. At times of price difference between node 1 and node 3 the flow between node 3 and node 2 could be zero or any positive number. In fact, it could even be negative. That is, flows between node 2 and node 3 could even be "counter-price". When flows are "counter-price" the inter-regional settlement residues are negative, giving rise to the problem of "negative settlement residues".

154. Under the present market arrangements, NEMMCO auctions the right to receive a one-way stream of payments corresponding to the residues arising on any given interconnector. These streams of payments are "one-way" in the sense that as long as the residues are positive, the holder will receive a positive payment from NEMMCO. But, in the event that these payments are negative, the holder is not obliged to make a payment to NEMMCO. As a result, if the residues become negative, NEMMCO risks incurring a substantial loss.

[64] ERIG (2006) summarises this problem as follows: "Because intra-regional congestion is not priced in the regional design of the NEM, there are implications for inter-regional trade as the volume risk associated with intra-regional congestion can manifest itself with the price risk of inter-regional congestion". As we will see shortly, the problem of lack of firmness is not just a problem arising from a lack of pricing of intra-regional congestion.

155. Although NEMMCO can tolerate small negative residues (perhaps by recovering them from positive residues at other times), large negative residues threaten the solvency of NEMMCO. For this reason, under the current NEMMCO policies, NEMMCO intervenes in the market, primarily by "clamping" (that is, limiting) the flows on those inter-connectors with negative settlement residues.[65]

156. Clamping is likely to directly reduce the efficiency of dispatch and significantly reduce the firmness of inter-regional settlement residues. Whether or not clamping is an efficient intervention depends, in part, on whether or not the counter-price flows are caused by mis-pricing. As we have seen above, mis-pricing reduces the short-term efficiency of dispatch. Clamping may partially offset some of that harm. However, as explained below, counter-price flows can also arise as an efficient outcome in a meshed network.

157. Negative settlement residues are just one, extreme, manifestation of the more general problem of lack of firmness discussed here. Negative settlement residues give rise to particular, acute problems in the market and therefore attract a disproportionate share of the attention. For this reason, negative settlement residues are sometimes viewed as being a problem in their own right. However, negative settlement residues are just an extreme form of the more general problem of lack of firmness.

Lack of firmness in meshed networks

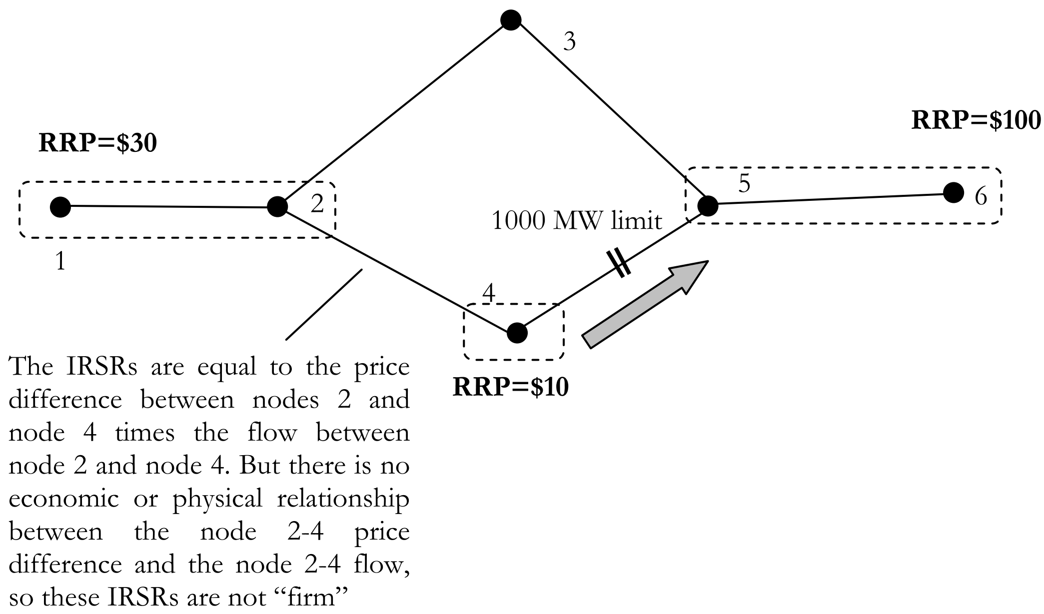
158. This problem of lack of firmness of IRSRs is not just a result of generator mis-pricing (as it was in the example above). In fact, the same lack of firmness arises in meshed networks *even when no generators are mis-priced.* To see this, consider again the looped network used in the earlier example.

159. As we saw earlier, when a constraint arises on a loop in a network, price differences arise at all the nodes on the network, regardless of the size or direction of the flow between those nodes. For example, in the network of figure 7 below, when the flow between nodes 4 and 5 reaches its physical limit, price differences arise at all of the nodes around the loop (nodes 2, 3, 4 and 5 in this example). These price differences will, in general, give rise to corresponding inter-regional settlement residues.

160. For example, in figure 7, inter-regional settlement residues will arise between node 2 and node 4 – but can these inter-regional settlement residues be used to hedge a transaction involving the purchase and sale of a fixed quantity of forward contracts at nodes 2 and 4? As before, we see that there is no necessary physical or economic relationship between the existence of a price difference between node 2 and node 4 and the direction or magnitude of the flow between node 2 and node 4. A market participant trading between node 2 and node 4 therefore cannot know the share of the IRSRs to purchase in order to obtain the hedge that he or she requires.

[65] See NEMMCO, "Operating Procedure: Dispatch: Document Number SO_OP3705", section 18, "Management of Negative Residues Process".

Figure 7: Inter-regional settlement residues in a meshed network



161. IRSRs are inherently not firm in the presence of loop flow – i.e., in a meshed network. Yet, as I noted earlier, transmission networks are inherently "meshed". As the number of pricing regions in the NEM increases, this problem of lack of firmness, and negative settlement residues, is likely to increase.

The economic harm from the hedging problem

162. The fact that the inter-regional settlement residues are not a "firm" hedging instrument in the financial or forward markets gives rise to real economic harm. This arises because, as noted above, in the absence of firm hedging instruments, traders cannot perfectly hedge inter-regional transactions in forward contracts. As a result, the prices of the forward or hedge contracts in each region will not be perfectly aligned with the underlying expected future spot prices.

163. The fact that the hedge prices are not aligned with the expected future spot prices has direct economic consequences. Generators will have a reduced incentive to locate or expand in generation-rich regions, even if there is adequate transmission capacity to export the resulting production and even if that location is the most efficient for the market overall. In addition, generators will have too strong an incentive to locate in demand rich regions, even if there is adequate import capability into that region and even if locating elsewhere is more efficient.

