Regional Electricity Trading: Opportunities and Challenges for Ontario

Brendan Ring, Larry Ruff and Usman Hannan

27 April 2008

Disclaimer: The views and opinions expressed in this draft paper do not necessarily reflect the views of the Independent Electricity System Operator.

1 INTRODUCTION

The purpose of this paper is to discuss the role of regional trade– defined as trade between a buyer in one jurisdiction (or market, or control area) and a seller in another[1] – in electricity, the costs and benefits of expanding regional trade, some of the obstacles to such expansion and what might be done to reduce these obstacles. Although the discussion often refers to Ontario and its context within Canada and North America, most of the discussion here is general and applies to any large, modern electricity system.

In discussing obstacles to regional trade and how these might be reduced, the emphasis here is on market and operational procedures rather than physical interconnection capacity. This emphasis is motivated by the fact that the procedural obstacles to trading across the boundaries or "seams" between markets often prevent regional trade from reaching the level and providing the benefits that should be possible given the physical capacity of the interconnections. Modifications to the way markets operate and are interface as a means of reducing the unnecessary costs and risks of regional trading caused by such "seams issues" should be a relatively easy and low-cost way to increase the volume and benefits of regional trade, particularly compared to the high costs and long lead-times involved in expanding physical transmission capacity. Ideally, improved market processes will create improved conditions for efficient intertie expansion when this is appropriate.

A major barrier to communicating the nature of seams issues is the complex nature of power systems and electricity markets. In an effort to address this problem, a substantial appendix is included with this paper which provides a primer on the economics of electricity trade. This primer introduces the nature of electricity markets and the benefits

---

[1] In this paper, a single electricity "market" is an electrical area or "jurisdiction" that is under the control of a single system operator (e.g., the IESO, or Hydro Quebec, or the New York ISO) and is subject to a uniform set of operational and pricing rules – technically referred to as a "control area". A "region" here is a combination of several jurisdictions/markets/control areas that allow trade with one another, thereby creating a regional system or regional market. The terms "jurisdiction" and "market" are used largely interchangeably, because even jurisdictions that do not have a formal spot market (for example, Quebec) must have rules and procedures regarding information flows and pricing that look much like market rules.

and costs of energy trade. Many concepts are introduced within the Appendix which are used within the body of the paper. Readers not familiar with some of the concepts referred to in this paper should refer to the Appendix. Section 2 summarises features of trade between electricity markets with particular attention on the differences between the Ontario and New York electricity markets. Section 3 catalogues the different types of seams issues that exist in general. A range of options for addressing seams issues which might be considered over the next five to ten years are presented in Section 4. Our conclusions follow in Section 5.

2 THE PRACTICE OF REGIONAL ELECTRICITY TRADE

The power system of the northeast United States and the southeast of Canada can be viewed as a single large scale power system. However this system is divided into "control areas" or jurisdictions within which one organisation has control over how that part of the regional system operates.

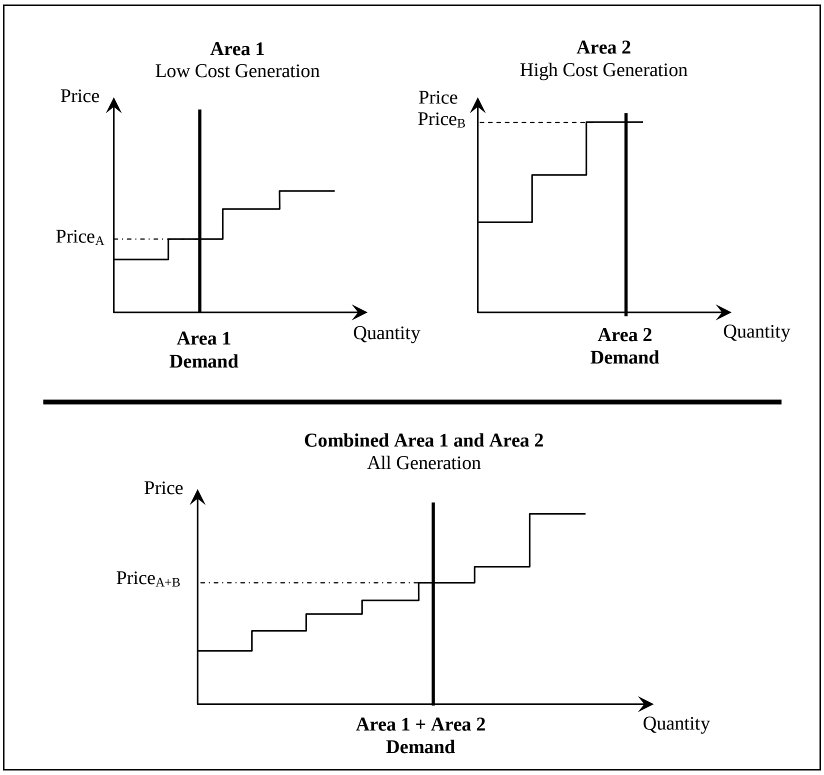

Figure 1: A stylised view of the jurisdictions around Ontario and New York.

Figure 1 shows control areas as circles containing their own transmission networks with heavier lines representing interties between jurisdictions. This system is electrically a single entity and could in principle be operated as a single entity without any concept of different control areas. In such an idealised system consider a supplier in the mid-west of the United States who has surplus low cost power and a consumer in New York City who has a need for more power and is willing to pay for it. The supplier could offer that energy into a market while the consumer in New York City could bid for that energy, and the market would match those trades and take care of all the transmission considerations automatically. Neither of the trading parties, or the operator of these idealised markets, would need to know anything about the other party.

In practice, the system operators in each jurisdiction aim (to the extent possible) to isolate their control area in real-time from unexpected power flows from outside by fixing the intertie flows for periods traditionally of one hour, though markets are moving to shorter periods. As long as the intertie flows are constant, a system operator can largely ignore the power system outside of its control area. Given this paradigm, let us consider again the trade from the mid-west to New York City.

The supplier might be in the market operated by the Mid West ISO (MISO) while the consumer is in the market operated by the New York ISO (NYISO). When power is injected into the network in the MISO market it does not flow directly to the customer in New Your City. Instead, loop flow comes into play.[2] Some of the MISO injection will flow to New York market to the south of Lake Erie and potentially via the market of the Pennsylvania-Jersey-Maryland Interconnection (PJM) while some will flow north of Lake Erie (via Michigan and across the Canadian border) to the Ontario market and onward (across the Canadian border again) to the New York market.

This transaction must be scheduled in four different markets, all of which have subtle differences in their scheduling processes and all of which requires the intertie flows to be fixed. It is not currently possible to have the transaction scheduled simultaneously in all those markets. Even once the transaction is scheduled, real-time events can mean that the scheduled transaction might be cancelled by one or more system operators, effecting the entire transaction. In completing this transaction some of the power is imported to Canada and then exported back to the United States, creating obligations in regard to international trade and taxes.

The efficiency of electricity trades between jurisdictions can only be as good as the quality of the intertie schedules. An efficient intertie schedule is just the intertie flows corresponding to the most efficient schedule for the entire interconnected system at any given time. Likewise, efficient pricing across interties must reflect the economics of the entire interconnected system at any given time. This means that the effects of loop flow, the complex “spring washer” effect that transmission constraints in loops have on prices, and dynamic changes in the power system must be correctly represented. This requirement means that flows and price differences across an intertie may be related to what is happening within a distant control area.

This ideal is not being achieved today. Only subsections of the region are represented in each market’s processes, while fixing intertie flows limits options for real-time trade to resolve localised supply and demand imbalances more efficiently. It is quite possible that the generators and the constraints which should be driving prices within a jurisdiction are not even represented. The general consequences of this are economically imperfect prices and real-time price spikes and an overall higher cost of energy.

---

2 Loop flow and the “spring washer” effect are described in the section A.2.2 in the Appendix.

2.1 THE ONTARIO MARKET AND REGIONAL TRADE

The Ontario electricity market considers all its market based transactions with other jurisdictions to be within its market scheduling process. The province of Ontario is treated as one region with a single energy price for the entire region, with the directly connected markets surrounding Ontario each represented as a separate region, or "intertie zone" with a single energy price.

