Andrew Truswell
Australian Energy Market Commission
PO Box A2449
Sydney South
NSW 1235

Reference: EPR0019

Dear Andrew,

Re: Transmission Frameworks Review Second Interim Report

Thank you for the opportunity to comment on the AEMC's proposed set of reforms to the NEM, known as the Optional Firm Access ("OFA") model. I have prepared a document which seeks to analyse and critique the OFA model. I would like to submit this document to the AEMC for your consideration. This submission reflects my own views and does not necessarily reflect the views of the Australian Competition and Consumer Commission, or the Australian Energy Regulator. I hope this analysis and critique proves useful.

While I commend the AEMC for its ambition in proposing a model designed to address a range of issues in the NEM, I am concerned that in some respects the model doesn't go far enough; in other respects, the model will fail to achieve certain objectives. At this stage it seems at least plausible that it will be possible to address the concerns in the NEM with a cleaner, simpler, and more sustainable long-term solution. As a result it seems too soon to be making a decision in favour of a framework such as the OFA model.

Please do not hesitate to contact me if you need further clarification.

Regards,

Dr Darryl Biggar

ANALYSIS OF THE AEMC OPTIONAL FIRM ACCESS MODEL

Darryl Biggar

4 October 2012

In the Second Interim Report of the Transmission Frameworks Review (TFR) the Australian Energy Markets Commission (AEMC) has proposed a set of reforms to the National Electricity Market, known as the "Optional Firm Access" (OFA) model. The OFA model is an attempt to set out a coherent package of interacting reforms in the NEM. These reforms would collectively represent a major change to the design and operation of the National Electricity Market (NEM). Given the scope of the proposed changes it is important that these claims be carefully analysed and tested.

The Second Interim Report suggests that the OFA model is designed to solve the following problems in the NEM:[1]

- The lack of efficiency in short-run dispatch outcomes due to a lack of effective locational price signals, resulting in disorderly bidding and inefficiency in dispatch;
- A potential reduction in the depth or liquidity in the hedge market that arises when generators and loads are situated at differently-priced locations (this issue arises when we seek to improve the locational price signals in the NEM to address the first problem);
- A lack of responsiveness to market conditions in the operational decisions of TNSPs; and
- A lack of efficiency and market-responsiveness in transmission investment decisions resulting in a lack of co-optimisation of generation and transmission investment decisions.

This paper analyses the OFA model in the light of these objectives.

This paper does not seek to identify all of the possible issues with the model. The model is broad in scope and will require substantial further development before implementation. I have not focused here on design or development issues which could be the subject of later consideration or refinement. For example, I have not addressed questions such as the design of the mechanism for allocating or pricing firm access rights[2], the definition of the "Firm Access Standard", the question of five-minute versus thirty-minute dispatch, or issues surrounding the transition to the new arrangements. Instead I have tried to focus on issues which go to the core elements of the proposal and whether those elements would work in principle, even if it was perfectly implemented.

[1] See SIR page 45. I will use the abbreviation SIR and TR for the Second Interim Report and the Technical Report, respectively. SIR: AEMC, "Transmission Frameworks Review: Second Interim Report", 15 August 2012. TR: AEMC, "Transmission Frameworks Review: Technical Report: Optional Firm Access", 15 August 2012.

[2] The AEMC proposes the use of Long-Run Incremental Cost.

Furthermore, there are other important issues which are not discussed in the Technical Report but which will require detailed further consideration. One fundamentally important issue is the operation of the OFA model under situations of market power.[3] At this stage it seems reasonable to first assess the OFA model in an ideal situation of a perfectly competitive market. If a market design does not work for a competitive market, we can reject it before we take the time to consider how it might operate in a market with market power. Later analysis however will have to fill this gap.

An important issue which is worth raising at the outset is how best to characterise and label the OFA model. In my view the OFA model is usefully (if slightly imperfectly) characterised as a form of nodal pricing for a subset of generators coupled with two new forms of transmission rights: An optional fixed-volume transmission right, and a non-optional non-firm transmission right.

In the Technical Report the AEMC sets out a list of potential concerns with nodal pricing.[4] These include, amongst other things, market power concerns and concerns about the availability of hedge contracts. But, as I have just noted, the OFA model is usefully characterised as introducing a form of nodal pricing. The same concerns that the AEMC raises with respect to nodal pricing therefore apply, to an extent, to the OFA model. In particular, market power concerns may continue to arise under the OFA model[5] and this will have to be analysed further. In addition, as we will see below, a primary concern discussed below is whether or not the OFA model will continue to ensure the availability of hedge contracts. By correctly characterising the OFA model as a form of nodal pricing plus transmission rights we are in a better position to identify its likely strengths and weaknesses.

The AEMC describes the OFA model as providing a form of "firm access". In my view the term "firm access" is somewhat misleading and therefore likely to be unhelpful. No generator in the NEM has a physical right to exclude other generators from using any part of the shared transmission network. Nor would it be desirable to allow such a right. No incumbent generator should ever be able to exclude another generator from connecting to the shared network provided the new generator is prepared to accept the local nodal price. The AEMC does not propose to establish some form of physical right to the network. Instead it merely proposes to establish a new form of financial right, as noted above. The use of the term "firm access" has the potential to obscure rather than bring out the key features of the model.

The AEMC emphasises that procuring firm access would be optional. The Technical Report emphasises that the "optionality" of the model is one of its key principles[6] - that is, generators are permitted but not required to obtain firm access. The optional nature of the access has been used as an argument in support of the claim that the model would be an improvement on the status quo: "How can providing generators the *option* to obtain firm access – an option which they do not have at present - leave them worse off than under the status quo?". However, in this sense the use of the word optional is misleading. Although generators have the option of not acquiring firm access, if another generator which shares a positive coefficient in a binding constraint equation purchases firm access, the first generator is left strictly worse off. In this

[3] The term 'market power' does not occur in the Technical Report except once, in connection with nodal pricing.
[4] TR, page 95.
[5] It is true that some generators (those with a negative coefficient in a binding constraint equation) are partially excluded from the OFA model. These generators may (or may not) coincide with the set of generators which have market power.
[6] TR, page 15.

sense, acquiring firm access is not optional. The AEMC recognises this.[7] It seems likely that affected generators will be forced into a "prisoners' dilemma" game in which the unique equilibrium is where all generators seek to procure firm access. In this sense, labelling the acquisition of firm access as optional is misleading. There is a sense in which we could reasonably refer to the model as "forced firm access".

One criticism that has been made of the OFA model is that it seems to implicitly assume that each regional transmission network takes the form of a "hub and spoke". The authors are clearly aware of the fact that all real networks are highly meshed (there are examples of meshed networks in chapter 12 of the Technical Report). Nevertheless, the model seems to assume that each generator can negotiate its firm access level with the TNSP independently of every other generator. This is not true in a meshed network. The amount that any one generator is able to produce at any given time will in general depend on the output of other generators and loads in the market. As a consequence, the amount of firm access which can be allocated (without having to rely on scaling) to any one generator potentially depends on the amount allocated to other generators and loads in the network. The proposal seems to envisage that each generator will negotiate access on a first-come-first-served basis.[8] It is not clear how this approach will ensure that the firm access rights will be allocated to those who value them most highly – or that the particular allocation chosen will maximise overall economic welfare. This is a fundamental issue which will need to be resolved.

But, as we will see, in my view there are more fundamental issues with the proposed approach which are worth highlighting. The next four sections discuss each of these issues in turn.

Issue #1: Achieving short-run efficiency in dispatch

One of the most fundamental objectives of a wholesale electricity market is achieving efficient short-run dispatch – that is, the efficient short-run use of the available stock of physical (generation, transmission, and load) assets.

It is well known that the NEM's regional pricing design does not necessarily result in efficient short-run dispatch outcomes whenever intra-regional constraints are binding. The reason is that, under regional pricing, remote[9] generators and loads in a region are paid the regional reference price rather than the correct local or nodal price. This "mis-pricing" results in disorderly bidding, inefficient dispatch outcomes, dispatch uncertainty for generators, incentives to manipulate bid parameters, and can result in counter-price flows on interconnectors.[10] In certain circumstances system security and reliability can be threatened. The only known way to correct the problems arising from mis-pricing is to ensure that at least the affected generators and loads face the correct local marginal price. Some wholesale electricity markets overseas which adopted regional pricing in the past have subsequently moved to nodal pricing.

The AEMC claims that the OFA model will lead to "more efficient dispatch of generators – by reducing the current incentives on generators to engage in disorderly bidding"[11]. In addition, "the OFA model would reduce the incentives for disorderly bidding by decoupling access to the

---

[7] TR page 83.
[8] See, for example, TR, page 47.
[9] As in the broader literature on congestion management in the NEM, "remote" in this paper is used in this paper to mean "at a location other than the regional reference node".
[10] See for example, Biggar (2006), "How significant is the mis-pricing impact of intra-regional congestion in the NEM?", 25 October 2006 (on the AEMC website).
[11] SIR page 45.

regional reference prices from an individual generator's dispatch level, and should therefore enhance productive efficiency".[12]

It seems correct that the OFA model will reduce the incentives for disorderly bidding by some scheduled generators (at least some of the time – see the discussion below). But it is important not to lose sight of the fact that the elimination of disorderly bidding does not ensure that economic efficiency is achieved. Only *scheduled* generators and loads (which submit bid and offer curves) can engage in disorderly bidding, whereas efficient short-run dispatch requires correct price signals to *all* generators and loads. The OFA model does not correct the price signals for non-scheduled generators, or scheduled or unscheduled loads. Furthermore, as discussed below, the OFA model does not eliminate the incentives for disorderly bidding for any generator which has a negative coefficient in a potentially binding constraint equation.