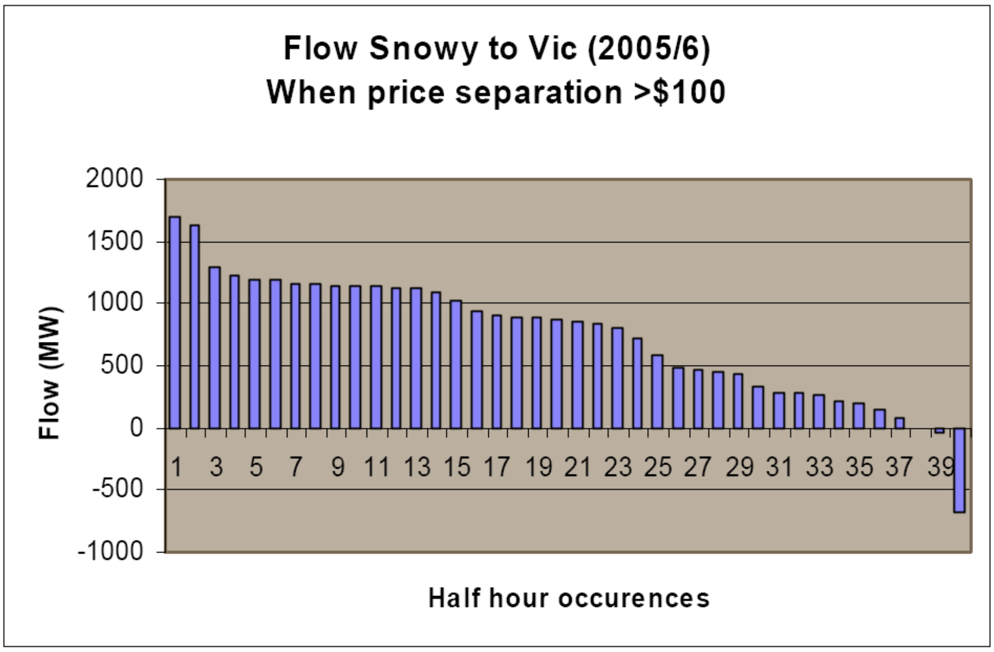
164. There may also be an impact on competition in the supply of hedge contracts. The existence of firm inter-regional hedging instruments allows traders to compete (up to the volume of the import capability) with local generators in the supply of hedge contracts. This may enhance competition, further driving prices down (in this case) in the hedge market.

165. There is a sense in which the hedging problem is really just another type of mis-pricing problem – however, in this case, not mis-pricing in the spot-market, but mis-pricing in the contract market. Since hedge prices reflect underlying expected future spot prices (but may trade at a premium or a discount to the expected future spot prices), efficient hedge pricing requires both (a) efficient underlying spot prices (the "mis-pricing" problem above) and (b) firm instruments for hedging inter-regional risk (the "hedging" problem discussed here).

166. Can we quantify the economic harm arising from the hedging problem? One first step is simply to explore the magnitude of the "lack of firmness" on the inter-regional settlement residues.

167. The report of the Energy Reform Implementation Group takes some first steps at quantifying the extent of the problem. They note that in submissions: "Energy Australia, EIIA, ERAA, MEU and Ergon … note that the lack of a reliable inter-regional hedging instrument creates risks for participants trading across regional boundaries". ERIG (2006) then goes on to examine the magnitude of the flow on different interconnectors at times of high price differences. As noted above, if the inter-regional settlement residue were firm, a price difference between two regions would only arise when the flow on the interconnector reached a fixed maximum limit. However, as the graph below makes clear, large price differences on the VIC-Snowy interconnector have arisen when the flow is at *any* level, from less than -500 MW to more than 1500 MW.[66]

Figure 2: Flows across Snowy-VIC when price separation between Snowy-VIC > $100 MWh



Source: ERIG (2006), page 174.

168. Further work is needed to determine whether or not it is possible to quantity the deviation in the hedge prices from the underlying expected future spot prices arising from this lack of firmness, and to quantify the economic harm from distorted generator and load decisions that results.

[66] Strictly speaking, price differences can also arise due to losses. ERIG (2006) have excluded most price differences due to losses by only looking at price differences above $100/MWh. However, there remains a small possibility that some of the remaining price differences are due to losses alone and not due to constraints.

169. Earlier we observed that a precondition for an efficient hedge market is that market participants must have access to the total residues in a manner which facilitates hedging. In the NEM, NEMMCO does not retain the residues for itself – the residues that it does retain are auctioned off in the form of IRSRs. So, there is a sense in which the market already has access to all the settlement residues – so why can't market participants achieve the hedges they need?

170. The answer, as we will see, is that the residues are not currently made available in a manner which facilitates hedging. In fact, under the current NEM arrangements, the total residues arising from any given transmission constraint may be paid out to generators and/or spread over several interconnectors.

171. As we saw earlier, each constraint equation in the NEM takes the following form:

$$\alpha_1 Q_1 + \alpha_2 Q_2 + \ldots \alpha_N Q_N + \beta_{A\to B} F_{A\to B} + \beta_{B\to C} F_{B\to C} + \ldots \le RHS$$

Where:

> α_i is the coefficient in the constraint equation on generator *i*, Q_i is the output of generator *i*, $\beta_{A\to B}$ is the coefficient in the constraint equation on the interconnector from region A to region B and $F_{A\to B}$ is the flow on the interconnector from A to B and so on; RHS is a number representing the physical limit in the network.

172. As noted earlier, every constraint equation has associated with it a value known as the "constraint marginal value". The constraint marginal value is zero unless the constraint is binding. When a constraint is binding the constraint marginal value is positive. As before I will use the Greek letter lambda (λ) to denote the constraint marginal value of the constraint above.

173. It turns out that the when a single constraint is binding, the total residues or merchandising surplus accruing to NEMMCO from that constraint is just equal to the constraint marginal value times the right-hand-side of the constraint equation – or $\lambda \times RHS$. It can be shown that under certain conditions (see box 5) this amount is "firm" in the sense that market participants can obtain a perfect hedge for a fixed quantity by purchasing a fixed share of these residues.

174. In addition, it turns out that when we multiply both sides of the constraint equation above by the constraint marginal value, the inequality (the less-than-or-equal sign) becomes an equality:

$$\lambda\alpha_1 Q_1 + \lambda\alpha_2 Q_2 + \ldots \lambda\alpha_N Q_N + \lambda\beta_{A\to B} F_{A\to B} + \lambda\beta_{B\to C} F_{B\to C} + \ldots = \lambda RHS$$

175. The important thing to note here is that, under the current NEM arrangements, the total residues ($\lambda \times RHS$) are not made available to the market as a whole, rather they are divided up into the parts on the left-hand-side of the equation above and paid out to generators and interconnector. Specifically, the amount $\lambda\alpha_1 Q_1$ is paid out to the generator at node 1, $\lambda\alpha_2 Q_2$ is paid out to the generator at node 2 and so on. Similarly $\lambda\beta_{A\to B} F_{A\to B}$ is paid out to the inter-

regional settlement residues for the interconnector between region A and region B, $\lambda\beta_{B\to C}F_{B\to C}$ is paid out to the interconnector between region B and C and so on.[67]

176. For example, suppose there was a single binding constraint equation of the following form:

$$0.2Q_1 + 0.4F_{A\to B} \leq RHS$$

177. We know that the total residues ($\lambda \times RHS$) associated with this constraint are firm, but the inter-regional settlement residue fund on the A-B interconnector does not receive all of these residues. Instead, the generator at connection point 1 is paid a share of the residues equal to $\lambda 0.2Q_1$. This amount depends on the output of the generator at connection point 1. The left-over residues (equal to $IRSR_{A\to B} = \lambda 0.4F_{A\to B} = \lambda RHS - \lambda 0.2Q_1$) are paid out to the inter-regional settlement residue fund.