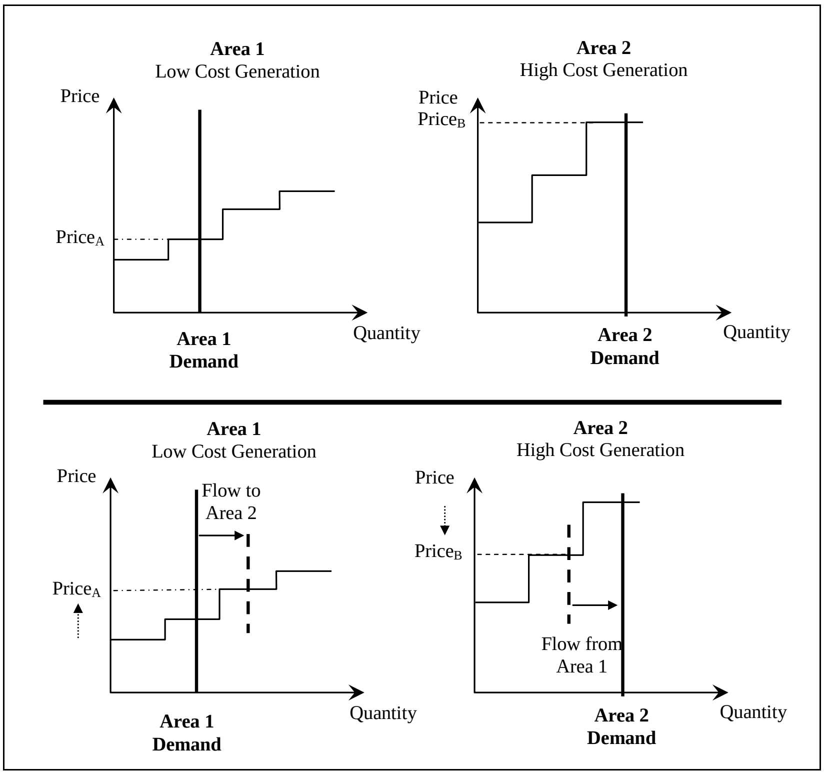

Figure 2: Network representation of the Ontario market.

The zones in the Ontario market are shown in Figure 2. Note that the connections with Quebec effectively allow generators in Quebec to disconnect from the Quebec power system and instead to connect to the Ontario power system, and hence these interconnections are not represented as interconnections with an external inter-zone within the Ontario market.

Energy scheduled across an interconnection is only scheduled in the Ontario market, which is then confirmed with the target jurisdiction to determine if matching transactions have been scheduled.

Each interconnection in Figure 2 has a capacity limit which can create price separation between Ontario and each of the intertie zones. An operational security limit is also imposed on the total rate of change of flow between Ontario and all inter-tie zones. This security limit can influence the price difference between the intertie zones.[3]

[3] As the security limit impacts the flows on multiple interties it might be though that this could create a linkage between the price differences between the different interties. However, the

Participants in the Ontario market can submit bids (export) or offers (import) to/from any of the zones: the single Ontario zone or the external intertie zones. A participant wishing to "wheel" power through Ontario can bid to buy that energy in one intertie zone while simultaneously offering to sell that energy in another intertie zone. There is no linkage between the offer and bid associated with a wheeling transaction, so it is possible for only one of the two legs to be accepted. Each external intertie zone is associated with a notional "boundary entity" rather than a physical generator or load. Only once a quantity of energy is scheduled at a boundary entity must a participant establish how that energy will be consumed or generated, and must schedule this accordingly with the system operator of the intertie zone.

An hour prior to the real-time trading hour the Ontario market determines a "pre-dispatch schedule" using its optimization algorithm, which determines prices and schedules for the trading hour based on the best available data at the time the pre-dispatch schedule is run.[4] The pre-dispatch schedule produces one set of schedule solutions that includes a full representation of the transmission limits within Ontario and the intertie zone flow limits. This "constrained" schedule provides an indication as to which generators will be operated in real-time. The pre-dispatch schedule also produces a set of schedule and price solutions which includes the intertie zone flow limits but ignores all transmission constraints within Ontario. This "unconstrained" solution provides forecasts of the single price in Ontario and each of the inter-tie zones. It also prices a number of classes of Operating Reserve. The pre-dispatch schedule gives generator outputs and intertie flow schedules for a full hour at a time, and indicates the energy price differences between Ontario and the inter-tie zones.

The Ontario market has a self commitment process where by participants decide themselves when to turn on and turn off generators (rather than this being centrally managed). For much of the time the market has operated this approach has left no certainty until real-time that a generator will be scheduled. Recently, the Ontario market has implemented a "cost recovery guarantee" for generators that commit their generators based on the pre-dispatch schedule. This gives generators protection against the possibility that generation that they commit based on pre-dispatch schedules will not be economic to run in real-time. However, this approach is more limited than a formal centralised unit commitment.

The pre-dispatch schedule was conceived as a planning tool. It provides insight into prices and how generators are likely to run. Generally, the pre-dispatch schedule has no impact on real-time schedules, with a separate model being used to determine those. However, as flows on the interconnections must be scheduled with the surrounding markets ahead of real-time, the pre-dispatch schedule for the interconnections is taken to be how the interconnections will be operated in real-time. The Ontario market operator informs each of the system operators in the inter-tie zones how much energy is to be

---

Ontario market has implemented pricing rules in such a manner that no such price linkages occur.

4 Pre-dispatch schedules are actually run for up to 24 hours ahead, but the hour prior to the trading hour has particular significance.

traded between the buying/selling party in Ontario and the buying/selling party associated with the boundary entity.

The Ontario market operates in real-time on a 5-minute basis, with the intertie flows being fixed at their levels for the associated hour. The system operator can advise generators to change their outputs in real-time based on information from its scheduling tools or to comply with power system operating procedures. This real-time scheduling process accounts for transmission constraints within Ontario. Immediately after the real-time interval energy and operating reserve prices are determined for the Ontario zonezone using the unconstrained model with the intertie flows fixed at the pre-dispatch level. Generators are paid based on these 5-minute prices while consumers are settled at a derived hourly price.[5]

The 5-minute real-time energy price for each intertie zone cannot be determined in the same way as the Ontario price is determined since the intertie flows are treated as constants in real-time. Hence the real-time energy price for each intertie zone is set equal to the 5-minute real-time Ontario price adjusted by the price differences between Ontario and the inter-tie zone as determined in the pre-dispatch schedule.

As an example, if a participant offersoffers to provide energy for $50/MWh in the New York intertie zone and the price in that intertie zone in the pre-dispatch schedule is $60/MWh for the hour then that offeroffer will be scheduled. If the Ontario price in the pre-dispatch schedule was $80/MWh for the hour then the price difference across the intertie is $20/MWh for the hour. In real-time, and for a given 5 minute interval, if the Ontario price is $65/MWh then the price in the intertie zone will be $20/MWh less, or $45/MWh. Clearly the participant scheduled at a price of $50/MWh faces a risk that the actual price it is paid will be less for at least some of the 5-minute intervals.

The Ontario market provides “intertie offer guarantee” to net imports (imports less exports) on interties to ensure that they do not suffer a loss over the hour for delivering according to their schedule because the real-time price they are paid (based on real-time pricing) causes them to make a loss (based on their offersoffers accepted in the pre-dispatch schedule). The intertie price guarantee’s remove incentives for importers to renege on trade which is uneconomic, and hence increases the security of the Ontario power system. However, these guaranteed payments are a product of a seams issue, and are a cost on the consumers of Ontario (who collectively fund these payments). No intertie price guarantee is provided by Ontario for exports – it is left to the external market receiving the energy to provide such guarantees.

Note that that the intertie offer guarantees are based on the Ontario calculated price for external markets. They are not based on the *actual* price in those markets. Hence those trading on the interties are exposed to the risks associated with the price difference between the Ontario calculated price for the intertie zone and the actual price for the intertie zone determined within the neighbouring market.

---

[5] Locational factors to account for the impact of transmission losses are not applied to prices in Ontario.

The Ontario market makes financial transmission rights available across the interconnections to the inter-tie zones. Prices do not vary by location within Ontario so that there is no need for financial transmission rights within the province.

The Ontario market allows for the import of operating reserve by it being offered at a boundary entity to serve Ontario's operating reserve requirements. Ontario does not facilitatefacilitate the import of operating reserve.

The Ontario market provides no payment for generator capacity. All payments are based on energy and operating reserve only. The market rules allow for such market if required. It is possiblepossible that the relatively low peak energy price pricein Ontario of $2000/MWh is inadequate to encourage optimal generation investment. While this might favour the activation of a capacity market, the Ontario government has demonstrated a preference for acquisition of new generating capacity under power procurement tenders.