Inefficient dispatch of scheduled generators

Under the OFA model, non-scheduled generators and scheduled or unscheduled loads are treated exactly the same as in the status quo – that is, they are paid (or pay) the regional reference price for their production (or consumption) irrespective of the presence of transmission constraints.

The situation for scheduled generators is, however, a little more complicated. As I understand the OFA model (as summarised in the appendix), the "entitlement" received by a scheduled generator depends on whether or not it has a positive coefficient in a binding constraint equation. If a scheduled generator has a positive coefficient, it receives an entitlement for that constraint equation which depends on the volume of firm access it has procured (or, if there are funds left over, a secondary entitlement which varies according to its "availability"). Otherwise, if the generator has a negative coefficient in a binding constraint equation, that generator receives an entitlement for that constraint equation which depends on its own production (as in the status quo).

As a consequence, whenever a scheduled generator has a positive coefficient in every binding constraint equation it faces the correct nodal price at the margin. It no longer has an incentive to distort its offer curves in the way known as disorderly bidding. But, any time in which a scheduled generator has a negative coefficient in a binding constraint equation, that generator will face a price which is *too low* relative to the correct local nodal price. Depending on the variable cost of the generator this can result in inefficient dispatch outcomes and disorderly bidding. In particular, such a generator may have an incentive to pretend to be unavailable. In extreme situations this can threaten reliability.

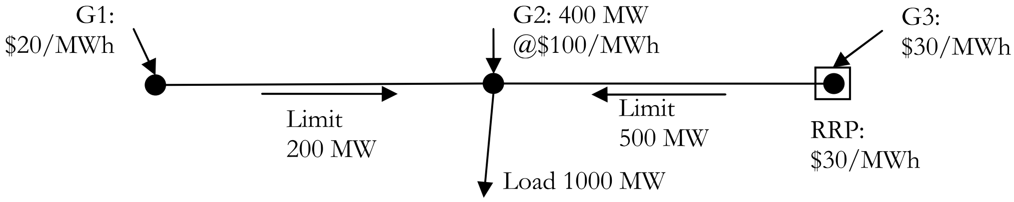
It is easiest to see this effect when there is a single binding constraint. In this case, any generator with a negative coefficient in that binding constraint equation will receive just the regional reference price for its output, rather than the higher local nodal price. This may put the generator in a position where it does not wish to be dispatched at the price that it is paid even though the output of this generator is valued by the market more than its variable cost. This is the well-known case of "constrained on" generation. In such circumstances, the generator will have an incentive to distort its bid (by raising the offer price to the market price cap) and/or to pretend to be unavailable. This results in inefficient dispatch and may threaten reliability.

It is sometimes suggested that situations where the local nodal price is above the regional reference price and the corresponding generators are "constrained on" are special or exceptional cases, which should perhaps be dealt with separately (that is, not through a market mechanism).

[12] SIR page 52.

But it is important to recognise that the mis-pricing arising in the OFA model that is discussed above can also arise when the local nodal price is *below* the regional reference price.

An example of this outcome is provided by the AEMC in the Technical Report, Figure 12.6, page 106, and is reproduced here below. In this example, there are three nodes and two links. There is load at node 2, which results in power flows from node 1 and node 3.



The relevant correctly-oriented constraint equations for this simple network are as follows:

$G_1 \leq 200$ (for the left hand link) and $L_2 - G_1 - G_2 \leq 500$ (for the right hand link).

In this example, both of these transmission constraints are binding. From the constraint equations above and equation (A 3) in the appendix we can determine that the local nodal price at nodes 1 and 2 are $20 and $100 respectively. But, in this example both G1 and G2 have a negative coefficient in the constraint corresponding to the limit between node 3 and node 2. Therefore both G1 and G2 remain mis-priced in the OFA model. The effective price faced at node 1 and node 2 are $p_1 = P_R - \lambda_{1\to2}$=$-50/MWh and $p_2 = P_R$=$30/MWh.

Let's look first at node 2. The OFA model pays G2 the regional reference price, $30/MWh. At this price, G2 does not want to produce and so engages in disorderly bidding. But if G2 does not produce there is not enough transmission capacity to service the load at node 2. Some load would have to be shed, compromising reliability.

Now let's look at node 1. Node 1 is also mis-priced. Generators at node 1 are paid an effective price of $p_1 = P_R - \lambda_{1\to2} = \$30 - \$80 = \$ - 50$. Since this price is below the variable cost of the generators at node 1, these generators have an incentive to distort their bid and pretend to be unavailable. Therefore, even if somehow the generators at node 2 could be induced to increase their production, there would still not be enough generation to meet the load. Some load would need to be shed. Even worse, since the effective price paid by G1 is negative, if G1 could install a large device for consuming electricity (perhaps a device for heating water or pumping water uphill), the generator would have an incentive to consume large amounts of electricity – further worsening the supply/demand balance in the network.

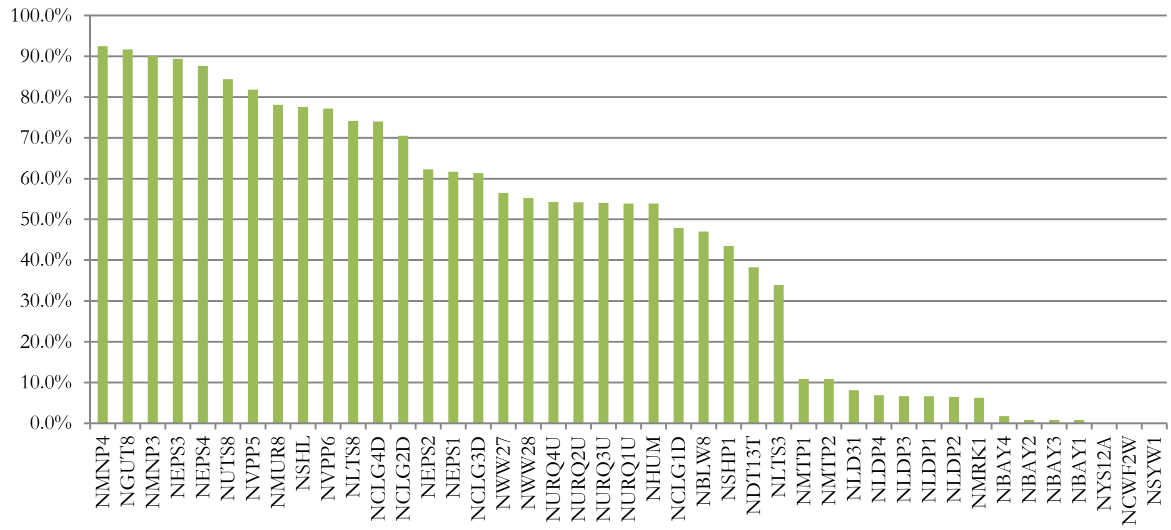
In this example, the mis-pricing results in pure social waste. Arguably this outcome is worse than the status quo. Certainly the mis-pricing in this example is not corrected by the OFA model.

Perhaps, it might be argued, situations where there is a negative coefficient in a binding constraint equation are rare or exceptional and can be ignored. Perhaps only a few generators have a negative coefficient in a constraint equation or, even if they do have a negative coefficient, perhaps those constraints do not bind very frequently? Perhaps most generators would be correctly priced under the OFA model.

We can check this by looking at the market data. The graph below shows, on the horizontal axis, the set of all connection points in NSW which were mis-priced at least once in the last five years (that is, which had a non-zero coefficient in a binding constraint equation). The vertical axis shows the proportion of time in which that connection point had a negative coefficient in a binding constraint equation (and which therefore would be mispriced under the OFA model). For example the connection point NGUT8 (Guthega) appeared in a binding constraint equation (and therefore was mis-priced) in 23,931 dispatch intervals during this period, and on 21,900 (more than 90 per cent) of these occasions NGUT8 had a negative coefficient in that constraint equation. NGUT8 would have been mis-priced during 21,900 dispatch intervals in the last five years under the OFA model.

As can be seen, the bulk of these connection points had a negative coefficient in a binding constraint equation around half of the time that the constraint equation was binding. Virtually all of these connection points have a negative coefficient at least once. Only two connection points (NCW2FW and NSYW1) would have been consistently correctly priced under the OFA model (and constraints affecting these generators were very rarely binding).

Figure 1: Percentage of negative-coefficient events in binding constraints by connection point in NSW



Intuitively, it is easy to see why negative coefficients in constraint equations are likely to be common: A network element will almost always have a maximum flow limit *in each direction*. Any generator which has a positive coefficient in the constraint equation for flow in one direction on the network element, will automatically have a negative coefficient for the constraint equation for flow in the opposite direction. As long as there is some controllable generation or load on each end of the transmission line (as is likely to occur in a meshed network) and some probability that the flow will reach its limit in either direction, every generator or load affecting flow on that network element will be mispriced under the OFA model under some conditions.

Inefficient dispatch of unscheduled generators and loads

In addition to the impact on scheduled generators it is also worth looking at the impact of the OFA model on unscheduled generators and loads. The OFA model explicitly excludes non-scheduled generators. The Technical Report states:

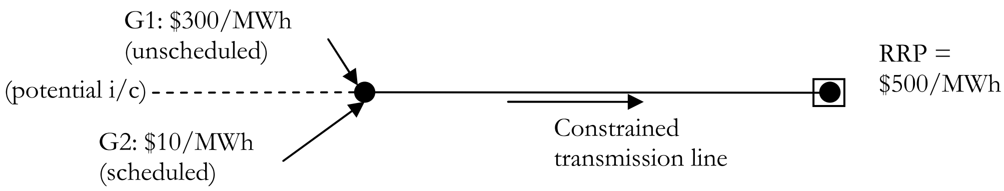
> "non-scheduled generators are not dispatched by AEMO and so cannot be constrained off by transmission congestion".[13]

[13] TR page 52.