178. What share of the IRSRs should a trader wishing to hedge a transaction between region A and region B purchase? The answer depends on the output of the generator at connection point 1. Unless the trader can perfectly forecast the output of that generator he/she will not be able to obtain a perfect hedge.

179. Any time that a constraint equation has one or more generator connection point terms on the left-hand-side of the constraint equation, together with at least one interconnector flow term, the inter-regional settlement residues for that interconnector will not be firm.

180. In fact, exactly the same argument applies when there are two or more interconnector terms on the left-hand-side of the constraint equation. For example, suppose that there was single binding constraint equation of the form:

$$0.4F_{A\to B} - 0.2F_{B\to C} \leq RHS$$

181. Again, we see that even if the total residues ($\lambda \times RHS$) associated with this constraint are firm, the inter-regional settlement residue fund on the A-B interconnector does not receive all of these residues. Instead, the residues accruing to the A-B interconnector fund depends on the flow of the interconnector B-C (since $IRSR_{A\to B} = \lambda 0.4F_{A\to B} = \lambda RHS + \lambda 0.2F_{B\to C}$). Again, unless the trader can perfectly forecast the flow on the interconnector B-C at the time this constraint binds, it will not be able to obtain the perfect hedge it requires.

182. We can conclude that any time a constraint equation has two or more interconnector terms on the left-hand side, the inter-regional settlement residues for those interconnectors will not be firm.

183. In fact, this analysis shows us that the *only* time that the inter-regional settlement residues are firm is when the constraint equation takes the form of a single term on the left-hand-side corresponding to the flow on a single interconnector. In other words, the *only* time that an IRSR is a firm hedging instrument is when the only constraint that is binding is of the form:

$$F_{A\to B} \leq RHS$$

[67] The total inter-regional settlement residue for an interconnector A-B is the sum of these payments for all binding constraints: $IRSR_{A\to B} = \sum_n \lambda^n \beta^n_{A\to B} F_{A\to B}$.

184. In this case, and only in this case, the total residues associated with this constraint (which as we have said, are firm) are paid out in full to the corresponding inter-regional settlement fund, so that the inter-regional settlement residues are firm.

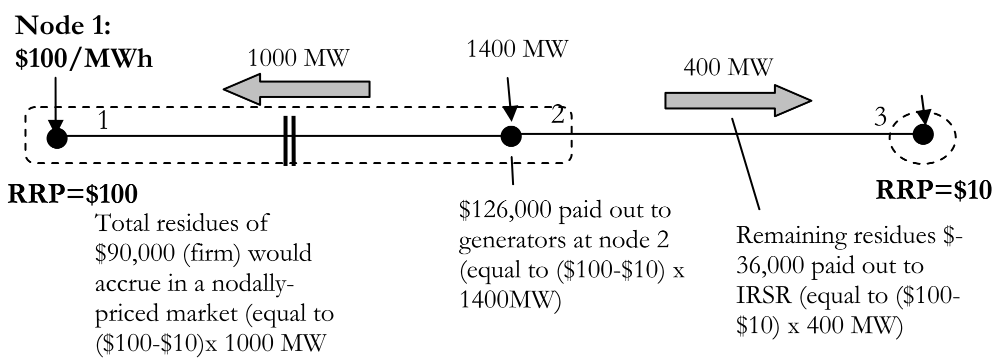
185. In summary, although the total residues associated with any given constraint are themselves "firm", under the current market arrangements these residues are divided up and paid out to those generators and interconnectors which are on the left-hand-side of the constraint equation. This disperses the residues in a way which is not useful for market participants to obtain the firm hedging instruments they need.

186. As before, these points can be made much clearer by using simple network examples. In the network set out in figure 8 (based, again, on the network of figure 1), let's suppose that the regional reference price at node 3 is $10 and the regional reference price at node 1 is $100. Since there are no constraints binding between node 2 and node 3 we know that in a market with locational marginal pricing these two nodes must have the same price ($10 in this case). In such a market, therefore, the total residues would be equal to the price difference between node 2 and node 3 ($90/MWh) times the flow on the constrained line (1000 MW), yielding total residues of $90,000/h.

187. Under the current market arrangements these residues are not directly made available to the market. Instead, under the current market arrangements, the generator at node 2 is paid out an amount equal to the difference between the regional reference price ($100) and the local marginal price ($10) times the output of the generator at node 2. If the output of the generator at node 2 is, say, 1400 MW, this generator receives residues equal to $90x1400 = $126,000/h.[68]

188. At the same time, if the output of the generator at node 2 is 1400 MW, the flow on the interconnector between node 2 and node 3 must be equal to 400 MW in the direction of node 2, so the residues on the interconnector are equal to the price difference between node 1 and node 3 ($90) times the flow on the interconnector (-400 MW) which is $-36,000/h. Since $90,000=$126,000+$-36,000, we can see that the total residues have been divided between the generator and the interconnector.

Figure 8: Lack of firmness arising from mis-pricing



189. We saw above that the correctly formulated constraint equation takes the form: $Q_2 - F_{A\to B} \le RHS$. If we multiply both sides by the constraint marginal value we find that: $\lambda Q_2 - \lambda F_{A\to B} = \lambda \times RHS$. In addition, in this simple linear network the constraint marginal value is just equal to the price difference between the regions: $\lambda = P_B - P_A$, so we can deduce that the inter-regional settlement residues between the regions are equal to the price difference

[68] These residue payments to this generator could be interpreted as an implicit "financial transmission right" to the regional reference node with a quantity equal to the level of output of this generator.

between the regions times the flow limit, less the output of the generator at node 2 times the price difference:

$$IRSR_{A\to B} = (P_B - P_A)F_{A\to B} = 1000(P_B - P_A) - (P_B - P_A)Q_2$$

190. As this example shows, the total residues (which are "firm" and equal to $1000(P_B - P_A)$ are divided up between the inter-regional settlement residues and the residues which are paid to the generator at node 2. Since the residues paid to the generator at node 2 depends on that generator's output, it follows that the share paid to the inter-regional settlement residues also depend on that generator's output and therefore cannot be "firm".

Box 5: Firmness of residues and the right-hand-side of the constraint equations

As noted in the text, we are particularly interested in determining whether or not it is possible to package the total residues in a manner which facilitates hedging. The "firmness" of the stream of residues depends on the "right hand side" of the corresponding constraint equation. The stream of residues will be firm if the "right hand side" of the constraint equation is a single, fixed number representing the physical limit on a particular element of the transmission network. In practice, however, the right hand side of a constraint equation may be variable and uncertain for two principal reasons:

(a) the transmission network may not be fully reliable – that is, there may be outages which were not forecast in advance which reduce the physical capability of a transmission network element, leading to a reduction in the constraint equation's right hand side;

(b) since, NEMMCO does not have direct control over the electricity demand of smaller electricity consumers, NEMMCO is not able to leave demand terms on the left-hand-side of the constraint equations. Instead, these are usually shifted to the right-hand-side of the constraint equations. Significant variability in demand could, in principle, lead to variation in a constraint equation's right hand side.