The Ontario market has some significant limitations with respect to regional trade. The market includes no representation of any connection between the intertie zones outside of Ontario. Further, as transmission constraints within Ontario are not reflected in the Ontario price, such constraints cannot impact prices outside of Ontario. It follows that spring washer type pricing effects stemming from Lake Erie loop flows are not correctly priced in the Ontario market, except to the extent that traders can reflect them in their bids and offers (which is unlikely).

The disconnection between how intertie flows are scheduled and how they are priced can create inefficient schedules. While the schedules do not properly reflect constraints outside of Ontario they do at least reflect constraints within Ontario. The unconstrained schedules used to generate prices not only ignore constraints within Ontario but use as their starting conditions unconstrained schedules from the previous trading period, and hence the prices the algorithm produces reflect nothing of the constraints within Ontario and only have a limited reflection of the actual power system state at the time.[6] The fact that the real-time inter-tie zone prices are based on a combination of real-time Ontario price and pre-dispatch intertie zone price difference creates a further inaccuracy.

Wheeling transactions through Ontario are further limited by the fact that there is no mechanism for ensuring that the both the import and export component are simultaneously scheduled. Participants can bid at extreme negative prices and offer at extreme positive prices to maximise the chance of being scheduled, but this forces them to commit to the transaction without knowledge of the price. A participant who is only interested in wheeling if prices are within a particular range is unlikely to want to use the existing mechanism. Participants effectively need the capability to define the price they wish to pay for the export from Ontario, and the maximum difference between the import

[6] This was a deliberate design strategy at the commencement of the market to ensure that compensation paid to those in Ontario who are scheduled inconsistently with the unconstrained prices are compensated for being "constrained on" or "constrained off" on the correct basis. To align the starting conditions of each constrained and unconstrained schedule would have undermined the constraint compensation mechanism.

price and the export price that the participant is prepared to accept, with the transaction being scheduled if and only if these conditions are satisfied.[7]

2.2 THE NEW YORK MARKET AND REGIONAL TRADE

The New York electricity market is connected directly to four jurisdictions: Ontario, Quebec, the ISO New England market and the Pennsylvania-Jersey-Maryland Interconnection (PJM).

Unlike Ontario, New York operates what is called a "nodal" energy market. That is, rather than having one price for the entire market, prices vary at every location in the network to reflect limitationslimitations on power flow due to transmission line physical limitations and operational constraints imposed to ensure power system reliability and security. This means that New York can use a single system representation to both schedule and price energy, largely avoiding the mis-match between prices and schedules encountered by Ontario. Like Ontario, New York schedules and prices operating reserve simultaneously with energy. In addition New York prices regulation, an ancillary service which absorbs small fluctuations between supply and demand moment by moment.

New York has a day ahead forward market with centralised unit commitment. This is the main mechanism by which transactions outside of New York can be scheduled.

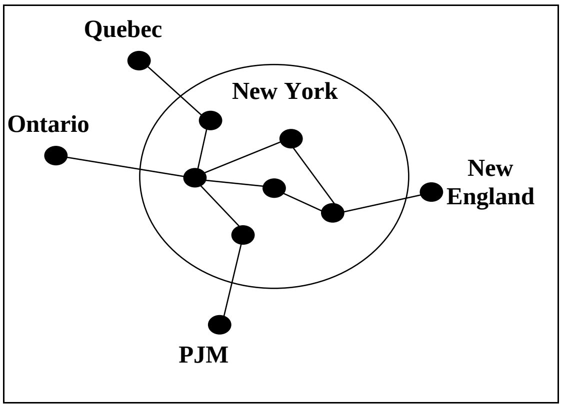

Figure 3: Network representation of the New York market.

Figure 33 shows the representation of the New York market used in the day-ahead market process. While all the transmission lines in New York are represented, only a single link is represented to each "proxy bus" outside of New York. Participants in the New York market can buy or sell energy at these proxy buses, and if scheduled the intertie flows are effectively scheduled in the New York market. Exporting participants must schedule their exports with the markets the power is flowing to. Like Ontario this representation suffers from the lack of representation of constraints outside of New York, however it differs significantly from Ontario in that prices can vary within the New York market to reflect constraints within the market.

[7] This problem is more difficult to solve than the current scheduling and pricing problem, which is the primary reason it was not adopted at market commencement.

It is possible to schedule imports and exports after the day-ahead market has closed. Indeed, this is how Ontario's exports to New York are scheduled. The New York system operator must be notified of the transaction and will enter it into a near real-time unit commitment process. If an intertie flow scheduled in either the day-ahead market or submitted subsequently cannot be accommodated then it will be scheduled off and the transaction rejected. However, day-ahead transactions are given higher priority in this process. The New York market requires that intertie schedules remained fixed for periods of 15 minutes.

While this process looks similar to that in Ontario, there are a number of fundamental differences. The Ontario process occurs one hour prior to real-time whereas most intertie flows scheduled in the New York market are scheduled a day prior to the trading day. As this New York process happens much earlier than the Ontario process, participants in the Ontario market that participate in the New York day-ahead market face significant uncertainty as to the cost of getting the energy to or from the Ontario border. Further, since Ontario schedules its flows to New York much closer to real-time, there is a greater risk that those schedules cannot be accommodated by New York.

Further, since the New York process includes a centralised unit commitment and a forward market, participants who are scheduled within New York not only have reasonable certainty that they are scheduled, but they also have an assurance of recovering the associated costs. This is because by settling based on the forward market participants scheduled by that mechanism are indifferent as to what subsequently happens in real-time. If these day-ahead transactions cannot be accommodated in real-time, the real-time market processes will provide compensation for this. Participants within Ontario do not have this same certainty, accept to the extent provided by the "cost recovery guarantee".

The technology used to solve unit commit problems in New York is well suited to the problem of scheduling wheeling transactions so as to ensure that both the importing and exporting leg of the transaction are scheduled simultaneously if the conditions of the bid are satisfied.

New York provides capacity payments for installed capacity. These payments are implemented on a zonal basis, with some zones in New York, like New York City having high payments for installed capacity, while other zones with ample capacity can have a payment of zero.

3 A CATALOGUE OF SEAMS ISSUES

The purpose of this section is to attempt a relatively comprehensive list of factors that create seams issues which limit the extent or efficiency of regional trade.

3.1 INEFFICIENT SCHEDULING AT MARKET BOUNDARIES

The scheduling of intra-jurisdiction trades has great scope for inefficiency. Suppose a trader has purchased energy in Ontario under a bilateral contract and wants to deliver that energy to a customer in New York. The energy must be scheduled in both markets. It is possible that the trader may succeed in having the energy scheduled in one market but not the other, which leaves the trader exposed to trading out of that position in the real-time market. Further trades may be curtailed in real-time due to changed circumstances.

Wheeling from one market, across a second market, to a third market is possible but the approach to this differs by market. In particular, Ontario's current approach does not link the import and export transaction meaning that only one part of the schedule may go ahead. Participants can largely eliminate this risk by pricing exports and imports at extreme bid and offer prices so they are likely to be scheduled at any price. In effect they become price takers. New York explicitly links the import and export legs of wheeling trade so that if one leg is scheduled then so is the other. It follows that very price sensitive wheeling transactions are unlikely to occur across Ontario but are more likely to occur across New York. Consequently transactions across both Ontario and New York are less likely to occur. In general the situation becomes even more complicated and fraught across multiple markets because of the totally independent pricing and scheduling processes.

It is worth nothing that the PJM and MISO markets treat imports and exports as fixed price takers. In effect, these are bilateral transactions outside of the market. While this facilitates easy trade, the efficiency of the bilateral trades will only be as good as the mechanism by which they are formed allows. A centralised market clearing process provides the greatest opportunity for matching buyers and sellers while accounting for transmission constraints.[8] There are generally still risks for participants trading under bilateral contracts as they are generally still exposed to locational price differences between the injecting point and the receiving point. Finally, the key features of the PJM and MISO approach can be achieved in the Ontario market just by using extreme bid and offer prices for the import and export legs.

In conclusion, with each market scheduling trades independently of the broader power system, and with risk exposure for traders with strong price sensitivities, there is little likelihood that the most efficient trades are occurring.