This statement seems to imply that there are no dispatch inefficiency consequences arising from non-scheduled generators. This conclusion would be quite incorrect. Non-scheduled generators respond to the wholesale market price they are paid. If that market price is wrong, non-scheduled generators will make inefficient production decisions. In fact, to an extent, the dispatch inefficiency problem of non-scheduled generation is even worse than for scheduled generation since non-scheduled generators have an effective "priority" in dispatch.

This can be illustrated with a simple two-node, one-link example as set out below. Let's suppose the RRP is $500/MWh. There is remote unscheduled generation with a variable cost of $300/MWh. At the high RRP, this generation would like to produce as much as possible. But unscheduled generation does not have to submit an offer curve – it can simply choose to produce as much as it likes. In effect, all the other scheduled generation has to accommodate the increased output of the unscheduled generation. Unscheduled generation and load has an effective priority in the dispatch process.

In this simple example, there is some other low-cost ($10/MWh) scheduled generation which is forced to back off. It cannot avoid being backed off even if it offers its output at $-1000/MWh. This is clearly inefficient, since a $300/MWh generator is displacing a $10/MWh generator. Again we see a pure social waste. Furthermore, if the unscheduled generation increases its output sufficiently this can cause power to flow counter-price on an affected interconnector (indicated by the dashed line in this example). Counter-price flows on interconnectors are another undesirable by-product of mispricing. In this example, the counter-price flows are caused by the actions of an unscheduled generator.



It might be argued that unscheduled generators are, by definition, insignificant and can be ignored. However, with increasing penetration of small "embedded" generation, such as roof-top solar PV, unscheduled generation is likely to become increasingly important in the NEM. This generation cannot make efficient production decisions if it does not face the correct price.

It is also important for loads to face the correct prices. One of the AEMC's Strategic Policy Objectives is to improve the level of demand side responsiveness in the NEM.[14] Improving demand-side responsiveness is, in part, about creating an environment in which loads are willing and able to respond to price signals.[15] The AEMC clearly recognises this. The term "price signals" appears 53 times in the AEMC Directions Paper in the Power of Choice review. But loads cannot respond to price signals they do not face. The OFA model does not propose to correct the pricing defects for loads. To an extent, dispatch inefficiencies will remain.

Consider for example, a case where a large load is located away from the regional reference node (RRN), as indicated in the simple network below. The large load might be an aluminium smelter, or some other large industrial facility. Let's suppose that the load is somewhat price sensitive and

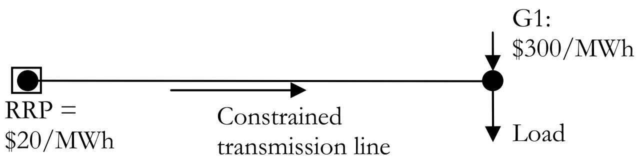
---

14 Strategic Priority Number 2 is "Building the capability and capturing the value of flexible demand". AEMC, "Strategic Priorities for Energy Market Development: 2011", 23 August 2011.

15 AEMC, 2012, "Power of Choice: Giving customers options in the way they use electricity", 23 March 2012.

will cut back its consumption of electricity substantially when the price rises above, say, $100/MWh.

Let's suppose there arises a constraint on power flows away from the RRN to the load node at a time when the regional reference price (RRP) is, say, $20/MWh. At this low level the load decides to produce at full output. However, due to the constraint on the transmission line, this load cannot be served by generation at the RRN. Instead, more expensive local generation must be dispatched. Let's suppose that the local generation is, say, a $300/MWh peaking plant. There is a loss of economic welfare equal to $300-$100 = $200 for each MW of consumption of the load. If the load consumes, say, 100 MW, there is a net loss of efficiency of $2,000 for each hour this situation persists.



I have focused here on the short-term dispatch efficiency consequences. But there are longer-term effects on investment incentives which are arguably even more important. In particular, the failure to send the correct price signals reduces the incentives for efficient location decisions by generators and loads. In addition, mis-pricing of generators and loads can lead to pressures for inefficient investment in transmission. In the example above the failure to provide the correct price signals to the load in aggravates the economic consequences of the congestion on the transmission link. This increases the pressure to augment the transmission link. If loads faced the correct price signals and were willing and able to respond to those price signals the need to augment the transmission and distribution network could be deferred.

In summary, I think it is important to be clear that the OFA model will not result in efficient dispatch outcomes. It does not resolve the mis-pricing problem for non-scheduled generators and loads. The OFA model partially resolves the problem of disorderly bidding for some scheduled generators, some of the time. Mis-pricing of scheduled generators remains whenever a generator has a negative coefficient in a binding constraint equation.

To an extent therefore the OFA model is a partial or 'band-aid' solution. It may be appropriate to implement partial or band-aid solutions in certain circumstances, for example, as an interim measure while fuller solutions are developed, or where all more complete solutions have been assessed and found to have critical drawbacks of their own. However, the OFA model, with its ambitious scope does not appear to be intended as an interim solution. Nor have the alternatives yet been fully assessed. If we are to adopt a complete and lasting reform to the pricing problems in the NEM, in my view, consideration should be given to correcting the mis-pricing problem for *all* market participants – that is, all generators and loads, whether scheduled or non-scheduled.

Nevertheless, some might argue that aiming for greater efficiency might risk making "the best the enemy of the good". Let's therefore go on to explore some of the other potential problems with the OFA model.

Issue #2: Facilitating the provision of hedges to generators and loads

In order to achieve efficient investment in generation (and, to an extent, in loads), there is a need for a deep and liquid market in forward contracts referenced to the local wholesale spot price faced by each generator or load on the network. The current NEM arrangements, under which all generators and loads in a region pay the same regional reference price, manages to achieve a reasonably liquid market for forward contracts in most regions of the NEM (at the expense of mis-pricing, as we have seen above). Any move to correct the mis-pricing problem inevitably involves a move to finer geographic differentiation of pricing and therefore raises the risk of fragmentation of the forward contract market, with a loss of depth and liquidity.[16] As we have noted above, the OFA model introduces a form of finer geographic differentiation of pricing (as we have seen it results in nodal pricing for some scheduled generators). This raises the question of whether the other features of the OFA model ensure that generators and loads have access to the hedge contracts which they require.

The AEMC claims that the OFA model results in "improved support for a deep and liquid contract market"[17] by creating "the ability for generators to hedge the risk of congestion"[18] and that firm access provides "the financial certainty for generators to offer forward contracts on a volume reflective of their access amount"[19]. This section of this paper tests these claims.

Generators and retailers seek to hedge their risks through forward contracts.[20] In the case of conventional thermal generators, these forward contracts will typically take the form of "swap" or "cap" contracts (or other variants). In principle a perfectly reliable generator can eliminate *all* of the wholesale spot price risk that it faces by selling the appropriate portfolio of forward contracts referenced to its local spot price.[21] Similarly, retailers can perfectly eliminate the risk they face by selling appropriate load-following hedge contracts referenced to their local spot price. Therefore, increasing the geographic differentiation of prices does not necessarily result in any change in the availability of hedge contracts provided there is a counterparty prepared to purchase the required hedge contracts referenced to the local spot price for generators and to sell the required hedge contracts referenced to the local spot price for retailers.

This observation is useful as it allows us to frame the problem. The underlying problem is not usefully characterised as creating a mechanism which allows "generators to hedge the risk of congestion"[22]. After all, generators and retailers can always perfectly hedge their risk by selling or buying the appropriate forward contract referenced to their local nodal price. The underlying problem is not that generators or retailers cannot sell the appropriate hedge contracts but that there is *no natural counterparty* for these contracts[23]. The problem therefore is creating an environment in which some other party (which I will refer to as a 'trader') is prepared to act as the counterparty for the contracts which generators and retailers require. Can we create an environment in which traders are willing and able to buy hedge contracts from generators,

---

16 The TR recognises this, noting in its discussion of nodal pricing that one of the effects of a move to more granular pricing is that "retailers and generators find it difficult to contract forward, since each faces a different spot price". TR page 95.

17 SIR page 45.

18 SIR page 48.

19 SIR page 48.

20 The material in this section is explored in more detail in a separate paper.

21 In practice, generators also face the risk of forced outages.

22 SIR page 48.

23 At least there is no natural counterparty other than the other generators and retailers at the same network location.

referenced to the local nodal price of the generators, and sell hedge contracts to retailers, referenced to the local nodal price of the retailer?

I discuss this issue in more detail in a separate paper.[24] That paper makes the point that traders who buy hedges from generators and sell hedges to retailers can, at best, eliminate all the *spot price* risk faced by these parties. The traders collectively cannot eliminate *all* risk – there remains the variation in cash-flow that would arise in perfectly integrated electricity industry.[25] So, refining the definition of the problem still further, we can ask the following: Can we create an environment in which traders are willing to act as the counterparty for the hedges that generators and retailers require, while taking on no more risk on themselves than would arise in a perfectly integrated electricity industry?

I suggest that framing the question in this way represents a useful and preferable way to address locational hedging issues. In a separate paper I argue that, in order for traders to act as a counterparty to provide the hedges that generators and retailers require, (a) traders must have access to some form of financial transmission right or instrument; (b) the total payout on all of the transmission rights that are allocated or issued must be *equal* to the settlement residues that arise from geographic differentiation of prices[26]; and (c) the transmission rights or instruments must be packaged or structured in a manner which allows traders to easily use this instrument to back the provision of hedges to generators and retailers.

As we have seen the OFA model does introduce a new form of transmission right, called the "firm access" transmission right. Does this new instrument facilitate the provision of hedge contracts to generators and loads?