In regard to the first reason above, it is widely accepted that unreliability in the transmission network will affect the firmness of hedging instruments based on the total residues. In fact, in the presence of transmission network unreliability it is not possible to offer fully firm hedging instruments without some party taking on risk (which may require an external source of funds). For the purposes of this paper I will put to one side issues arising from transmission network reliability. The implications of transmission network unreliability are raised below as an issue for further research.

In regard to the second source of lack of firmness above (relating to demand terms on the right hand side of the constraint equation) we may note that this issue is closely related to the mis-pricing issue. If no connection points were mis-priced, there would be no local demand terms on the right-hand-side of any constraint equations. Therefore, correcting the mis-pricing problem through say administrative region boundary changes (which is the focus of this paper) will also solve this source of lack of firmness. There is a need for further empirical research exploring the extent to which demand variability affects the right-hand-side of constraint equations.

For the purposes of this paper I will simply assume that the right-hand-side of the relevant constraint equations are largely fixed, representing the fixed physical limit on the underlying transmission network elements. Under this assumption both the total residues and the residues associated with any particular binding constraint are "firm".

Where are the "missing residues"?

191. Up until this point I have simply argued that a share of the total residues are paid out to mis-priced generators. But there is no such explicit payment made under the current NEM arrangements? In what sense can we say that payments are made to mis-priced generators?

192. Earlier we noted that there is a close relationship between a generator's local nodal price and the regional reference price, as follows:

$$P_i = RRP - \sum_n \lambda^n \alpha_i^n$$

193. If we multiply both sides of this equation by the output of the generator at node *i*, Q_i we find that the sum of the residue payments mentioned above $\sum_n \lambda^n \alpha_i^n Q_i$ is just equal to the price difference between the regional reference price and the local nodal price.

$$\sum_n \lambda^n \alpha_i^n Q_i = (RRP - P_i) Q_i$$

194. In other words, although there is no explicit payment of the residues to mis-priced generators, there is an implicit payment of precisely the right magnitude. Paying mis-priced generators the regional reference price rather than their local nodal price is entirely equivalent to making a payment to each generator at connection point *i* equal to $\lambda \alpha_i Q_i$ for each binding constraint.

195. Viewing the problem in this way we can see why mis-priced generators (such as the generator at node 2 in figure 8 above) have an incentive to distort their bid. A mis-priced generator can be thought of as receiving two payments – a payment equal to the local marginal price times its output and a payment equal to its share of the residues times its output. Since these residues are valuable and since its share of the residues varies with its output, this generator has an incentive to distort its bid in order to capture a larger share of the residues.

196. This is an example of a more general principle: when a generator can, by taking some action, increase its entitlement to receive a positive share of valuable residues, that generator will have an enhanced incentive to take that action. For example, if the entitlement to receive a share of residues is based on the generator's output in previous periods, the generator will have an incentive to keep its output up in order to receive a greater share of the residues in future.

197. This analysis also suggests a mechanism for "firming up" the residue funds, by requiring mis-priced generators to make payments equal to $\lambda \alpha_i Q_i$ for each binding constraint. This is, in fact, precisely what happens under the current Snowy CSP/CSC trial and is what is proposed to happen more generally under the CSP/CSC mechanism or the constraint-based residues approach.[69]

What else can we learn by observing constraint equations?

198. Earlier we saw that by observation of the form of a constraint equation we can determine which generators will be mis-priced when that constraint equation is binding. From the analysis in the previous section we can see that, in addition, observation of the form of a constraint equation also can directly inform us as to the likely firmness of the inter-regional settlement residues.

199. In particular, as we have seen, whenever, a constraint equation includes both a term involving an interconnector and a term involving a generator, or when a constraint equation includes two or more interconnector terms, the residues accruing to an interconnector will not be firm.

[69] See Biggar (2006b)

200. Both of these sources of lack of firmness – that is, when a constraint equation includes terms involving a generator and an interconnector, and when a constraint equation includes terms involving two or more interconnectors – can lead to negative settlement residues. (As emphasised earlier, negative settlement residues are just one, extreme, manifestation of the problem of lack of firmness).

201. In fact, if a constraint equation includes terms involving two or more interconnectors, negative settlement residues *must* arise, as long as the flow on at least one of those interconnectors has a sign opposite to the sign of its coefficient in the constraint equation. If there are two interconnector terms in a constraint equation with opposite signs on their coefficients, negative settlement residues must arise whenever the flow on those two interconnectors has the same sign. As we will see, this explains why negative settlement residues arise in the Snowy region when the Murray-Tumut constraint binds, and flows are uniformly northwards or southwards through the Snowy region.

202. There may be other useful information that it is possible to glean from observation of the form of a constraint equation. For example, it may be possible, through inspection of the constraint equations, to say something about how generators with market power will exercise that market power – whether they will exercise market power to cause a constraint to bind, or whether they will exercise market power to prevent a constraint from binding.

203. Suppose a constraint equation includes a term involving a notional interconnector. When that constraint binds, the price difference between the two regions joined by that interconnector has the same sign as the coefficient on the interconnector term in the constraint equation.[70] If the coefficient is positive, the "to region" will have a higher price than the "from region" and vice versa if the coefficient is negative.

204. If a generator located in, say, the "from region" of a given interconnector has market power, it will exercise that market power to relieve or "unbind" a constraint which has a positive coefficient on that interconnector. Similarly, that same generator will exercise its market power in order to bind a constraint which has a negative coefficient on that interconnector.

205. Finally, inspection of the constraint equations also sheds some light on whether the price paid to a generator at a specific connection point will go up or down following a region boundary change. The price paid for a generator's output at a particular connection point is equal to the price at the regional reference node for that connection point. Therefore, the change in the price paid at a connection point following a region boundary change is equal to the price difference between that connection point's original reference node and the reference node of that connection point following the region boundary change. It is straightforward to work out the difference in price between any two nodes by inspection of the constraint equation. Therefore, it is possible to work out whether or not a generator located at a connection point will be made better off or worse off (in the sense of receiving a higher or lower price) following any given region boundary change.

206. These principles can be made clearer by considering a specific example. The following constraint equation is one of a set of constraint equations used to represent a thermal network limitation between Murray and Tumut for flows in the northerly direction. Under the current region boundary and interconnector configuration, the correctly oriented version of this constraint (known as H>>H-64_B") takes the form:

$$
\begin{aligned}
&-0.81 \times Q_{LT} + -0.792 \times Q_{UT} + 0.165 \times Q_{BLW} + 0.504 \times Q_{HUM} + \\
&0.79 \times F_{SNY \to NSW} + -0.164 \times F_{VIC \to SNY} + 0.16 \times F_{VIC \to SA(ML)} \le RHS
\end{aligned}
$$

[70] Assuming that the constraint equation is formulated in the "less than" form.