3.2 INACCURATE PRICING AT MARKET BOUNDARIES

Pricing within a market is only as accurate as the prices reflect the actual schedules in that market. New York's nodal prices are probably a quite accurate representation of the economics of its dispatch. Ontario meanwhile employs an unconstrained pricing representation which does not represent the true schedule. This means that the Ontario market does not accurately transform the price from an importing intertie zone to reflect the correct import of Ontario constraints in the price at an exporting intertie zone. The Ontario Energy Board (2006) has observed that the hourly unconstrained price in Ontario can be sufficiently low as to encourage exports to New York when the actual cost in Ontario of producing that energy is uneconomically high. While Ontario pays compensation for this, the resulting schedule is not actually efficient. Nodal pricing would avoid this issue.

---

8 New York's day ahead forward market can be seen as a mechanism for forming multi-lateral trades which is efficient, with those trades generally being price takers in real-time.

The actual prices employed at settlement in Ontario are a combination of components determined at different times. Further, while Ontario employs limits on the rate of change of total import and export to Ontario, these are not reflected in intertie zone prices, when it could be argued that they should be. As noted earlier, the approach to pricing in Ontario is perfectly adequate in an isolated market but clearly have limitations for regional trade.

The approach of representing external control areas as being connected to markets by radial lines, with no representation of the broader external network, produces prices which fail to represent the economics of loop flow across multiple markets. In the case of Ontario the Lake Erie loop flow should have a material impact on prices in Ontario if constraints occur within external control areas, but this is not represented. Further since no constraints within Ontario are represented the prices at the boundaries of the Ontario market will be inefficient prices with respect to the broader network. The materiality of this effect depends on how much impact the schedules in one control area have on constrained flows in other control areas.

Even if a trade is scheduled across and between Ontario and New York, there are discontinuities in the prices between the Ontario inter-tie zone for New York and the actual price in the New York market for the connection points with Ontario. While each market has mechanisms for compensating traders for transactions wholly within their markets, traders bear the cost of the differences between the markets.

The practice of fixing intertie flows for an hour reduces the efficiency of the dispatch in real-time and can trigger price spikes. Patton (2003) observed exactly this behaviour between New York and New England. Patton argued that flows were often running in the opposite direction to the price differences, price differences were arising when no constraints existed[9], and that flows were not reacting to prices over-time. The causal relationship between price differences and flow were not strongly apparent. Patton ascribed these outcomes to the need to fix scheduled ahead of time and the scheduling and pricing risk exposure that participants face at the boundaries of markets.

Participants who have strong price sensitivities, and who only want to trade within a narrow range of prices, are unlikely to participate due to the uncertainties of being scheduled and the uncertainty in the prices faced.

3.3 DIFFERENT APPROACHES TO FORWARD SCHEDULING AND FORWARD MARKETS

Ontario does not operate a day-ahead forward market, while New York does. A key feature of forward market – as opposed to pre-dispatch schedule – is that transactions in a forward market are settled at the forward market price. That is, prices and schedules are determined simultaneously and therefore are consistent with each other and with the costs participants incur. Participants scheduled in a forward market are only exposed to real-

[9] Coincidentally, were broader North American loop flow effects properly represented, this type of behavior should actually be expected occasionally. While the intertie between two markets may not be at a flow limit, the flow is effectively constrained by a constraint on another path.

time markets to the extent that their schedules are changed in real-time. New York's forward market therefore provides participants with great certainty with respect to their scheduling and with respect to their earnings on the day prior to the trading day. Participants in Ontario must commit to courses of action prior to real-time without being certain what the prices will be or how they will be scheduled (though can revise their costs on the trading day via their bids and offers). Both approaches are valid and work perfectly well in isolated markets, but in the context of interconnected markets are likely to undermine the efficiency of cross-border trade.

Ontario has partly mitigated this discrepancy via offering cost recovery guarantees to major generators based on day-ahead pre-dispatch schedules. However this is a very restricted and limited mechanism when compared with a day-ahead forward market. It is understoodunderstood that Ontario is also considering reducing its pre-dispatch scheduling time frame from 1 hour to 15 minutes so as to align the periods for which intertie flows must be fixed with that for New York.

The trade related issues with the Ontario market primarily reflect a lack of a forward market which settles transactions ahead of real time, reducing participant exposure to subsequent changes,

3.4 Barriers to Ancillary Service Trade

The markets in Ontario and New York each schedule and price energy and operating reserve together. Each market must procureprocure enough operating reserve to satisfy the requirements set b the system operator based on criteria defined by industry regulatory bodies such as the North American Electric Reliability Council (NERC) and the Northeast Power Coordinating Council (NPCC). To meet the reliability standards, a system operator must ensureensure that operating reserve is surplus to its requirements before it is offered into another market.

NPCC requirements only allow 50 MW of ten-minute response operating reserve for a control area to be provided automatically from outside.[10] This has allowed Ontario to reduce its local operating reserve requirements by about 4%. This relatively small allowance highlights a practical difficulty with the trade of operating reserve. If too much operating reserve is imported then a control area may be short due to a transmission failure or even a curtailment of an intertie flow. This creates a very significant risk for a system operator, potentially requiring that it carry operating reserve simply to cover the loss of access to external operating reserve. This could cause the cost of operating reserve in the importing market to increase significantly. This effect already exists within jurisdictions – operating reserve providers may fail to respond or be unavailable for transmission related reasons – and is factored into current scheduling processes, but the risks are significantly magnified if a significant proportion of operating reserve is coming from outside the jurisdiction. While not ultimately a reason to prevent market based trade of operating reserve, concerns of this nature may restrict the extent to which operating reserve trade is allowed, potentially limiting its value.

[10] Sourced from OEB (2006)

An import of operating reserve would effectively require that the intertie(s) have sufficient unutilised capacity in the import direction (including switching from exporting to importing quickly) to ensure that if the external operating reserve is activated that it can be delivered. This also requires that intertie flows can change in real-time to support the delivery of energy from the operating reserve source. Both of these considerations will complicate the scheduling of intertie flows and their management.

The value of traded operating reserve is a function of the inter-market energy price difference which, because of seams issues, may not always reflect the true economics of trade. As the value of operating reserve tends to be lower than the value of energy, reflecting the difference in a generators cost and the prevailing price, operating reserve trades are likely to be more sensitive to seams issue related price inaccuracies than energy market trades. It is argued above that market seams are not favourable to very price sensitive trade of energy, so unless this can be corrected is unlikely that a robust trade in operating reserve can be achieved.

3.5 COST BARRIERS DUE TO TRANSMISSION TARIFFS

A seams issue not discussed in our earlier examples concerns transmission tariffs. Each jurisdiction has some mechanism for charging users of its transmission grid toward the recovery of the costs of that transmission grid. A component of these costs is recovered from those who wheel energy across the transmission grid. This means that a trader that causes power to flow across 5 jurisdictions will be charged transmission usage by each jurisdiction. The concern is that this “rate pancaking” is a potential barrier to trade.

Transmission network cost recover is contentious at the best of times. The fundamental problem with transmission cost recovery is that the decision to build the assets has already occurred in the past so the transmission cost recovery is largely about recovering a sunk cost. If this cost is charged to *any* user of the transmission grid based on that users energy usage then this cost will distort energy usage. If the network user reduces its consumption because of the higher delivered cost of energy, this does nothing to reduce the network sunk costs to be recovered. The implication is that transmission network costs should not be recovered based on instantaneous electricity flows. In practice, however, all transmission use of system (TUOS) charging regimes around the world have at least some component of the charge associated with electricity usage.

If it is accepted therefore that some degree of distortion of electricity trade is unavoidable from current TUOS practices then the issue becomes one of how to minimise any additional distortions in regional trade. Rate pancaking of itself is not an issue if the overall cost is equitable relative to how other grid users with similar levels of usage are charged.

To avoid distortion it is desirable that transmission network operators and their TUOS charging mechanisms are ring-fenced from any energy trading activities. However, in some jurisdictions where electricity reform is less advanced the electricity industry is still essentially vertically integrated. This means that the potential exists for high transmission charges to be levied which do not impact the bottom line of the local vertically integrated utility but which does generate income from inter-jurisdictional trade. Such an outcome would be anti-competitive in the context of the broader market, and would be detrimental to trade.

Another source of distortion can arise from different jurisdictions having different TUOS mechanisms. This could result in trades across one jurisdiction carrying a higher proportion of the total TUOS cost of that jurisdiction than trades across another jurisdiction. As these TUOS costs raise the cost of delivered energy then they distort the relative competitive of suppliers in different locations.