According to the paragraph above, the first question we have to ask is whether the total payout on all of the transmission rights is *equal to* the settlement residues that arise from nodal pricing. In the OFA model, by design the total payout on the set of all transmission rights is equal to the settlement residues. If there are funds left over after paying out on the firm access transmission rights (which might happen if the volume of firm access transmission rights allocated is small), the OFA model envisages that the remaining funds will be paid out in the form of the secondary non-optional rights. On the other hand, if there is a shortfall in the funds required to support the firm access transmission rights (which, putting aside outages, might happen if the volume of firm access transmission rights allocated is larger than the network can handle), the OFA model envisages that the payments will be scaled back.

There is a sense, therefore, in which the first condition above is always satisfied. However, we can and should go further. It seems that the intention of the OFA model is that the firm access transmission rights will, at least in normal operating conditions, have a payout which is as similar as possible as a *fixed volume* transmission right. The expectation seems to be that these fixed volume transmission rights are sought by traders in order to hedge the risks they face (this is discussed further below). If the TNSP allocates too large a volume of these fixed volume transmission rights, the total payout obligation will exceed the settlement residues and therefore the volume of each transmission right will have to be scaled back. This scaling back of transmission rights reduces their value as a hedging device. Therefore, I will assume that the TNSP seeks to allocate no more of these rights than can be accommodated on the network without scaling.

---

24 Biggar, D., "Designing Transmission Rights to Facilitate Hedging in the NEM", October 2012.

25 This variation in cash-flow arises from variation in demand and supply conditions – such as changes in load which affects both the level of revenue received in the industry as a whole and the cost of generation to meet that load.

26 The settlement residues are also known as the merchandising surplus or the congestion rents. For the purposes of this paper I will take these terms to be synonyms.

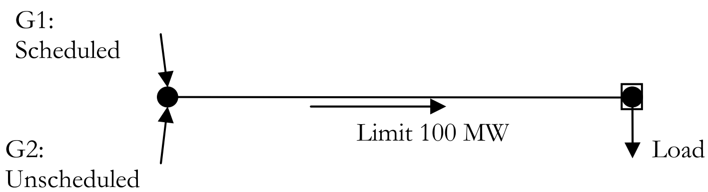
It is well known from the theory of fixed-volume transmission rights, that the total payout on the transmission rights will always be less than the settlement residues (and therefore no scaling would be required) provided the full set of transmission rights corresponds to a set of feasible flows on the network – that is, provided the transmission rights correspond to a physical flow which can itself be accommodated on the physical network. This is known as the "simultaneous feasibility" test. So, let's tentatively assume that the TNSP chooses to allocate no more firm access transmission rights than can be simultaneously accommodated on the underlying physical network. Putting aside outages (which are discussed further below), this condition guarantees that the TNSP will always have sufficient funds from the settlement residues to meet its financial obligations under the firm access transmission rights without engaging in scaling. But we noted earlier that in order for the traders to provide the hedges which generators and retailers require, they must have access to a flow of funds, through the transmission rights, which is *equal* to the settlement residues. Is it the case that when the TNSP imposes this requirement (which is necessary to prevent scaling back of the firm access rights), the total payout obligation on these firm access rights will be equal to the settlement residues?

I claim that when we impose this condition, there will always arise some situations where the payout obligation on the firm access transmission rights falls short of the settlement residues. In this case, the remaining funds are paid out in the form of the secondary rights, which are of little value for hedging purposes. In other words, there is no way to define the fixed-volume transmission rights so that the payout on those transmission rights is equals equal to the settlement residues. There is therefore no way for traders to collectively provide all the hedges which generators and loads require while taking on the minimum possible risk on themselves.

This point might at first appear technical, but can be further clarified using some simple examples.

Let's start by observing how the problem arises in simple radial networks when we exclude some generators or loads.

Let's take the simplest possible case of a two-node, one-link network. The link has a flow limit of 100 MW. There are two generators located "behind" the constrained transmission line, one scheduled and one unscheduled. The unscheduled generator has an output which varies between, say, 0 MW and 100 MW.



The relevant constraint equation for this simple network is:

$$G_1 + G_2 \leq 100$$

Again, let's suppose that the TNSP wants to allocate as large a volume of firm access transmission rights as possible, but subject to the requirement that it will never have to scale back the volume due to a lack of funds. How much entitlement can we allow generator 1 to purchase while ensuring that the payout obligation under the transmission rights is less than the settlement residues? We can ensure that this condition is satisfied provided the entitlement to generator 1 plus the output of generator 2 is less than 100 MW. But generator 2 is excluded from the OFA scheme (its entitlement is equal to its actual output) therefore the entitlement allowed to generator 1 must satisfy:

$$E_1 \leq 100 - G_2$$

But, since generator 2 is assumed to be able to produce up to 100 MW, the entitlement to generator 1 must not exceed 0 MW. At any time when the output of generator 2 is less than 100 MW the payout on the transmission right allocated to generator 1 will be less than the settlement residues. It is not possible to allocate transmission rights which in some sense "match" the maximum flows on the network.

The TNSP could increase the entitlement to generator 1, above 0 MW, to say 50 MW. But, in doing so, it would have to scale back the payout on the firm access transmission right whenever the output of generator 2 increased above 50 MW.

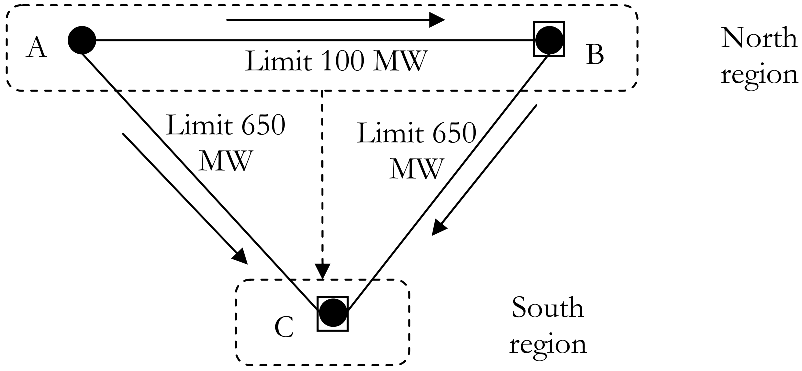
As this example shows, it is not possible to ensure that the payout on the firm access transmission right matches the settlement residues under all market conditions while avoiding scaling. There is a trade-off: The TNSP may ensure that the firm access transmission rights have payout corresponding to a guaranteed fixed volume (as the AEMC envisages is required for hedging purposes). But if it does so, it has to choose a volume of rights which at least in some circumstances results in a total payout which falls short of the settlement residues. On the other hand, if the TNSP allocates a higher volume of transmission rights, there will inevitably arise circumstances where scaling is required, reducing the value of the firm access right as a hedging instrument.

In this example the limitation on the entitlements that can be offered arises from the presence of an unscheduled remote generator. But exactly the same problem would arise in a network with a remote load, or a remote generator with a negative coefficient in a binding constraint. The decision to exclude loads, unscheduled generators, and generators with a negative coefficient in a binding constraint in the OFA model implies that it is not possible to allocate a set of entitlements which ensures that the payout obligation under the corresponding transmission rights is equal to the settlement residues.

But what if we did include all generators and loads in the OFA model? What if the OFA model were made completely comprehensive, covering all generators and loads? Would the approach to transmission rights proposed by the AEMC allow us to define a set of firm access transmission rights which yield a payout equal to the settlement residues without scaling?

The answer is no. It is not always possible to define a set of fixed-volume transmission rights in such a way that the payout obligations on the transmission rights are always equal to the settlement residues.

This is illustrated in the following example. This is a network with three nodes, three links, and two regions. (The outcome here does not rely at all on the fact that there are two regions. Exactly the same issues would arise if there were a single region.) Nodes A and B are in the "North" region (with node B the regional reference node), and node C is in the "South" region. There is a single notional interconnector between the North region and the South region. The load is located at the regional reference nodes (B and C). There is a limit on flow from A to B (an intra-regional constraint) and from A to C and B to C (which are mixed constraints).



If all the links A-B, A-C, and B-C have identical electrical characteristics the correctly-oriented constraint equations for this network are as follows:

$$\frac{1}{3}G_A + \frac{1}{3}F_{N\to S} \leq 650 \text{ and } \frac{2}{3}G_A - \frac{1}{3}F_{N\to S} \leq 100 \text{ and } -\frac{1}{3}G_A + \frac{2}{3}F_{N\to S} \leq 650$$

It follows that we could achieve a payout on the entitlements equal to the settlement residues if and only if we can find a pair of entitlements which satisfy the following equations:

$$\frac{1}{3}E_A + \frac{1}{3}E_{N\to S} = 650 \text{ and } \frac{2}{3}E_A - \frac{1}{3}E_{N\to S} = 100 \text{ and } -\frac{1}{3}E_A + \frac{2}{3}E_{N\to S} = 650$$

But there is no pair of entitlements which satisfy all three equations. At most two of the equations can be satisfied at one time. For example, the first two equations could be satisfied by choosing an entitlement for the generator of $E_A = 750$ and an entitlement for the interconnector of $E_{N\to S} = 1200$. But then we would have

$$-\frac{1}{3}E_A + \frac{2}{3}E_{N\to S} = 550 < 650$$

The conclusion that we reach is the following: even in the theoretically ideal world of nodal pricing where all of the generators and loads and interconnectors are included in the mechanism, it is not possible to define a set of entitlements for fixed-volume transmission rights which ensure that the payout on the transmission rights is equal to the settlement residues, so that all of the settlement residues are made available to the market without scaling. Even in a theoretically ideal world the approach proposed by the AEMC in the OFA model cannot ensure that traders will be able to provide the hedges that generators and loads require.[27]

Even if the OFA model did make available, through the firm access transmission rights, all of the settlement residues, this is still not a sufficient condition for achieving a deep and liquid hedge market. As I noted above, in addition in my view there is a second requirement: we require that the model should package those settlement residues in a manner which allows traders to provide hedges associated with a range of common transactions.