Where

> Q_{LT} and Q_{UT} are the output of the Lower Tumut and Upper Tumut generators, respectively, Q_{BLW} and Q_{HUM} are the output of the Blowering and Hume (NSW) generators respectively, $F_{SNY \to NSW}$ and $F_{VIC \to SNY}$ are the flows on the Snowy-NSW and VIC-Snowy interconnectors, respectively, and $F_{VIC \to SA(ML)}$ is the flow on the Murraylink DC interconnector between VIC and SA.

207. Following the above principles, we can see that when this constraint equation binds, which only happens when flow is in the northerly direction (and in the absence of any other congestion management mechanism such as the CSP/CSC derogation at Tumut and the Southern Generators' proposal):

(a) Upper and Lower Tumut are both mis-priced (constrained on) (since the coefficient on Q_{LT} and Q_{UT} is negative);

(b) Hume and Blowering are both mis-priced (constrained off) (since the coefficient on Q_{BLW} and Q_{HUM} are both positive);

(c) Neither VIC-Snowy nor Snowy-NSW residues will be firm (since the equation includes both generator terms and interconnector terms); and

(d) Negative settlement residues will arise when flow on VIC-Snowy and Snowy-NSW are in the same direction (that is, when flow is northwards through the Snowy region or southwards through the Snowy region) (since the VIC-Snowy and Snowy-NSW interconnector terms have the opposite sign).

(e) Snowy Hydro, located in the Snowy region, has an incentive[71] to exercise any market power it has to prevent this constraint equation from binding (since the coefficients on the VIC-Snowy and Snowy-NSW terms are such that the Snowy region has a lower price than either of the other regions when this constraint binds).

(f) Finally, when this constraint binds, the price at the NSW RRN must be higher than the Snowy RRN (since the coefficient on the Snowy-NSW interconnector is positive) and the price at the Snowy RRN is less than the price at the VIC RRN (since the coefficient on the VIC-Snowy interconnector is negative). Therefore, we can deduce that a region boundary change which, say, places Tumut generation in the NSW region and Murray generation in the VIC region will increase the price paid to *both* Murray and Tumut generation when this constraint binds.

208. Just how serious is this hedging problem? As before, one way to ask this is to ask: how many constraint equations in the NEM constraint library will give rise to non-firm inter-regional settlement residues?

209. Again, let's restrict attention to those constraints which were binding in the 2005/06 financial year. Of the 747 such constraint equations, there are 98 (13.1%) which have at least one interconnector term on the left-hand-side *and* at least one generator term on the left-hand-side. These constraints were binding for around 11% of the total constraint-minutes. Furthermore 127

---

[71] At least, in the absence of any other congestion management mechanisms, such as the CSP/CSC derogation at Tumut.

(17%) of the constraint equations have two or more interconnector terms on the left-hand-side. In all these cases the corresponding inter-regional settlement residues will be non-firm.

210. There is some overlap in these figures (since some constraint equations have both generator terms and two or more interconnector terms on the left-hand-side). As a rough guess, we might expect that around 20-25% of the time when a constraint binds, at least one inter-regional settlement residue fund will be non-firm.

Summary of the hedging problem

211. In brief, this section makes the following key points:

- A key component of an efficient electricity market is an efficient "forward", "contract" or "hedge" market. One essential precondition for efficiency in the hedge market is that market participants must have access to the total residues (or "congestion rents" or "merchandising surplus") in such a way that facilitates hedging.

- A hedging instrument is "firm" if traders can purchase a fixed share of the instrument in advance and obtain a perfect hedge for a transaction which involves buying a swap in one region and selling a swap for the same quantity in another region. By entering transactions of this kind, traders bring the prices of hedge contracts more closely in line with their underlying expected future spot prices.

- The only instrument available in the NEM for hedging such transactions – the inter-regional settlement residues – are not, in general, "firm". The IRSRs are firm in a simple radial network with constraints aligned with the region boundaries. In this case, the price difference between the regions times the flow between those regions in just equal to the price difference between the regions times the flow limit between those regions. In all other instances the IRSRs are non-firm – a price difference between two regions may arise when the flow between those regions is zero, positive, or even negative (counter-price).

- Negative settlement residues give rise to particular problems for NEMMCO, which leads them to directly intervene in the market. Negative settlement residues are, however, not a separate problem in their own right, but a symptom of the more general "lack of firmness" problem.

- It is possible to determine when the lack of firmness problem will arise by observation of the constraint equations. Specifically, the IRSRs are not firm when there is a generator connection point term along with an interconnector term on the left-hand-side of a binding constraint equation or when there are two or more interconnector terms on the left-hand-side of a binding constraint equation. In these cases, the total residues, which are firm (under certain assumptions), are spread between the IRSR fund in question and other generators or between IRSR funds.

- The residues which are paid to mis-priced generators are not paid explicitly, but implicitly. The share of the residues which is paid to mis-priced generators is precisely equal to the difference between regional reference price and the local nodal price times the output of each generator. The current regional pricing arrangements can therefore be thought of as nodal dispatch and pricing, supplemented with a payment to each mis-priced generator making up the difference on its output between the local nodal price and the regional reference price.

2. Region boundaries and the formulation of constraint equations

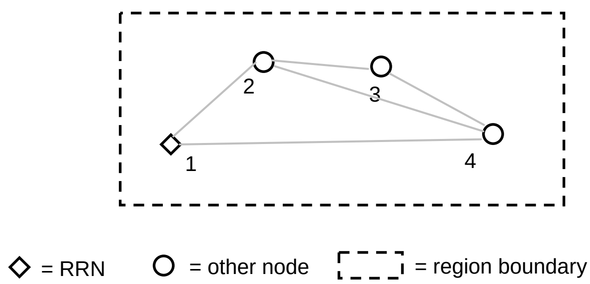
212. In the previous sections we observed that it is possible to obtain some indication of the likely pricing and hedging outcomes that will arise in a given configuration of administrative region boundaries and interconnectors by inspection of the correctly-oriented form of the constraint equations for that configuration.

213. In this section, we will see that, if we ignore losses, it is relatively straightforward to work out the appropriate correctly-oriented formulation of the constraint equations for any given configuration of regions and interconnectors. Therefore, we can, in principle, easily obtain an indication of the likely pricing and hedging implications of any given region boundary change.

214. I will focus here on the two simplest possible cases of the division of an existing region and the merger or two regions. These are the two simplest possible forms of region boundary change which we could imagine. Can we make any general statements about these simple region boundary changes? For example, is it the case that a division of an existing region only improves pricing outcomes?

The pricing and hedging impacts of the division of a region

215. Let's examine first the impact of a region division on the constraint equations. Suppose we have a network comprising a single region and four connection points labeled 1-4. Let's suppose that connection point number 1 is the regional reference node, and the other connections points are "remote intra-regional connection points", as illustrated:

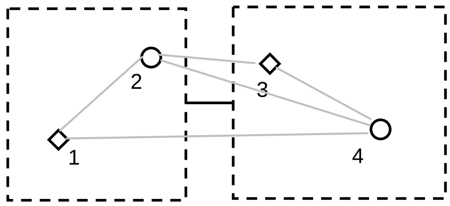


216. In this network, the correctly oriented constraint equations will take the form:

$$\alpha_2 Q_2 + \alpha_3 Q_3 + \alpha_4 Q_4 \leq RHS$$ [72]

217. Now suppose that this region is divided into two regions labeled A and B, with a new interconnector created (labeled "A-B"), with connection point 1 the regional reference node in the A region and connection point 3 the regional reference node in the B region, as illustrated:

[72] Note that the coefficient on connection point 1 (the RRN) is zero as required for the constraint equation to be correctly oriented.