3.6 DIFFERING PRODUCTS AND PRICING METHODS

Products traded in electricity markets can differ between markets in the form, definition and obligations associated with being paid to provide that product. Such differences create barriers to trade and competition when trading across market boundaries.

The Ontario electricity market does not make capacity payments whereas the New York market does. If a generator in Ontario is competing with a generator in New York to provide a customer in Michigan then they are not competing on quite the same basis. The generator in Ontario needs a higher average energy price to break even than an equivalent generating in New York that receives a capacity payment. In some circumstances, these differences could create an advantage for one of the two generators.

Environmental considerations may eventually create similar issues in regard to accounting for emissions and renewable energy. If these are accounted for differently in different places then this will restrict the ability to trade these products between markets.

3.7 DECENTRALISED MANAGEMENT OF TRANSMISSION RELIABILITY

If a transmission line fails, then power flows will instantly redistribute themselves across the available pathways. To avoid such sudden changes violating the operating limits of transmission lines, power systems are operated so as to ensure that the power flows are feasible even if a number of different transmission lines become unavailable.[11] In effect, if one of the lines of concern were to fail then there would be enough capacity available in the system to accommodate the changed flows.

The limits on flows can include limits on interconnector flow which capture key features of potential contingencies outside of one market. However, limits imposed on

---

[11] The blackout which struck the north eastern United States and Canada in 2006 provides an example of what happens when this process fails. Some see interconnection between systems as being a cause of this cascading black out. At a basic logical level this is true – it could not have cascaded if there were no interconnections. However interconnections generally improve the reliability of power systems. For instance, a drought stricken region with significant hydro generation can avoid energy shortages by importing power from interconnected regions. Overall, it would seem that improved interconnection and trade should actually help to reduce the frequency of supply shortfalls and blackout events within each region, albeit with the risk of more widespread blackouts when systems and procedures fail.

interconnections do not necessarily reflect the same limits that would be imposed were all markets operated as one – since a greater scope could be allowed for inter-jurisdictionaljurisdictional changes in power flow and responses from other jurisdictions. In effect, the control area boundaries create artificial constraints on trade.

3.8 DECENTRALISED TRANSMISSION PLANNING

The development of the transmission system within a market needs to be centrally coordinated to ensure the correct trade-offs are made between transmission and generation investment, and to ensure that adequate transmission investment is made to support efficient operation of energy markets. The favoured model for transmission planning and investment involves a Regional Transmission Operator (RTO) taking responsibility for operating the transmission network and specifying what transmission augmentations are needed, which private industry actually builds those assets and receive a regulated return. This allows private enterprise to own assets while maintaining centralised and independent control over the planning and operation of the network.

The absence for coordinated regional and, in the case of Ontario and New York, international, transmission investment will mean that transmission investment will be piece-meal and sub-optimal. This will distort generation and demand side investment decisions and will ultimately raiseraise costs.

3.9 LIMITATIONS OF RISK MITIGATION MEASURES

Currently it is not possible to buy a single financial transmission right across multiple markets. Instead, financial transmission rights have to be purchased in different markets (potentially based on different processes and different timelines).

3.10 CURRENCY, TAX, AND REGULATORY DIFFERENCES

Energy market structural, design and implementation related seams issues are further amplified by regional trade occurring in different currencies, being subject to divergent tax regimes, and the application of different regulatory and compliance requirements in different markets.

3.11 COMPLEXITY AND LACKLACK OF TRANSPARENCY

Seams have created a situation where market operators are tuning their rules and processes to better address seams issues. Typically, this adds complexity to the design of their markets which would not exist in a single market. Further if different markets uses different methods to resolve the same seams issues then this complexity compounds. This added complexity can make it more difficult for participants to understand the market and to interpret the outputs of the market.

In the case of Ontario it can be argued that the policy decision made at the outset of the market that a uniform price would apply within Ontario, and the lack of a forward market, have given rise to prices which are not transparent and a host of complicated scheduling processes and ad hoc settlement features. Given the position of Ontario on a significant route of power flow around Lake Erie this lack or transparency will be feeding into the scheduling and pricing processes across the north east of the United States and the south east of Canada.

While transparency issues are not unique to Ontario, and a number of states and provinces surrounding Ontario are less reformed and consequently less transparent, moving the Ontario market towards the designs of New York and PJM would improve transparency.

4 APPROACHES FOR ADDRESSING SEAMS ISSUES

This section explores a number of options for addressing seams issues. These options are presented as a continuum from radical and comprehensive solutionssolutions to more incremental changes.[12] The more comprehensive solutions will certainly provide greater benefit than incremental change, although at a greater cost. While it is unclear which approach provides the greatest improvement on a unit cost basis in the short term, the more comprehensive solutions are likelylikely to be more attractive in the long term as, if planned for, they can be factored into the normal cost of market infrastructure upgrades.

It should be stressed that the system and market operators in North America are working actively to reduce the impact of seams issues. However, they are subject to many factors beyond their individual and collective control which limit what they can do. In addition to basic physical system limitations they are limited in what they can by: what their market participants are prepared to support and fund; the extent of broader economic and trade barriers; and the willingness of governments and government agencies to make the broader economic and regulator changes required for addressing seams issues. The agenda’s of individual players in the electricity industry and those of governments are not always aligned with achieving greater societal benefits through remedying seams issues.

4.1 EXPANDING THE SCOPE OF MARKETS

In North America the various system operators are working together to find means of over-coming the limitations created by seams. To the extent that seams issues are gradually addressed over time, the ideal final result should be no different to a single market. To the extent that barriers are not broken down, seams issues will never fully be addressed. An obvious question then, is why not just merge the existing markets? Why not have a single electricity market incorporating the eastern United States and Canada?

A common reaction to the idea of merging markets is that it is socially, practically, commercial, politically and/or constitutionally impossible. Concerns which are cited include: loss of national control over energy; impacts on electricity prices and competition; complexities of currency and tax issues; the large numbers of state/provincial and federal bodies that would need to cooperate; and so forth. While these complexities are real, resolving them is by no means impossible.

---

[12] Pierce et al (2006) provide more detailed and specific recommendations on addressing seams issues in the Canadian context.

On 1 November 2007 an electricity market called the "All Island Market" commenced in Ireland, spanning both the Republic of Ireland and Northern Ireland. This is a market that crosses a national border and incorporates one country and partpart of another, while facilitating a single market operator, two system operators, two different tax regimes, two different regulatory regimes, and two different units of currency. While Ireland still has some distance to travel in reforming its electricity sector, it provides evidence that there is nothing intrinsically impossible about applying a single market across large regions of Canada and the United States.

Experience in Australia also provides a good example. The states within Australia can be thought of in a similar light as a state of the United States or a province of Canada. During the mid-1990's electricity markets had been developed separately and independently in each of New South Wales and Victoria. Given these states were electrically connected, and that Victoria was also connected to South Australia and Queensland was soon to connect to New South Wales it became natural to think of forming a single national market. The federal government took the lead in promoting a national market. While the federal government could give some force to encouraging change through its power to regulate inter-state commerce, there was implicit agreement between the state governments that such a move was in their interest. Indeed, since the market commenced Tasmania has interconnected with the mainland and has joined the market. The national electricity market is now the largest control area in the world, stretching from Queensland in the north of the continent, across New South Wales, to South Australia and Victoria in the south, with Tasmania joined by an interconnection. This market has resulted in significant efficiencyefficiency improvements in the sector.

To achieve such a merger their needs to be political consensus from the various governments involved. Ireland has shown that this is possible between two sovereign countries while Australia has demonstrated this in the context of a federation of states.

How would a merged market impact those currently operating the regional markets? The whole point of expanded markets would be to increase trade. As market operators earn their revenue through trade there are commercial benefits in existing market operators joining a larger market. Even if a new market operator emerges for this expanded market, the existing market operators could remain as stakeholders and could continue to earn revenues based on trade through their current geographic areas.

It seems unlikely that a single system operator would be an acceptable solution for an expanded market. National and state governments would have some concern about this and there are likely to be good engineering reasons to maintain regional control.[13] However, maximum efficiency gains could be realised through having a high level system operator with responsibility for managing constraints on power system operation that are best managed at the overall power system level, such as inter-jurisdictionaljurisdictional flows, while jurisdicationaljurisdicational system operators continue to control those matters specific to their areas. An appropriate partitioning of roles between these two

[13] In particular, the function of power system network control equipment will be geared to a specific control area's needs and it may not be possible to redefine a control area without significant investment in new network and control equipment.

levels of system operation has the potential to achieve efficiency gains across the entire interconnected network.