[27] It is possible to define an entitlement which ensures that the total payout on the transmission rights is equal to the settlement residues if we allow the entitlement to be varying with the output of some generators or loads, or with the identity of the binding constraint equation. The problem with this approach is that it does not facilitate hedging – that is, it does not package the settlement residues in a manner which allows traders to provide the hedges that generators and loads require.

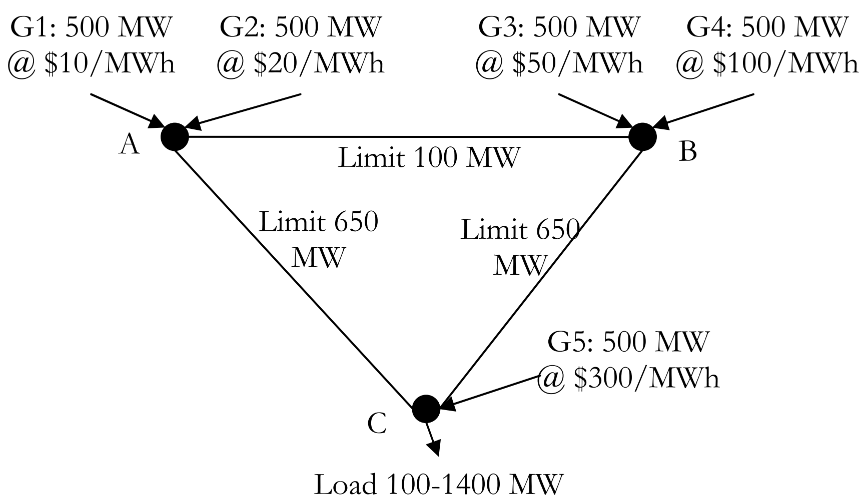
The OFA model creates a form of fixed-volume transmission right, with parallels to a swap contract. A fixed-volume transmission right pays out the price difference between the reference node and the local nodal price multiplied by a *fixed* volume. A fixed-volume transmission right is a useful instrument for hedging a fixed-volume transaction between a generator and a retailer.

However, some (or perhaps most) generators in the market have a volume of production which *varies with market conditions* such as demand and/or the local wholesale spot price. Most generators in the market will increase their production if the local wholesale spot price goes high enough. Some generators, such as peaking generators, will produce nothing at all on most days of the year, but will produce as much as they are able at times of high spot prices.

Let's consider the position of a trader who seeks to sell a cap contract to a peaking generator and then to match that contract with a corresponding contract with a retailer with exactly the same production and consumption profile. Can this trader reduce the risks it faces to the theoretical minimum using a fixed volume transmission right?

The answer is no. A fixed-volume transmission right pays out a price difference between two nodes multiplied by a fixed volume at any time of the day or night. But the generator is only producing (and therefore the hedge is only required) at times when the local spot price is high. A financial instrument such as the firm access transmission right proposed by the AEMC which pays out the price difference between two nodes at *any time* of the day or night may *increase* the risk faced by this trader.

This can be easily illustrated with a simple example. Consider the following simple three-node, three-link network. There are five generators, each with a distinct variable cost. The load varies from 100 to 1400 MW. Each of the three transmission links has its own maximum power flow.[28]



The efficient pricing, dispatch and flow outcomes under optimal dispatch in this simple network are set out in the following table:

[28] This network example is simplified in various important ways. For example, there is a single load node. There is no intermittent generation. Generation and transmission outages are ignored. The transmission limits are simple thermal limits which can be expressed as a single fixed number. None of these assumptions will remain true in the actual NEM.

Load	Dispatch (MW)					Disp.	Flows			Prices			Cong.
(MW)	G1	G2	G3	G4	G5	Cost	A->B	A->C	B->C	A	B	C	Rent
100	100	0	0	0	0	1000	33.3	66.7	33.3	10	10	10	0
200	200	0	0	0	0	2000	66.7	133.3	66.7	10	10	10	0
300	300	0	0	0	0	3000	100.0	200.0	100.0	10	10	10	0
400	350	0	50	0	0	6000	100.0	250.0	150.0	10	50	30	6000
500	400	0	100	0	0	9000	100.0	300.0	200.0	10	50	30	6000
600	450	0	150	0	0	12000	100.0	350.0	250.0	10	50	30	6000
700	500	0	200	0	0	15000	100.0	400.0	300.0	10	50	30	6000
800	500	50	250	0	0	18500	100.0	450.0	350.0	20	50	35	4500
900	500	100	300	0	0	22000	100.0	500.0	400.0	20	50	35	4500
1000	500	150	350	0	0	25500	100.0	550.0	450.0	20	50	35	4500
1100	500	200	400	0	0	29000	100.0	600.0	500.0	20	50	35	4500
1200	500	250	450	0	0	32500	100.0	650.0	550.0	20	50	35	4500
1300	500	150	500	150	0	48000	0.0	650.0	650.0	20	100	180	156000
1400	500	150	500	150	100	78000	0.0	650.0	650.0	20	100	300	312000

Now let's suppose that, in keeping with the OFA model, the TNSP makes available a form of fixed-volume financial transmission right from each generator's local node to the load node (which in this case is node C).

Let's suppose that (following some allocation process) the TNSP makes the following transmission rights available to the market: 750 MW of firm transmission right from node A to node C and 450 MW of firm transmission right from node B to node C.[29]

As we have seen, a central claim of the AEMC is that the OFA model creates the ability for generators to hedge the risk of congestion. Let's consider a transaction between a generator and a retailer under which the retailer agrees to purchase all of the output of the generator, say, G2. A trader seeks to provide a hedge to generator G2 which eliminates all its risk, while simultaneously providing a hedge to the corresponding retailer to eliminate its risk. The trader then seeks to back these hedges using some fixed volume of firm access transmission rights. In doing so, the trader would like to reduce the risk it faces to the minimum theoretical level (which is the level that would arise if the generation and retailer were vertically integrated).

The second column of the following table shows the net profit of the trader in the hypothetical case in which the generator and the retailer were vertically integrated. This column reflects the minimum theoretical risk which cannot be removed by contracting between generators and loads. The remaining columns show the net profit of the trader after it has procured various different levels of firm access.

[29] These numbers have specifically been chosen so as to make maximum use of the available transmission network without having to resort to scaling back the volume of transmission rights. It is not possible to increase the rights allocated against the node B price without reducing the rights allocated against the node A price if the total payout of the transmission rights is to be less than the settlement residues. There are a number of other possible combinations of transmission rights which also satisfy this condition. For example, the combination of 650 MW from A to C and 650 MW from B to C is another possible combination. Any allocation process for the firm access rights must mediate between these alternative combinations.

Load	Residual risk (integrated entity)	Trader profit with 0 MW firm access	Trader profit with 150 MW firm access	Trader profit with 250 MW firm access
100	$0	$0	$0	$0
200	$0	$0	$0	$0
300	$0	$0	$0	$0
400	$0	$0	$3,000	$5,000
500	$0	$0	$3,000	$5,000
600	$0	$0	$3,000	$5,000
700	$0	$0	$3,000	$5,000
800	$1,750	$1,000	$3,250	$4,750
900	$3,500	$2,000	$4,250	$5,750
1000	$5,250	$3,000	$5,250	$6,750
1100	$7,000	$4,000	$6,250	$7,750
1200	$8,750	$5,000	$7,250	$8,750
1300	$5,250	-$18,750	$5,250	$21,250
1400	$5,250	-$36,750	$5,250	$33,250

What volume of "firm access" should the trader procure? Since G2 produces as much as 250 MW, the trader might seek to procure 250 MW of access rights from node A to node C. Does this (or any) volume of access rights reduce the risk faced by the trader to the minimum theoretical level? As can be seen in the table above, if the trader acquires 250 MW of access rights, the trader's risk is not reduced to the minimum theoretical level (that is, the numbers in column 5 are not the same as the numbers in column 2). The same holds true no matter what level of access rights the trader acquires. There is no level of access rights which this trader can purchase which reduces its risk to the minimum theoretical level.

We can carry out the same exercise for transactions involving any of the generators in this example. All of the generators in this simple network example have a volume of output which varies with the local nodal spot price. The fixed-volume transmission rights created in the OFA model do not facilitate the provision of hedges between generators and loads where the volume of that transaction varies with the wholesale market price. The OFA model therefore does not facilitate the provision of hedges in this example.

In a separate paper I propose an alternative approach to the design of transmission rights. This alternative approach allows the trader to acquire a portfolio of rights with a volume which varies in line with the output of any given generator in the market.[30] This approach allows traders to reduce the risk they face in hedging transactions (such as those above) to the theoretical minimum.

In mentioning this alternative approach, I am not advocating that it should be adopted instead of the AEMC proposal. The alternative approach requires further thinking and analysis before it could be developed into a full-blown proposal. However I am merely seeking to point out that there may exist even better approaches to managing inter-nodal pricing risk following a move to nodal pricing. Good public policy requires that we identify and explore alternatives. In my view, it has not yet been established that the OFA model is the best approach to addressing the hedging issues that arise from a move to nodal pricing.

[30] Biggar (2012), "Designing Transmission Rights to Facilitate Hedging in the NEM".

However, rather than asking whether or not we can do better than the OFA model, it might be argued that we should only ask whether or not the OFA model is better than the status quo?