218. It is straightforward to work out that the new correctly oriented constraint equations for this new network are as follows:[73]

$$\alpha_2 Q_2 + (\alpha_4 - \alpha_3) Q_4 - \alpha_3 F_{A \to B} \leq RHS$$

219. Using the principles above, we can see that a region division of this kind:[74]

(a) eliminates any mis-pricing that was present at the new regional reference node in the newly created region, no matter which constraint equation is binding (since there is no term involving connection point 3 in the above constraint equation);

(b) eliminates mis-pricing at "remote intra-regional" connection points in the newly created region in those constraint equations in which the remote intra-regional connection points happen to have the same coefficient as the new regional reference node in the former constraint equations (since in these cases the coefficient on connection point 4 is zero, since by assumption $\alpha_4 = \alpha_3$);

(c) creates new mis-pricing at remote intra-regional connection points in the newly created region in those constraint equations in which there was no mis-pricing of those connection points originally, but there was mis-pricing of the new regional reference node in the former constraint equations (since in these cases $\alpha_4 = 0$ and $\alpha_3 \neq 0$ so $\alpha_4 - \alpha_3 \neq 0$);

(d) has no effect on any mis-pricing that was present at remote intra-regional connection points in the original region (since the coefficient on connection point 2 remains unchanged); and

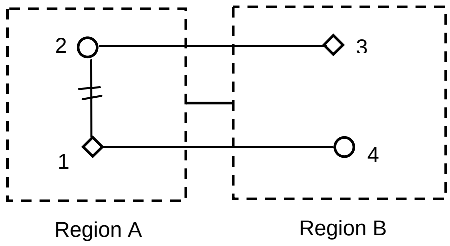
(e) will not yield firm residues on the new A-B interconnector as long as there is any mis-pricing of remote intra-regional connection points in either of the two regions (that is as long as either the coefficient on connection point 2 or connection point 4 is non-zero).

(f) results in a lower price for generation in the new region B if and only if the coefficient on connection point 3 is positive (and vice versa).

[73] This equation is obtained by subtracting a multiple of the energy-balance equation for the second region, in such a way as to eliminate the coefficient on connection point 3.
[74] Many participants (including both Snowy Hydro and Macquarie Generation) have made reference to the recommendation by CRA that "no region shall have a maximum demand of less than 200 MW". However, in my view, this recommendation is not justified by any theory of which I am aware. This is also the view of Eraring Energy who write: "this recommendation by CRA is in fact quite baseless with their report providing neither elaboration nor justification". (Eraring Energy, submission to the AEMC 22 March 2006, page 3).

220. It is important to recognise that there is no guarantee that a region division will improve the mis-pricing problem. Although a region division does eliminate mis-pricing at the regional reference node in the new region and at other nodes in the new region with the same coefficient in the constraint equation, the region division *creates* new mis-pricing at nodes which previously were not mis-priced.

221. To provide an example of how this might work, consider the following simple network:



222. The correctly oriented constraint equation for the constraint between node 1 and node 2 prior to the region division is $-z_2 - z_3 \leq RHS$. After the region division, the correctly oriented constraint equation for this constraint is $-z_2 + z_4 + F_{A\to B} \leq RHS$. It is clear that mis-pricing at node 3 has been replaced by mis-pricing at node 4. The clear conclusion is that we cannot be sure that an arbitrary division of an existing region will improve the mis-pricing problem.

The pricing and hedging impacts of a merger of two regions

223. Now let's examine the impact of a merger of two regions on the constraint equations. Let's suppose that we have two regions joined by a single interconnector, as illustrated above. Now let's consider merging these two regions, retaining connection point 1 as the regional reference node in the new merged region (i.e., just the inverse of the region division above). Let's suppose the constraint equations in the original two-region network have the following generic form:

$$\alpha_2 Q_2 + \alpha_4 Q_4 + \beta_{A\to B} F_{A\to B} \leq RHS$$

224. It is straightforward to work out that the corresponding constraint equation for the merged region is:

$$\alpha_2 Q_2 - \beta_{A\to B} Q_3 + (\alpha_4 - \beta_{A\to B}) Q_4 \leq RHS$$

225. (It is easy to check by comparing this equation with the equation above that the process of dividing and then merging two regions restores the original constraint equations).

226. Again, using the principles above, we can see that a merger of two regions

(a) creates new mis-pricing at the former regional reference node of the eliminated region under any constraint equations which also previously included a term involving the interconnector (since the coefficient on connection point 3 is equal to the coefficient on the interconnector in the former constraint equation).

(b) creates new mis-pricing at the former remote intra-regional connection points in the eliminated region in those constraint equations in which there was previously no term involving those connection points but there was a term involving the interconnector (since if the coefficient on connection point 4 in

this case is equal to the coefficient on the interconnector in the former constraint equation).

(c) eliminates mis-pricing at the former remote intra-regional connection points in the eliminated region in those constraint equations in which the coefficient on that connection point matched the coefficient on the interconnector (since in this case $\alpha_4 = \beta_{A \to B}$ by assumption)

(d) has no effect on any mis-pricing that was present at remote intra-regional connection points in the non-eliminated region (since the coefficient on connection point 2 is left unchanged).

(e) results in a higher price for region B generation if and only if the coefficient on the interconnector term is positive (and vice versa).

227. In addition, a merger of two regions has implications for hedging – specifically it eliminates the need for market participants trading between the two regions to make use of a (possibly non-firm) inter-regional settlement residue to hedge their risks, but introduces new dispatch risks for mis-priced generators.

228. Of course, any possible region boundary change that might be considered is likely to be somewhat more complicated than a simple region division or the merger of two regions, as illustrated above. Nevertheless, if we ignore losses, it is possible, via an extension of this sort of analysis to determine the appropriate form of the constraint equations following any given change in the configuration of regions or interconnectors. As we have seen, inspection of the appropriately formulated constraint equations, allows us to go some distance towards analyzing the implications of any given region boundary change.

229. However, this analysis has its limitations. It is not possible, for example, through this analysis to assess the likely impact on average prices, flows, dispatch or overall market efficiency following any given region boundary change. The actual pricing, dispatch and hedging outcomes that will arise in the market depend on precisely which constraints will bind and the frequency and duration of those constraints. There are a very large number of constraint equations, each of which has its own pricing, dispatch and hedging implications. The change in the incentives brought about by the new region boundary, will lead to new bidding and dispatch outcomes, and therefore to different flows on the transmission network. Following a change to the definition of region boundaries or interconnectors, new constraints may emerge as important, while other constraints which were significant under the old regions may become insignificant.

230. Therefore, while it is possible to make statements of the form "*if* constraint X binds, the pricing, dispatch and hedging implications will be as follows…", it is not possible to state with certainty (at least not without significant additional analytical or modelling analysis) that constraint X will bind with certainty.

Should we increase the number of pricing regions?