4.2 A PARALLEL INTEGRATED MARKET

Rather than immediately merging existing market, an interiminterim approach would be to establish a market that operates across the existing North American markets but which runs in parallel to the current markets.

This "Parallel Integrated Market" (PIM) would determine prices and schedules across the interconnected markets based on the actual bid and offer data submitted by participants in those markets. System operators would need to provide information about the state of their networks to the operator of the PIM. For example, when a participant bids to buy or sell energy in the Ontario market, those same bids will be entered into the PIM. Participants in each jurisidication jurisidicationwill still be settled based on their local market prices and schedule, but will also have available the schedules and prices that would have occurred had the participant been trading in the Parallel Market.

Unlike a simulation, PIM is an actual market clearing process based on real bids and offers that are binding in their respective geographic markets. This means that the PIM can provide accurate information about how an enlarged market would perform. The PIM could provide information on how much efficiency is lost through seams issues by providing an efficient reference point. By studying the behaviour of the PIM solutions, participants involved in inter-jurisdictionaljurisdictional trade would have a better understanding of what the most efficient power flow solutions are, and could use that information to better structure their bids so as to capture those efficient gains. The attractiveness of the PIM model is that it provides a mechanism for facilitating improved efficiency of trade via market integration without actually requiring market integration.

Taken a step further, the PIM could be used as a pre-scheduling instrument for regional trade. If the PIM includes a day-ahead pre-scheduling process then that process will effectively determine efficient transmission flows and price differences between markets. Individual markets, if they wish, could adopt the PIM solutions as an official pre-dispatch schedule, including the intertie flow solutions. This would allow factors such as loop flow and the spring washer effect to be properly accounted. While it would be necessary for each individual jurisdiction to manage deviations from these pre-schedules in real-time, the PIM's more efficient starting point should aid efficiency of trade.

A given jurisdiction could even use the PIM as its real-time scheduling tool and take its results as its official market schedule. This would be particularly useful for states and provinces that are yet to commence markets as the cost of doing so could be significantly reduced. Further, it would allow existing market operators to stay with their current designs if they wished, or transition to the PIM design at any time of their choosing.

The PIM could also be used as a means of defining financial transmission rights across market boundaries, though the details of how this might work have not been explored,

It would not be a trivial exercise to establish the PIM. The IT system requirements would be of a similar scope as those in any existing North American market although the data volumes would be at least one or two orders of magnitude greater than any existing market. The PIM design would have to include features to translate different bid and offer structures within each jurisdiction into a common format. The cost of establishing the PIM would need to be shared by those jurisidctionsjurisidctions across which it operates, so it is unlikely that it could be established without widespread support for the concept.

A first step to exploring a PIM design might be to select a number of representative historic days from a range of markets and to simulate them individually and together as a single market using the most compatible modelling tools available (probably a commercial market simulation product). This would provide incite into the potential benefits to be gained from a PIM.

4.3 INTRA-HOUR TRANSACTIONS (OR VIRTUAL REGIONAL DISPATCH)

Patton (2003) identified a lack of casual relationship between schedules and prices on interties between New York and New England. Subsequently, in his market monitoring reports for New York, Patton proposed an approach for improving the price responsiveness of inter-tie flows within the hour. This was initially called Virtual Regional Dispatch but is also known as Intra-Hour Transactions.

The concept is that the system operators would monitor the price differences across interties and would adjust the flow rates within the hour to better align flows with the price differences. This would allow a lower cost market to provide more energy to a higher cost market, and hence increase the efficiency of flows. In effect the system operators would be matching buyers and sellers in the two markets, rather than having participants explicitly trade across the interties.

This approach has the potential to improve dispatch efficiency, but is a relatively limited and imperfect approach. It still takes no account of the broader flow dynamics between markets which could legitimately imply that the flow should be counter to the price differences. As the approach tends to rely incremental adjustments revised iteratively over time it is unlikely to converge to an ideal schedule, especially if numerous control areas were doing this simultaneously. A PIM based approach would be superior though necessarily more complicated.

An approach like this would probably not work at all for Ontario given its current pricing system. The unconstrained prices would fail to reflect intra-Ontario constraint costs so it would be very difficult to define the correct price difference.

4.4 STANDARD MARKET DESIGN

The differences in features in adjacent markets are a significant factor in seams issues. During the 1990's different North American markets evolved along different paths. These differences came about because of local policy differences, differences in the structure and nature of the local power industry, and different ideas as to how to address problems encountered during the design and implementation phases. Consequently there are no two markets which are the same. These differences have already been highlighted in the earlier comparison of the Ontario and New York markets.

The Federal Electricity Regulatory Commission (FERC) in the United States has promoted the concept of a "standard market design" to address these differences. The standard market design requires that markets must support a host of standard features and accommodate a range of trading mechanisms, such as bilateral energy trade, forward markets, spot trade etc. The standard market design is not a fixed design but is essentially a functional requirement. Markets still have some freedom as to how that functional requirement is satisfied in practice.

Major markets like New York and PJM have markets which are broadly standardised. Efforts have been made to align Ontario's market more with the design of its neighbours but with limited success. Standard market design is not a costless exercise. Significant and costly changes to IT systems, including market scheduling algorithms are required. The current design of the Ontario market is sufficiently different from New York and PJM's design, most particularly because of uniform pricing, that only piece-meal approximations can be made without high costs. A day-ahead market in Ontario would help considerably but again with considerable costs. The establishment of a day-ahead market in Ontario have been confounded by what can be best described as industry apathy with respect to funding it. Not withstanding the willingness of the Ontario market participants to support change, a move by the Ontario market to a forward market and nodal pricing would be a significant factor in reducing seams issues.

The basic concept of the standard market design is undoubtedly one of the key ingredients in improved trade efficiency. However boundaries will still exist between such standard markets. Thus while seams issues related to different market design concepts may be mitigated in part, many seams issues would still remain to be resolved as a manner of implementation detail.

4.5 REMOVING COST OF TRADE AND REGULATORY BARRIERS

Various jurisdictions in the United States are moving to simplify their transmission tariff structures. For example, PJM and MISO have agreed that their local customers only will pay for the cost of existing transmission assets, with importers from the other jurisdiction facing no additional cost. However, it is still proposed that where new transmission assets are built which can benefit importers then importers should share some of those costs.

Standardised transmission charging regimes across the interconnected markets would help to minimise the barriers created by these charges. Such standardisation should include ring fencing of the transmission businesses of vertically integrated utilities from energy trading businesses. A potential TUOS approach based on that employed by PJM and MISO would be to charge all existing asset costs to users within the jurisdiction but to share the cost of future augmentations between jurisdictions and with inter-jurisdictional traders.

Moves towards pan-market transmission investment planning would ensure that transmission capacity is better aligned with the needs of the traders. Some form of international RTO could be established to provide a broader view of how the inter-connected transmission network should be expanded. Such an international RTO could help in defining how TUOS costs for future transmission augmentations are allocated between jurisdictions and traders. Note that this would just be a planning body; it would not actually own or operate transmission lines.

Short of currency and tax union, the only thing that can really be done about currency and tax issues is to design features into markets that recognise these risks and provide some means for mitigating them. For instance, setting standard reference currencies and currency conversion times might at least reduce the complexity of handling these risks.

Regulatory compliance practices and processes could be made more standardised. For example, regulatory bodies in the different provinces and states, as well as federal bodies could work jointly towards standardisation along the lines of that envisaged in standard market design.

4.6 IMPROVED INFORMATION, PLANNING AND COORDINATION: A CASE STUDY IN OPERATING RESERVE TRADE

If markets stay much as they are today, with no further move to market standardisation or integration, what more could be done to limit the impacts of seams issues? The answer is to improve the information available to system and market operators, and hence to the market, so as to better facilitate trade.

Consider the situation with operating reserve. Conceptually, markets can trade operating reserve across interties quite easily.