The problem with this line of argument is that good public policy requires that we assess the costs and benefits of *all* the options which can address the underlying public policy problem. It is not good public policy to select one option and then compare only that option to the status quo. There may be other options which achieve the objectives more effectively or at a lower cost. There is an analogy here with the RIT-T. In assessing a transmission augmentation we expect transmission businesses to consider *all* alternatives which deliver the required level of services. We would not be satisfied if a transmission business selected one alternative and claimed that the benefits exceed the cost, leaving other possible alternatives untested. In my view it is too early to reject analysis or consideration of at least some potential alternatives to the OFA model.

Nevertheless, we might ask whether or not the OFA model is better than the status quo. This is a very difficult question to answer. Under the status quo arrangements in the NEM, generators and loads *automatically* receive a transmission right with a volume which perfectly matches their own output. This allows them to perfectly hedge the risk of trading with the regional reference node. However, granting generators and loads a financial transmission right with a volume which matches their own output gives rise to the problem of mis-pricing and disorderly bidding.

It is not immediately clear whether the harm from mis-pricing and disorderly bidding (which results in inefficient dispatch and volume-based dispatch risk) is larger or smaller than the harm that would arise from a move to fixed-volume transmission rights (which, as we have seen, might limit the ability for generators or retailers to find counterparties for hedge contracts referenced to their local nodal price). Under the OFA model, traders may prove reluctant to act as the counterparty and to provide hedges for at least some transactions. This may reduce the liquidity in the hedge market. The original NEM designers seem to have favoured the depth in the hedge market over the problem of disorderly bidding. The OFA model partially solves the problem of disorderly bidding, but it may lead to a reduction in the availability of hedge contracts. At this stage, it is not clear to me that the OFA model represents an improvement over the status quo.

Issue #3: Improving the responsiveness of TNSP actions to market conditions

As discussed above, under normal market conditions I would expect that the level of firm access rights provided to the market will be chosen in such a way that the payout obligation of the system operator arising from the transmission rights is less than or equal to the settlement residues. However, this condition will not necessarily continue to hold when an outage occurs on the transmission network. In this case the settlement residues will typically fall short of the payout obligations under the transmission rights.

The AEMC proposes that under the OFA model, TNSPs would be required to fund a portion of this shortfall. The AEMC claims that this would "create financial incentives on TNSPs to maximise network availability when it is most valuable" and that it would "provide a strong signal to TNSPs to manage the network consistently with the way in which capacity is valued by the market".[31]

It is certainly desirable to establish financial rewards and penalties on TNSPs which induce them to take actions which promote economic welfare. Since wholesale market conditions vary rapidly from one dispatch interval to another, it makes sense for those financial rewards and penalties to

[31] SIR page 54.

vary with market conditions.[32] However economic efficiency requires that the magnitude of the financial reward or penalty on a transmission business should be directly commensurate with the economic benefit or harm caused. If the financial incentive is too large, the TNSP will be induced to take certain actions to avoid outages even when the cost of those actions is larger than the economic benefit. If the financial incentive is too small the TNSP will forego potentially beneficial actions to reduce outages even when the economic benefit of those actions exceeds the cost.

The AEMC proposes to make TNSPs liable for a portion of the shortfall in the settlement residues relative to the financial payout under the transmission rights. The question for us, therefore, is whether or not it is efficient to make a TNSP liable for a fixed portion of the shortfall in the financial payout obligation under the transmission rights. Is it the case that we can choose that proportion in such a way that the resulting financial incentive on the TNSP is equal to the underlying economic benefit or harm?

The answer is no. We cannot set the financial reward or penalty equal to a fixed proportion of the shortfall in funds on the transmission rights and have that reward or penalty be equal to the underlying economic benefit or harm.

This can easily be seen with simple examples. Let's focus first on the case where an outage on a transmission line reduces the line's capacity but does not cause the line to trip out of service entirely. The reason for the distinction is that when a line trips out of service there is an immediate redistribution of power flows around the remaining network which complicates the analysis (discussed further below).

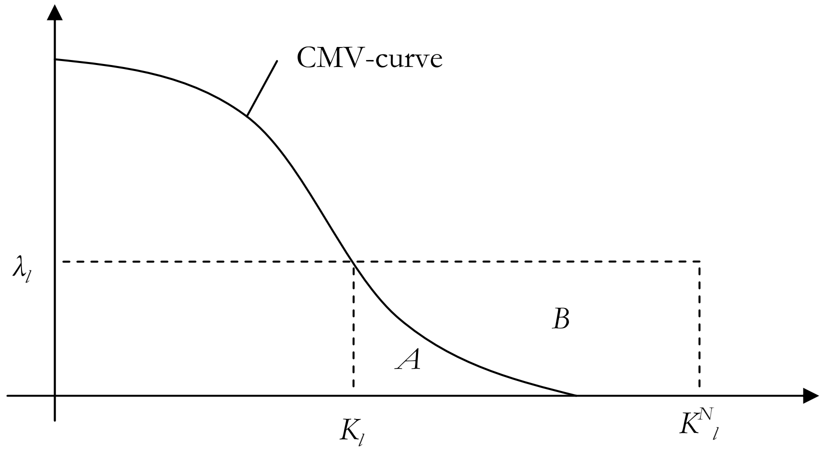
Let's assume for simplicity that there is just a single potentially binding transmission constraint (labelled *l*). Let's suppose that the normal network capacity associated with this transmission constraint (that is, the right hand side of the constraint equation) is K_l^N. Let's suppose that all of this capacity has been allocated in fixed-volume transmission rights. The payout obligation on the transmission rights is then $\lambda_l K_l^N$ where λ_l is the constraint marginal value for the binding constraint (the AEMC documents refer to this as the "flowgate price").

But now suppose that through some action or inaction of the TNSP, the out-turn network capacity K_l turns out to be lower than the promised network capacity K_l^N. The total settlement residues are therefore $\lambda_l K_l$. The shortfall in the funds required to finance the transmission right obligations is therefore:

$$\lambda_l (K_l^N - K_l)$$

We would expect that the lower the out-turn network capacity the higher the constraint marginal value. We can draw a curve which reflects, for each level of out-turn network capacity, the resulting constraint marginal value, as illustrated below:

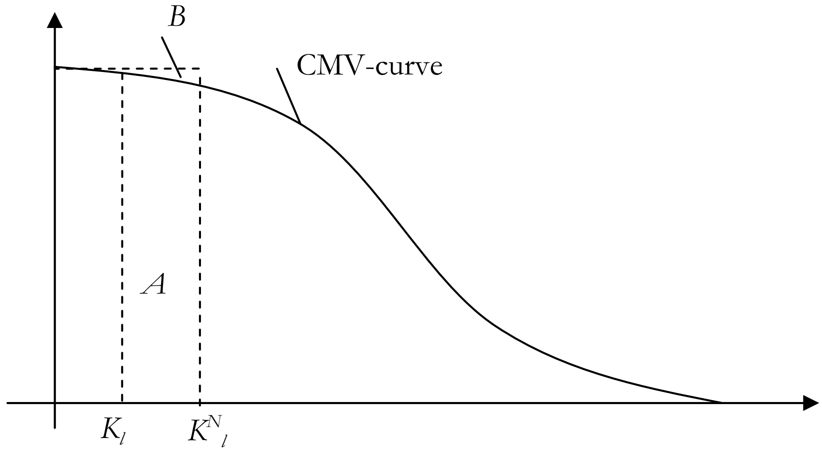
[32] The AER has for a long time sought to develop and improve its "Market Impact" indicators as part of the Service Target Performance Incentive Scheme on transmission businesses.



It turns out that the overall economic harm arising from a reduction in capacity from and K_l^N to K_l is the area under the CMV-curve – which is the area A on the diagram. But the shortfall in funding obligations is $\lambda_l(K_l^N - K_l)$, which is the area A+B on the diagram. We can conclude that if the TNSP were made liable for the full shortfall in funds, the TNSP would be systematically over-incentivised to prevent the outage. Economic efficiency would not be achieved.

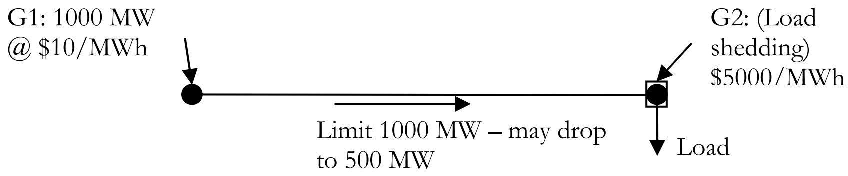
But let's suppose that the TNSP is only required to make up a proportion of the shortfall. For example, let's suppose that the TNSP is required to make up, say, 50 per cent of the shortfall. Does this guarantee an economically efficient outcome?

Again the answer is not necessarily. As the next diagram shows, in some cases the reduction in economic welfare can be as large as the full shortfall in funds. In this case, exposing the TNSP to only a proportion of the shortfall will result in the TNSP be systematically under-incentivised to prevent the outage.



We can conclude that making the TNSP liable for any fixed proportion of the shortfall in funds required to finance the transmission rights will result in a financial penalty which may be larger or smaller than the underlying economic harm. It is not possible to use the shortfall in funds on transmission rights to achieve efficient incentives on transmission businesses.

These points can be made clearer with some simple network examples. In the following simple network there is a single unreliable transmission link. This link normally has a capacity of 1000 MW – but this may reduce to 500 MW due to action or inaction by the TNSP. The TNSP has issued 1000 MW of transmission rights on this link.