231. Does this analysis shed any light on the appropriate configuration of the administrative pricing regions in the NEM? Alternatively, does this analysis shed any light on which changes to the existing region boundaries should be approved? We might make the following points:

232. Although – in a radial network –it is known that pricing and hedging outcomes will improve if the region boundaries are aligned with the physical constraints in the network, in a meshed network we cannot say *a priori* that any given region boundary change (including a region division or the merger or regions) will make the existing mis-pricing or hedging problems better or worse – it depends on the precise configuration of the network.

233. It is always possible to solve the mis-pricing problem by moving to small enough regions. However, doing so creates new prices and new price-risks which need to be hedged. Creating new regions also creates new hedging instruments to hedge those risks but, as we have seen, there is no guarantee that these new hedging instruments will be firm. As long as there remain some mis-priced generators, or as long as the region division gives rise to loop-flow between the regions, the inter-regional settlement residues will not be a firm instrument for hedging the risks of inter-regional transactions.

234. Hedging is important. It is possible that efficiency in the hedge market (which impacts on location and expansion decisions) is even more important than efficiency in the spot market. New price risks should not be created unless we simultaneously create a mechanism which allows market participants to hedge those risks.

235. At the same time, the ineffectiveness of the existing hedging instruments may be creating pressure for the merger of regions – limiting the need for hedging. This creates its own problems as we have seen above.

236. The mis-pricing and hedging problems are closely related and should be solved simultaneously. It does not make sense to give further consideration to region boundary changes in the NEM without simultaneously developing separate mechanisms for addressing the hedging problems that might result.

Box 6: Is the argument against more regions primarily a matter of transactions costs? Or is it due to the ineffectiveness of inter-regional settlement residues as a hedging device?

A few market participants have argued that an increase in the number of regions would increase the risk or transactions costs associated with a generator or retailer trading inter-regionally and therefore would reduce the volume of inter-regional trade. This view, for example, seems to be reflected in the AEMC's draft determination on the abolition of the Snowy region:

> "[A]n increase in the number of regions could increase transactions costs for managing inter-regional price risk. This could reduce inter-regional contracting and promote alternative strategies such as participants trading *intra*-regionally or vertically integrating their operations."[75]
>
> "an increase in the number of regions … can also raise transactions costs for participants wishing to enter into inter-regional financial contracts for electricity. This can lead to a decline in interregional contracting, more geographically-specific development of new generation plant and increased impetus for vertical integration."[76]

The same view is expressed by Southern Hydro and Macquarie Generation who commissioned a study by Firecone Ventures. Firecone observe that:

> "an increase in locational pricing will increase the transaction costs in the contract market and reduce its efficiency".[77]
>
> "The introduction of a large number of regions would greatly increase the possible number of significant price separations that participants would be exposed to, and hence the number of products required to hedge inter-regional price risk. This could be expected to result in a loss of liquidity in the trade of such products."[78]

That study concluded:

> "an increase in locational pricing in the spot market, either as a permanent change or as a transitional measure, is likely to result in a greater level of inter-regional price risk … and greater difficulty in pricing the risks".[79]

This line of argument seems to run as follows: a substantial increase in the number of regions will increase the number of hedging instruments available and increase the number of hedging instruments required in order to hedge any given inter-regional transaction. This, in turn, increases the transactions costs associated with a given inter-regional transaction, reducing the willingness of market participants to enter into such transactions.

But this argument seems to overstate the importance of transactions costs relative to the value or effectiveness of the hedging instruments themselves. Presumably the effectiveness of the hedging instruments is at least equally important, or even more important than the transactions costs associated with the mere number of such instruments.

If a market participant can, by combining two or more hedging instruments together, obtain the effective hedge that it desires, it is far more likely to enter into an inter-regional transaction than if it only had one, highly imperfect instrument available to hedge that transaction. Conversely, a change in region boundaries which eliminated one hedging instrument, leaving only imperfect or ineffective hedging instruments, might have a significant chilling effect on inter-regional trade, despite the apparent reduction in transactions costs.

The best way to improve inter-regional trading is to improve the quality of the instruments for hedging

---

75 AEMC (2007), page 12.
76 AEMC (2007), page 56.
77 Firecone (2006), page 1.
78 Firecone (2006), page 18.
79 Firecone (2006), page 25.

inter-regional trading risks. The absolute number of such instruments is a secondary consideration.

In my view, although the focus on transactions costs is misplaced, the underlying concern that increasing the number of regions might hinder inter-regional trade is legitimate.[80] The analysis in this paper shows that there is a sense in which this conclusion is likely to be true. The division of an existing region creates new pricing risks for some generators between their location and the former regional reference node. As long as the newly-created inter-regional settlement residues are an imperfect instrument for hedging these risks, the division of a region may reduce the efficiency of the hedge market.

Unfortunately, however, the precise impact of a region boundary change on the hedge market is unclear. A region division may reduce or eliminate mis-pricing, thereby "firming up" the inter-regional settlement residues at the same time as it introduces new pricing risk. A region division may also introduce new loop flow between regions, again leading to the lack of firmness on inter-regional settlement residues.

Because of the lack of clarity of the impact of region boundary changes on the hedge market, policies to fix the mis-pricing problem (such as region boundary changes) should be progressed simultaneously with policies to solve the hedging problem.

[80] The AEMC notes in its draft determination: "As noted in a recent paper prepared by Firecone, a large increase in the number of regions could, other things being equal, have a number of detrimental effects on the market that may more than offset the positive impact of more regions on reducing mis-pricing in the NEM". AEMC (2007), page 12.

3. Conclusions

237. This paper has attempted to set out the current level of understanding regarding the conceptual analysis of changes to the definition of administrative pricing region boundaries and interconnectors in the NEM.

238. In brief, this paper has made the following key points:

- All of the issues arising from changes to administrative region boundaries and "congestion management" relate to two underlying problems – the "mis-pricing" problem and the "hedging problem".
- The mis-pricing problem arises when the administrative pricing regions do not align with the pricing regions created by transmission constraints. In this case some generators are constrained on or off and may have an incentive to distort their bids, leading to dispatch inefficiency, poor location decisions and a lack of "firmness" on the inter-regional settlement residues.
- In addition to short-term dispatch efficiency, it is equally important to ensure efficiency in the forward contract or "hedge market". Efficiency in that market requires access to firm instruments for hedging the risk of price differences arising between different locations. In the presence of mis-pricing or loops between regions, the inter-regional settlement residues are not a "firm" instrument for hedging the risks of inter-regional trading.
- Observation of the form of the "correctly formulated" constraint equations provides some guidance as to the likely extent of the mis-pricing and hedging problems. A given constraint equation will lead to mis-pricing at all those connection points which have a non-zero coefficient in that constraint equation. Similarly, when a constraint equation includes both connection points and an interconnector term, or terms in two or more interconnectors, the inter-regional settlements will not be a "firm" hedging instrument.
- It is straightforward to transform constraint equations correctly formulated for one set of region boundaries to constraint equations correctly formulated for another set of administrative region boundaries. Analysis of the transformed constraint equations sheds some light on the likely pricing and hedging outcomes following any given potential region boundary change.