To facilitate trade in operating reserve a system operator needs to have confidence that:[14]

- Adequate operating reserve will be available to its market after allowing for exports and imports of operating reserve.
- Adequate unutilised intertie capacity in the importing direction (including reducing flows in the exporting direction) exists on the interties required to deliver energy from operating reserve providers when activated.
- Adequate operating reserve will be available to its market in the event of an intertie failure. The market might for example import on multiple interties but only plan for the loss of one.
- Changes in real-time power flow between markets owing to scheduled intertie flows being cancelled will not restrict the availability of operating reserve.
- The source of the operating reserve is actually capable of delivering that operating reserve (i.e. it is not scheduled in real-time differently so as to remove the capacity needed to provide the operating reserve).

---

[14] The system operator, market participants, and regulatory and government bodies must also have acceptance that once local operating reserves are committed to a particular external jurisdiction, that jurisdiction must be given first call on that resource even if it means experiencing an operating reserve shortage in the local market. Of course, to the extent that the external jurisdiction has surplus operating reserve, it would be desirable for there to exist a market based mechanism for the operating reserve to be purchased back.

All the system operators in the interconnected regionregion trading operating reserve have to be simultaneously confident in respect of all these issues.

It would seem that the best way to facilitate greater operating reserve trade would be to schedule and fix operating reserve inter-market trades at the time of fixing the intertie schedules, and to put in place more sophisticated protocols for curtailing energy transactions in real-time so as to preserve adequate operating reserve in each jurisdiction.

Once operating reserve is scheduled, the source providing that operating reserve would have limits placed on changes in its schedule so as to preserve its operating reserve availability to other markets.

Alternatively, the actual providers of operating reserve might just sell to their local markets, with the overall market surplus and shortfall operating reserve being traded between markets without regard for who the actual supplier is. This might allow for real-time trade of operating reserve (within the available transmission constraints) between markets so that the system operators can better match their requirements, but without requiring the ultimate providers of the service to know or care about this trade.

To facilitate such arrangements pre-dispatch planning arrangements, data exchanges, procedures, and software aiding scheduling decisions would all need to be adjusted.

More generally, other *ad hoc* solutions could be developed by system and market operators on a bilateral or multilateral basis to better improve their knowledge of each others situation, to develop tools that allow them to resolve the trade-offs between their different interests, and to ultimately increase their knowledge as to how much additional opportunity for trade can be “squeezed” from the existing structural paradigm. This would require significantly higher data exchanges between system operators and much higher levels of co-operation. A limiting factor in what can be achieved is that system operators are very unlikely to accept situations that could create uncertainty as to which system operator has ultimate responsibility and power for addressing a particular situation.

Whether it is worth moving forward with *ad hoc* solutions ultimately depends on the potential market benefits relative to the cost and complexity of implementing mechanisms to support such trade.

5 CONCLUSIONS

The trade of electricity and related services between jurisdictionsjurisdictions has the potential to lower the average cost of energy across North America, through improved efficiency in the short run by using the lowest cost mix of resources, and through more efficient investment in the long run. However increased trade will create winners and losers within subsections of the market. The objective of policy should be to expand trade (and other beneficial economic activities) broadly, so that most individuals win more than they lose from the resulting aggregate economic growth, while maintaining social safety nets for the minority that does less well.

While greater integration between markets has occurred over the last two decades, limitations of trade are created by the "seams" between these markets. Seams issues can broadly be viewed as arising from: differences in market designs; differences in market timeframes; differences in the environment in which the market sits; and the disjoint created by the need to trade across two distinct markets. Seams manifest themselves through energy scheduling and trade, trade in ancillary services, transmission operation and planning, economic regulation, and even through the complexity of mechanisms to work around seams issues.

Seams issues would be best addressed by simply merging markets and standardising regulatory practices. However, while by no means impossible, this solution requires a uniform position amongst a majority of participants, market and system operators, state and federal governments, and government agencies. It is therefore the most challenging option, but is also likely to produce the greatest benefits.

Standardising market designs, transmission planning and regulatory practices would go someway to reducing seams issues. In terms of open markets in the east of North America, Ontario differs from what might be thought of as the current standard design through its lack of a forward market and absence of nodal pricing. While efforts have been made to develop a forward market in Ontario, this has been undermined by industry apathy in funding it. Even with a standard market design the balkanisedbalkanised nature of markets and the lack of an overall coordinating framework will continue to limit trade.

This paper a Parallel Integrated Market (PIM) has been proposed which would allowallow the entire integrated network to be solved as a single market in parallel to the operation of the jurisdictionaljurisdictional. The PIM could account accurately for the scheduling and pricing effects of constrained looped flows, a feature not represented today which mean that today's market prices are inaccurate. The PIM would provide an efficient means of scheduling regional trade which could guide decisions within the jurisidictions, or which could even be adopted as the pre-dispatch or even real-time solution for individual jurisdictions.

Merging the markets, standardising their designs, and the PIM all require varying degrees of cooperation and commitment from industry and governmental bodies. Without this cooperation, the next best option is more localised arrangements to improve information exchanges and inter-market and inter-system operator coordination.

Fifteen years ago there were no competitive electricity markets in North America but a need for such markets was recognised. The transition to competitive electricity markets was seen as being a very difficult transition, even though it had been achieved in other countries. However, within 5 years a number of markets were operating. Today an analogous situation exists with respect to seams. A need for greater regional coordination of markets is recognised in North America.

The challenge for the next five years is to see how far industry and governments can move in overcoming seams issues. How far they are willing to move will be the major factor in determining the degree to which seams issues can be addressed or even eliminated.

About the Authors

Dr Brendan Ring is a director and principle of Market Reform. He is an expert on locational pricing of electricity and the design and implementation of electricity and gas markets. Dr Larry Ruff is special advisor to Market Reform. He is an economist specialising in market design and public policy with over 25 years experiencing in electricity market design. Usman Hannan is market development analyst with the Independent Electricity System Operator in Ontario and an adjunct lecturer in economics and trade at the University of Waterloo, Canada.

6 REFERENCES

T. W. Bessemer & F. X. Shields, In Defence of Markets, *Public Utilities Fortnightly*, June 2007.

OEB, Monitoring Report on the IESO-Administered Market Electricity Markets for the period from November 2005 to April 2006, Market Surveillance Panel of the Ontario Energy Board, June 14, 2006.

D. B. Patton, Coordinated Interchange Recommendations, Presented to New England RTO Working Group, March 13, 2003.

R Pierce, M Trebilcock, & E Thomas, Regional Electricity Market Integration: A Comparative Perspective, *Competition and Regulation in Network Industries*, Vol 8, No 2, June 2007

R Pierce, M Trebilcock, & E Thomas, Beyond Gridlock: The Case for Greater Integration of Regional Electricity Markets, *C. D. Howe Institute Commentary*, No 228, March 2006

E.G. Read and B J. Ring, Dispatch Based Electricity Spot Pricing, In M. Einhorn and R. Siddiqi (editors), *Electricity Transmission Pricing and Technology*, p. 183-206. Kluwer Academic Press, 1996.

A. THE FUNDAMENTAL ECONOMICS OF ELECTRICITY TRADING

This appendix reviews the fundamental economics of electricity trading, first within a single market and then between different markets in a region. The focus is on trade in electrical energy, but with some discussion of trade in other commodities such as capacity and ancillary services. The purpose is to define some basic trading concepts and to describe in qualitative terms the costs and benefits of regional trade in electricity.[15]

A.1. THE ELEMENTS OF A MARKET-BASED ELECTRICITY SYSTEM

Reliable and economical short-run operation of an interconnected electricity network requires that a single entity – a system controller, operator or dispatcher – know what is happening and what is about to happen physically on the system at all times and be able to change physical plans and operations quickly when necessary to maintain reliable

[15] No attempt is made to provide justification for the current paradigm of electricity market relative to other options. Bessemer and Shields (2007) provide the case for electricity markets.

operations. But the essential feature of a market is that market participants are free to act in their individual interests in response to market prices and to enter into enforceable private contracts with other market participants. Reconciling the need for central operational control and flexibility with the need for commercial freedom and stability was long regarded as impossible – and is still one of the central challenges in electricity market design. This section summarizes how the concept of a central spot market and decentralized contracts helps meet this challenge.

A.1.1. The Need for a Central Spot Market

The need for a centralized reliability function exists in many network industries other than electricity, such as commercial air travel and natural gas pipelines. If the network interactions can be managed without much effect on costs or the competitive situation of individual network users, the central reliability function can be separated from the economics and trading of the commodity itself. But the physical interactions among the parts of an electricity system are highly complex and virtually instantaneous, and reliability actions can have a strong effect on overall costs and, in a competitive market, individual network users. Under these conditions, trying to separate reliability from economics greatly increases costs and limits competition.