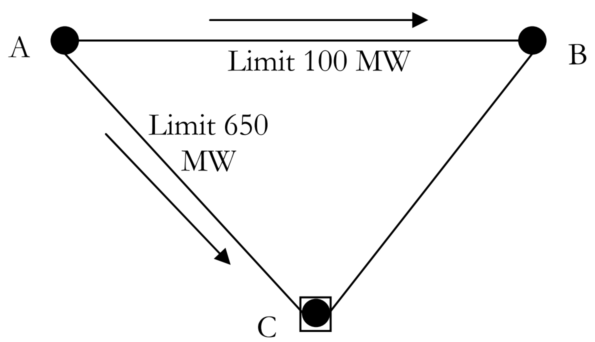
Let's suppose the load at the right hand node is initially 501 MW. The failure of the link will result in a binding constraint and a price difference of $5000-10=$4990/MWh. The shortfall on financial obligations on the transmission rights is therefore (1000-500)=500 times $4990 or $2,495,000 per hour. However, the economic harm from the action is the reduction in load of 1 MW, with an economic cost of $5000-$10=$4990/MWh (there is a savings in generation cost of $10/MWh from not having to serve this load). In this circumstance the TNSP is *significantly over-incentivised* to prevent this outage.

Let's suppose that, recognising this possibility, the AEMC only requires the TNSP to make up, say, half of the shortfall in funds. Now let's suppose that the load at the right hand node is 999 MW. The failure of the link again results in a shortfall on the financial obligations of $2,495,000 per hour, but the economic harm is the shed load of 499 MW with an economic cost of 499 times $4990 or $2,490,010 per hour. If the TNSP is only responsible for the half of the shortfall the TNSP will be *significantly under-incentivised* to take an efficient action.

Again we see the key result – if we make a TNSP liable for any fixed proportion of the shortfall in funds required to finance transmission rights the TNSP will be either under-incentivised or over-incentivised to take the efficient action, depending on the circumstances.

The examples above focused on the case of a partial outage of the transmission line. The case where an outage causes the complete loss of a transmission line is more complicated because the loss of a transmission line causes power flows to change on the network. In fact it is possible for the outage of a line to *improve* overall economic outcomes.

This might happen, for example in the following three-node network. When the line from A to B is in service it restricts the amount that generators at A can produce. In these circumstances, taking the line from A to B out of service may increase overall economic welfare. Yet, it may still be desirable to maintain the line from A to B in service for reliability reasons: if the system operator is unable to bring the line from A to B back into service quickly, it may be preferable to maintain this line in service to protect against the risk of an outage on one of the other lines which might lead to load shedding.



There is value in continuing to explore ways to make TNSPs more responsive to market conditions, to ensure that they make the maximum network capability available at times when the that capability is most valuable to market participants. However in my view the arguments above have shown that exposing TNSPs to the risk of a shortfall in congestion rents relative to the financial access right obligations does not bring about efficiency in decisions to maintain network capability.

As noted earlier, in a separate paper I have proposed an alternative design of transmission rights. One of the consequences of that proposal is that there is a direct link between the financial shortfall arising on the set of all the hedge contracts in the market and the total economic harm from a transmission outage. Therefore, in that alternative approach it is in principle economically efficient to make TNSPs liable for the shortfall in both hedge contracts and financial transmission rights as a mechanism for incentivising TNSP behaviour. As noted above, I do not raise this in order to advocate for this alternative proposal. I merely raise it to highlight that there are likely to be alternative approaches to incentivising TNSPs which can achieve better outcomes than would arise in the OFA model.

Issue #4: Improving the efficiency of generation and transmission investment

Under the OFA model, generators will be allowed to procure firm access from the TNSP who will then be obliged to upgrade the network to provide the corresponding level of access. The AEMC writes:

> "The purchase of firm access by generators would fund and guide network expansion, with TNSPs required by the firm access standard to plan the network to meet all firm access concurrently"[33].

TNSPs "must expand the network to accommodate firm access".[34] The AEMC claims that this will lead to more efficient transmission investment and should "encourage co-optimisation of transmission and generation investment".

Economic efficiency requires that generation and transmission investment is coordinated. Is it the case that generator-led procurement of firm access will necessarily result in overall efficient co-optimisation of generation and transmission?

In my view the answer is no. Intuitively, the reason is that a transmission augmentation creates winners and losers. Allowing any sub-group of the total set of winners and losers to drive

[33] SIR page 50.
[34] TR, page 95.

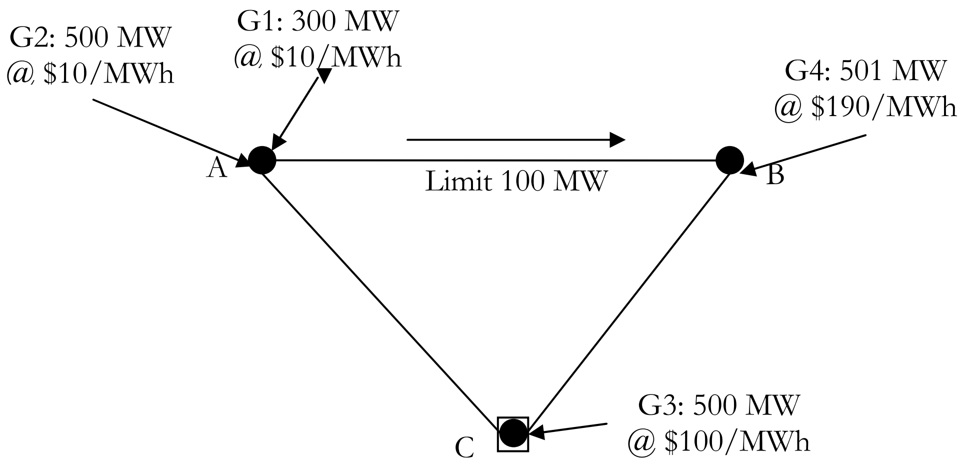
transmission expansion risks permitting transmission augmentations which are privately beneficial but not socially desirable.

In particular, a transmission augmentation will usually benefit generators in an exporting region and loads in an importing region. Conversely, an augmentation will usually hurt generators in an importing region and loads in an exporting region. Allowing generators in an exporting region (or loads in an importing region) to drive transmission expansion will inevitably result in transmission expansion which is socially inefficient. In particular, generators in an exporting region will benefit from a transmission augmentation which increases the price they receive, even if there is little or no social benefit from that augmentation. This is known as the "business stealing" effect of a transmission augmentation.[35]

We can illustrate this result using the following three node network. In this network there is 800 MW of $10/MWh generation at node A. The link from A to B has a limit of 100 MW. There is also 501 MW of generation at node B. With this configuration, the constraint between A and B is binding and the nodal price at A is $10/MWh. In the OFA model there are also some transmission rights from node A to node C (the reference node). But what volume of such transmission rights can be offered? The volume that can be offered depends on the output of the generator at node B. We may presume that the TNSP will take a conservative view on the output of the generator at node B when making a decision as to how much firm access transmission rights to make available. In particular, let's take the worst case scenario, where the output of the generator at node B is zero. In this case only 300 MW of firm access rights can be offered at node A. Let's suppose that generator G1 obtains all of these firm access rights, leaving generator G2 exposed to the local nodal price ($10/MWh).

Let's assume that at a particular time, the generator at B is producing, say 499 MW. This allows G2 to also produce almost at its capacity (499 MW), but G2 only receives the local nodal spot price ($10/MWh). Now let's assume that G2 is allowed to bring about an augmentation to the link from A to B by procuring another few MW of firm access transmission rights. The TNSP responds by augmenting the link A to B slightly. But now the link from A to B is no longer binding. The local nodal spot price at A jumps up to $100/MWh. Generator G2 receives a benefit of ($100-10) times 499 MW per hour or $44,910 per hour. Yet, since the size of the augmentation is small, the social benefit from the augmentation will be small. The private benefit to generator G2 may well be many times larger than net social benefit.

35 See Hogan, "Transmission Benefits and Cost Allocation", 31 May 2011.



This example illustrates the general principle that a transmission augmentation will in general create winners and losers. As a consequence the total private benefits will almost always exceed the net or social benefits. Allowing any subgroup the power to bring about a transmission augmentation will not achieve efficient transmission investment.

At present transmission investment is primarily carried out by regulated companies who do not respond directly to market signals. There are clear potential benefits from developing more market-based signals for transmission investment. However, at this stage I am not aware of any mechanism in the theoretical literature which links the allocation of fixed-volume transmission rights with market-led transmission investment and which achieves efficient transmission investment decisions. To my knowledge no such theory exists.

In a separate paper I have proposed an alternative design of transmission rights. One of the implications of that design is that, when the transmission network is augmented the payout to the market intermediaries introduced earlier, known as traders, changes by precisely the social value of the augmentation. In other words, under this alternative proposal, transmission investment should be approved if and only if it is supported by a coalition of all the traders. Further work is needed to explore the implications of this model. However, at this stage we can note that if the problem is achieving market-led investment in transmission it is not yet clear that the OFA model will achieve this or that it is the best way to achieve this.

Conclusion

The AEMC has proposed one possible set of measures to address a range of real issues in the NEM. The AEMC has put up this package of measures as an alternative to the status quo and, to a certain extent, has discouraged discussion of the components in isolation.[36] This paper has analysed the proposal of the AEMC and finds that, for a variety of reasons, it will not necessarily promote overall economic efficiency. Specifically, I have raised the following concerns:

36 "We ... would urge respondents to avoid promoting adoption of elements of the optional firm access proposal in isolation of what is an integrated and interdependent package". SIR page vii.

- Although the model does resolve the disorderly bidding problem for some generators, the model does not address the problem of mis-pricing of scheduled or unscheduled loads, unscheduled generators, or scheduled generators with a negative coefficient in a binding constraint equation. An important element for improving demand side responsiveness in the NEM is exposing loads to the correct wholesale market price. The AEMC has emphasised the importance of correct price signals in its Power of Choice review. The OFA model does not achieve this. Consideration should be given to expanding the model to include all generators and loads.