Region boundary changes should not be considered without policies to address the hedging problem and vice versa

239. As we have seen, the hedging and the mis-pricing problems are related. Policies to resolve these issues should be progressed simultaneously. In particular, changes to region boundaries should be considered simultaneously with separate mechanisms for addressing the hedging problems that might result:

- Any move to increase the number of regions will introduce new pricing risks and new inter-regional settlement residues. But those new inter-regional settlement residues will not necessarily be "firm". An increase in the number of regions therefore risks reducing the efficiency of the hedge market and possibly giving rise to negative settlement residues.
- Conversely, a reduction in the number of regions could eliminate the need for some generators to make use of inter-regional settlement residues and could eliminate negative settlement residues, but would also be likely to increase the extent of mis-pricing.

240. In general, it is not possible to address two public policy problems with a single policy tool. Two public policy problems require two policy solutions. Changes to region boundaries should be considered in conjunction with some other mechanism for resolving the hedging problem – such as the constraint-based residues approach mentioned earlier.

241. The AEMC is currently in the process of carrying out a review of congestion management issues in the NEM (the "Congestion Management Review"). That review may result in the development of new mechanisms to correct the mis-pricing and hedging problems in the NEM.

Further work

242. Although progress has been made in the analysis of congestion management and region boundary issues in the NEM, there are many theoretical questions which remain to be answered, such as the following:

- What are the theoretical conditions required for efficiency in the hedge market? How might we measure the economic impact of inefficiency in the hedge market?
- How "non-firm" is the right-hand-side of the constraint equations? Will the right-hand-side of the constraint equations be acceptably firm in a market where only generators receive the nodal price (and consumers pay the regional reference price)? Do we need nodal pricing for load to make the right-hand-side "firm"?
- What is the contribution of transmission network outages (which have been put to one side for the purposes of this paper) to the overall level of "non-firmness"? Should we be pursuing mechanisms which explicitly provide insurance against certain transmission outages? To what extent do mechanisms such as "Financial Transmission Rights" provide insurance against transmission outages? To what extent is that a key feature of such mechanisms?
- What is the contribution of losses to the overall level of non-firmness? (Losses also give rise to price differences between locations, and the IRSRs are not a "firm" instrument for hedging these losses).
- Under what circumstances is the firmness of hedging instruments affected by the exercise of market power? If a generator owns production capacity at opposite ends of a potentially constrained link, can that affect the firmness of hedging instruments?
- How does the exercise of market power impact on the liquidity of the hedge market? Is it possible that a correction of the mis-pricing problem might increase the ability of a firm to exercise market power to the extent that it reduced the liquidity of the hedge market?

236. These questions are left for future research.

References

AEMC, (2007), "Abolition of Snowy Region: Draft Rule Determination", 19 January 2007, available at: http://www.aemc.gov.au/pdfs/reviews/Abolition%20of%20Snowy%20Region%20and%20Alternative%20Snowy%20Region%20Boundary/aemcdocs/011Draft%20Rule%20Determination%20-%20Abolition%20of%20Snowy%20Region.PDF

AER, (2006a), "Investigation Report: The Events of 31 October 2005", October 2006, available at: http://www.aer.gov.au/content/index.phtml/itemId/705274/fromItemId/656186

AER, (2006b), "Compliance Bulletin: Complying with dispatch instructions", No.1, December 2006, available at: http://www.aer.gov.au/content/index.phtml/itemId/692887/fromItemId/658743

Alberta, (2003), "Transmission Development: The Right Path for Alberta: A Policy Paper", Alberta Energy, Energy Business Unit, November 2003, available at: http://www.energy.gov.ab.ca/docs/electricity/pdfs/transmissionPolicy.pdf

Biggar, Darryl (2006a), "Congestion Management Issues: A Response to the AEMC", 12 April 2006, available at: http://www.aemc.gov.au/pdfs/reviews/Congestion%20Management%20Review/issuespaper/017AER%20Submission.PDF

Biggar, Darryl (2006b),"Solving the Pricing and Hedging Problems in the NEM using Constraint-Based Residues", 20 October 2006

Biggar, Darryl (2006c), "Congestion Management Issues: How Significant is the Mis-Pricing Impact of Intra-Regional Congestion in the NEM", 25 October 2006, available at: http://www.aemc.gov.au/pdfs/reviews/Congestion%20Management%20Review/Supplementary%20Material/An%20Assessment%20Of%20the%20Significance%20of%20Intra-Regional%20Congestion%20In%20The%20NEM%20(Independent%20Consultancy%20Report%20By%20Darryl%20Biggar).pdf.

CRA (2002), "Constraint Representation in the NEM", 8 January 2002, available at: http://www.mce.gov.au/assets/documents/mceinternet/Constraintrepresentation20041130110033.pdf

CRA (2003), "Constraint Orientation: Principles and Pricing Implications", 28 March 2003, available at: http://www.mce.gov.au/assets/documents/mceinternet/ConstraintOrientation20041130105608.pdf

CRA (2004a), "NEM – Transmission Region Boundary Structure", September 2004, available at: http://www.mce.gov.au/assets/documents/mceinternet/CRARegionalStructureReviewPaper20041123152720.pdf

CRA (2004b), "A Transitional Non-LMP Market for California: Issues and Recommendations: Final Report", 16 September 2004, available at: http://www.caiso.com/docs/2004/09/17/2004091712202513134.pdf

CRA (2005), "Constraint Support Pricing: Implementation of Snowy Proposal", March 2005

ERIG (2006), "Energy Reform Implementation Group: Discussion Papers", November 2006, available at:

http://www.erig.gov.au/assets/documents/erig/ERIG%5FDiscussion%5FPapers20061117171022%2Epdf

Firecone Ventures (2006), "The impact of locational pricing on the contract market", November 2006, available at: http://www.aemc.gov.au/pdfs/reviews/Congestion%20Management%20Review/issuespaper/029Macquarie%20Generation%20and%20Snowy%20Hydro%20Supplementary%20Submission%20On%20impact%20of%20locational%20pricing%20on%20the%20contract%20market%20-%2022%20December%202006.pdf

Group submission (2004), "Response to CRA Report on NEM – Transmission Regional Boundary Structure", 12 November 2004 (the Group is AGL, Delta, Loy Yang, Macquarie Generation, Stanwell, Yallourn, Powerlink and TransGrid), available at:

Harvey, Scott and William Hogan (2000), "Nodal and Zonal Congestion Management and the Exercise of Market Power", 10 January 2000

IES (2004), "Regional Boundaries and Nodal Pricing: An Analysis of the Potential Impact of Nodal Pricing and Market Efficiency: Final Report", 2 December 2004, available at: http://www.mce.gov.au/assets/documents/mceinternet/ACCCAttachmentAIESreport20041222115205.pdf

IES (2006), "Modelling of Transmission Pricing and Congestion Management Regimes: Report on the estimation of dynamic inefficiencies", 22 December 2006

Joskow, Paul and Jean Tirole (2000), "Transmission Rights and Market Power on Electricity Power Networks", *Rand Journal of Economics*, 31(3), Autumn 2000, 450-487

Joskow, Paul and Jean Tirole (2004), "Reliability and Competitive Electricity Markets", 26 October 2004.