- The proposed model creates a form of fixed-volume financial transmission right. However it will not be possible to choose the volume of the transmission rights which can be offered under the proposal in such a way that the payout obligation on the firm access transmission rights is just equal to the settlement residues. Furthermore, fixed volume transmission rights are not a useful way of packaging the settlement residues to allow traders to back transactions involving a price-dependent volume, which is the case for most generators in the NEM. Consideration should be given to re-designing the transmission rights to better match the transactions carried out by market participants.

- The model exposes TNSPs to some of the shortfall in settlement residues brought about by transmission outages. But the shortfall in settlement residues is not related in a linear manner to the economic benefit from avoiding outages. TNSPs will be either over-incentivised or under-incentivised to prevent outages. Consideration should be given to alternative approaches (including alternative designs of transmission rights) under which the shortfall in funds is commensurate with the underlying economic benefit.

- The model proposes to require TNSPs to expand the network in response to requests from generators. But exporting generators may have an incentive to augment the network to enhance the local marginal price even if the augmentation is socially inefficient. In other words the AEMC proposal will not lead to efficient coordination of generation and transmission investment. Consideration should be given to alternative approaches which might allow for market-driven transmission investment.

In short, the AEMC can be commended for putting up an ambitious model designed to address many protracted problems in the NEM. However, before implementing such a major reform it seems desirable to me to clearly identify alternative approaches which may address the issues raised. There is a need for substantial analysis of these alternatives before a decision can be made between this – or any other – package of reforms and the status quo.

Appendix A: The mathematics of the OFA model

This appendix is an attempt to set out the key equations of the OFA model in one place.

Let's suppose we have a wholesale electricity spot market, with a similar market design to the NEM, but with scope for increased geographic differentiation of charges. Generators and loads are indexed by *i* and are assumed to be grouped together into "regions" labelled *r*. In each region there is a designated node known as the regional reference node. There are notional interconnectors between the regions, labelled *c*. Interconnector *c* joins the region *fr(c)* to the region *to(c)*.

Generators and loads located at the regional reference node are assumed to buy and sell wholesale electricity at the wholesale spot price at the regional reference node (known as the regional reference price) denoted P_r^R. In contrast, remote generators and loads may buy or sell electricity at a price which may be different to the regional reference price. Generators and loads at node i will be assumed to transact electricity at the local nodal price P_i. The notional interconnectors are assumed to (in effect) purchase electricity in the from-region at the regional reference price and sell the same volume of electricity in the to-region at the regional reference price.

As noted in the text, the OFA model is equivalent to full nodal pricing plus a form of transmission right allocated to all generators and loads. For scheduled generators there are two forms of transmission right – a "firm access" transmission right and a secondary transmission right.

Each generator or load either is allocated or will procure an entitlement on the *l*th constraint equation denoted E_{li}. The total payment to the generator or load from the transmission right (on top of the revenue received by selling its output at the local nodal price) can be written as follows:

$$TR_i = \sum_l \lambda_l \alpha_{li} E_{li} \qquad \text{(A 1)}$$

Here λ_l is the constraint marginal value (the AEMC use the terminology "flowgate price") and α_{li} is the coefficient of this generator in the *l*th constraint equation (the AEMC refer to this as the "participation factor"). The entitlement to a load can be written the same way, but with a negative sign. The total payment to the generator or load from the settlement mechanism is therefore:

$$P_i G_i + \sum_l \lambda_l \alpha_{li} E_{li} = P_r^R G_i + \sum_l \lambda_l \alpha_{li} (E_{li} - G_i)$$

This is intended to be the same as equations 2.1, 2.2, 2.3, 2.5, 2.8, 4.1 and the equations in section 12.2.9 in the Technical Report.

If the entitlement is set at a value which is independent of the constraint which is binding we have $E_{li} = E_i$, and the payment on the transmission right is just equal to the difference between the regional reference price and the local nodal price times the entitlement:

$$TR_i = \sum_l \lambda_l \alpha_{li} E_{li} = \sum_l \lambda_l \alpha_{li} E_i = (P_r^R - P_i) E_i$$

Similarly, we can define an entitlement for an interconnector denoted E_{lc}. The payout on the interconnector transmission right is then:

$$TR_c = \sum_l \lambda_l \beta_{lc} E_{lc} \tag{A 2}$$

Here β_{lc} is the coefficient of the interconnector *c* in the *l*th constraint equation. As before, if the entitlement is set equal to a fixed value $E_{lc} = E_c$, the payout on the interconnector transmission right is then just the price difference between the regions multiplied by the interconnector entitlement:

$$TR_c = \sum_l \lambda_l \beta_{lc} E_{lc} = \sum_l \lambda_l \beta_{lc} E_c = (P^R_{to(c)} - P^R_{fr(c)}) E_c$$

Under the status quo in the NEM, the entitlement for every generator is just equal to its actual production $E_{li} = G_i$. Similarly, the entitlement for every load is just equal to its actual consumption: $E_{li} = L_i$. Similarly, the entitlement for every notional interconnector is equal to its actual flow: $E_{lc} = F_c$.

This has several consequences. One important consequence is that every generator and load faces the regional reference price. The revenue paid to every generator is equal to:

$$P_i G_i + TR_i = (P_i + \sum_l \lambda_l \alpha_{li}) G_i = P^R_r G_i$$

The AEMC proposes to change the current definition of the entitlements for scheduled generators. The entitlement for scheduled generators will now be:

$$E_{li} = \begin{cases} G_i, & \text{where } \alpha_{li} < 0 \\ E_i & \text{otherwise} \end{cases}$$

Here E_i is the "agreed access level" of generator or load *i*. In addition, there is scope to define an entitlement for each notional interconnector.

Under some circumstances the payout obligation on the transmission rights might be less than the available settlement residues. In this case the AEMC proposes to distribute these remaining funds to scheduled generators in the form of a second, non-firm transmission right. Specifically, the AEMC proposes to have a second transmission right which pays out the amount:

$$TR_i^{NF} = \sum_l \lambda_l \Phi_l \alpha_{li} A_{li}$$

Where Φ_l is a parameter (defined later) and:

$$A_{li} = \begin{cases} 0, & \text{where } \alpha_{li} < 0 \\ A_i, & \text{otherwise} \end{cases}$$

Here A_i is the "availability" of the generator. The parameter Φ_l is chosen so as to ensure that the total payout obligation across all the transmission rights is precisely equal to the settlement residues.

It is worth proving an important result: under optimal dispatch (where the prices, dispatch, and flows are chosen to maximise economic welfare) the settlement residues are equal to the sum of the constraint marginal value multiplied by the constraint right hand side. Let's define the congestion rent as follows:

$$CR = \sum_l \lambda_l K_l$$

Here K_l is the constraint limit or right-hand-side. Under optimal dispatch the following conditions hold: From the first-order conditions and complementary slackness conditions:

$$P_r^R - P_i = \sum_l \lambda_l \alpha_{li} \tag{A 3}$$

$$P_{to(c)}^R - P_{fr(c)}^R = \sum_l \lambda_l \beta_{lc}$$

$$\lambda_l (\sum_i \alpha_{li}(G_i - L_i) + \sum_c \beta_{lc} F_c) = \lambda_l K_l$$

If we multiply the first of these equations by $G_i - L_i$ and sum over *i*, and then multiply the second equation by F_c and sum over *c*, and then add the result, using the last equation we find that the settlement residues are equal to the congestion rents:

$$\begin{aligned} SR &= \sum_i (P_r^R - P_i)(G_i - L_i) + \sum_c (P_{to(c)}^R - P_{fr(c)}^R) F_c) \\ &= \sum_l \lambda_l (\sum_i \alpha_{li}(G_i - L_i) + \sum_c \beta_{lc} F_c) \\ &= \sum_l \lambda_l K_l = CR \end{aligned}$$

Furthermore, for a given set of entitlements let's define K_l^E as follows:

$$K_l^E = \sum_i \alpha_{li} E_{li} + \sum_c \beta_{lc} E_{lc}$$

Multiplying both sides of this equation by the constraint marginal value and summing over all constraints we find that the total payout on the transmission rights is given as follows:

$$PO = \sum_i TR_i + \sum_c TR_c = \sum_l \lambda_l K_l^E$$

On the other hand, we proved above that the actual settlement residues or congestion rent can be written as follows:

$$SR = \sum_l \lambda_l K_l$$

Let's define the payout obligation on the firm access transmission rights attributable to constraint *l* as $PO_l = \lambda_l K_l^E$. Similarly, the settlement residues attributable to constraint *l* is $SR_l = \lambda_l K_l$.

Where the payout obligation for a particular binding constraint is greater than the settlement residues (where $PO_l > SR_l$), the AEMC proposes to scale the entitlements. Specifically, where $K_l^E > K_l$, the AEMC proposes to scale the entitlements by the factor:

$$k_l = \frac{K_l}{K_l^E}$$

So that the payout obligation on this binding constraint is equal to the settlement residues for the corresponding constraint:

$$PO_l = \lambda_l K_l'^E = \lambda k_l K_l^E = \lambda K_l = SR_l$$

Where the total payout obligation for a particular binding constraint is less than the settlement residues (where $PO_l < SR_l$) the AEMC proposes to pass the surplus back in the form of the secondary right. This right is chosen so that

$$\Phi_l \lambda_l \sum_i \alpha_{li} A_{li} = SR_l - PO_l$$

Or

$$\Phi_l = \frac{(SR_l - PO_l)}{\lambda_l \sum_i \alpha_{li} A_{li}} = \frac{(K_l - K_l^E)}{\sum_i \alpha_{li} A_{li}}$$