Vertically integrated electricity utilities have always combined, not separated, reliability and economics. Utility dispatchers use well-established economic dispatch processes to determine how to meet demand at least cost subject to reliability constraints and then issue dispatch instructions to plant operators whose job it is to follow those instructions, with all costs pooled into a system-wide revenue requirement that is recovered from captive customers. One of the main arguments against competitive electricity markets has always been, and still is, that such a process will not work unless all generation is owned or otherwise closely controlled by the same entity that determines the dispatch.

Hidden in the above argument against competitive electricity markets is the implicit assumption that there can be no *prices* that would motivate profit-seeking generators to do what is necessary for the reliable and economical operation of the system as a whole. As long as a central controller must be continually determining and directing the actions of individual generators there is no time or scope for Adam Smith's "invisible hand" to determine prices and motivate profit-seeking generators to take the required actions. If only spontaneously generated, decentralized markets are acceptable, a competitive market in electricity is at best highly inefficient and potentially unreliable, leaving no effective alternative to integrated monopoly utilities.

This logical Gordian knot preventing competition in electricity was severed when it was realized that a highly *visible* hand – the required central dispatch process – could be used to determine the prices that Smith's invisible hand could not. A modern central economic dispatch process is logically and operationally very similar to an electricity spot market: the operating cost data it uses are similar to market bids; the mathematical cost-minimizing process is logically equivalent to a market-clearing auction; the least-cost dispatch quantities are the offers that clear in the auction; and the "shadow prices" determined by the mathematics of the process are the market-clearing prices. In principle, all that is required to convert such a process into a real market is a settlement system to determine and manage the payments among market participants implied by the physical dispatch and the prices. In practice, of course, an integrated dispatch/spot market process is more difficult to design and less ideal than the simple theory suggests. But the first step toward making competition in electricity a practical reality is to accept the idea that the required central controller should operate, or at least closely cooperate and share information with, a central spot market.

The idea of combining the central system control process with a centralized spot market was first applied crudely in Chile in 1982, and then in a somewhat less crude form in the England and Wales Pool in 1992.[16] It took some years for the idea to be accepted in North America, but today all electricity markets in which competition is reasonably efficient and effective use mathematically sophisticated optimization techniques to determine both the physical dispatch and the associated market-clearing spot prices. The Ontario IESO and the system operators in all Ontario's major trading partners – i.e., PJM, New York, New England, and the Midwest ISO, but not Hydro Quebec – use some version of such an integrated dispatch/pricing process, albeit with significant differences in detail.

A.1.2. Spot Prices and Contracts

When an economic dispatch process is converted into an energy spot market, a generator's energy offer – which may be limited by market rules to reduce market power – is assumed to represent the avoidable (or marginal) energy cost for that generator at different levels of output. The spot price of energy (in $/MWh) at each time and place is (approximately, given the complexities of reality) the cost (in $) of meeting a 1 MWh increase in demand at that time and place.

In the absence of transmission constraints (and ignoring numerous other complexities such as limits on generator operation and system security issues), the market operator determines a least-cost dispatch by stacking generators in order of their offer cost.[17] The lowest-cost generation will be loaded – or cleared in the market – first, then the next-most-expensive generation will be loaded/cleared, and so on until total generation equals total demand. The spot price of energy in the market is the marginal cost offer of the highest-bidding generator needed to meet demand.[18]

When all generation is paid the market-clearing spot price for energy, the highest-cost generation covers its own short-run fuel and other avoidable costs, while lower-cost generation is paid more than its short-run costs and hence earn some operating profit. These operating profits from spot sales of energy (and ancillary services) at least help to

---

[16] In Chile, the central spot market was (and still is) open only to generators and used simple pricing rules, not a sophisticated mathematical optimization program, to determine prices that were used to settle imbalances between contract positions and actual operations. The England and Wales Pool was open to all market participants, but also used *ad hoc* pricing rules; it was replaced in 2001 with the New Electricity Trading Arrangements.

[17] Dispatchable loads may also be scheduled based on their bid prices. A load that can be dispatched off at low price will be scheduled before a higher priced generator is scheduled on.

[18] If dispatch load is present, then this may alternatively set the price.

cover a generator's fixed and capital costs. Some markets pay generators nothing more, relying entirely on spot prices – which must occasionally be very high – to compensate generators, while other markets, particularly those that limit generator bids or spot energy prices, must make additional payments for "capacity" or "capability".[19] Which ever approach is used, generators must recover adequate revenue in the short run to ensure adequate generation investment over the long run.[20]

On the consumer side, paying the spot energy price ensures that wholesale buyers (which may include some large industrial consumers) are exposed to the actual cost of changes in consumption and hence have incentives to use and to conserve energy efficiently. Although smaller consumers are often sheltered from spot prices by regulated tariffs or by "full requirements" contracts; the retailers supplying such consumers do pay the spot price and hence in principle have incentives to encourage efficient conservation by their customers. In practice, however, such upstream incentives are often weak. Greater exposure of end users to spot energy prices would encourage cost-effective conservation and reduce total energy costs.[21]

Although an electricity spot market is the key to reconciling centralized coordination with commercial freedom in short-run operations, an efficient market must have contracts to protect market participants and investors from volatile and unpredictable spot prices. In the typical wholesale electricity market, almost all energy is traded under bilateral contracts between generators, consumers and various types of intermediaries such as traders and retailers. These contracts can be for any time period from hours to many years, can contain essentially any commercial terms the parties agree and can be combined into complex chains.

Whatever the details of the contracting web, a system of contract ultimately defines the prices and amounts of energy that a generator is required to deliver to its contract customers in each trading period (usually five minutes to an hour). But the dispatch/spot market process outlined above pays no attention to the prices or quantities in bilateral contracts when deciding which generators should run at any time to meet total demand reliability and at least cost. Although total energy produced must equal total energy consumed (plus thermal losses) in each trading period, there is no necessary relationship between the amount of energy produced by each generator and the amount of energy that generator is contracted to deliver to its contract customers. Those generators that are

[19] The National Electricity Market in Australia does not make "capacity" payments but allows prices to rise to $10,000/MWh to ensure that generators that might only run for a few hours per year can receive a reasonable return on their investment. As many North American markets have significantly lower upper limits to prices some form of "capacity" payment is required instead, where these payments are paid to generators even when they do not generate.

[20] It is normal in power systems to have between 10 and 20 percent more generation capacity than the anticipated peak system load (or consumption plus losses). This "reserve margin" ensures that system load can be satisfied even allowing for planned generator maintenance and a reasonable level of unplanned generator outage.

[21] Large consumers may be directly exposed to the spot price, at least on the margin, and hence have incentives to conserve efficiently.

dispatched to produce less than their contracted quantities are, in effect, meeting their contract obligations by using excess energy produced by lower-cost generators to meet their contract obligations and, in a competitive market, must pay for that energy.

The prices used to value the generators' "overs" and "unders" are the spot prices determined by the dispatch/auction process. The market must maintain a settlement system that keeps track of which market participants are long and which are short in each trading period, prices that energy at the spot price for that period and manages the resulting payments among market participants (including the risks that some market participants may not pay).[22]

A.2. TRADING WITHIN A SINGLE MARKET

To understand the seams issues in regional electricity trading it is important first to understand the workings of an electricity market without seams. In this section we introduce the fundamental concepts of trade between two areas within a single market, i.e., within a part of the system that has a single market operator applying a single set of rules and procedures; later we consider the more difficult issues that arise when each of the areas is a separate market with its own market operator and rules/procedures, so that trade between them is "regional trade" in the sense used here. For this discussion, we assume all system operators use a market of the type outlined above to determine the dispatch and spot prices. We initially ignore the complexities caused by transmission and other complicating features, which are added later.

A.2.1. Trading Electrical Energy

Imagine a simplified electricity market with two areas: Area 1, which has more than enough low-cost generation to meet its peak demand; and Area 2, which can meet its own demand only by using relatively high-cost generation. There is a single market operator for the two areas, so generators in each area submit their bids to and receive their dispatch instructions from that market operator using the same procedures and rules.

Suppose first that the two areas of our market are isolated from each other, so that no trade is possible between them. Because there is a single market, generators in each area submit their bids to the same market operator in the same way, but the market operator must match supply to demand for each area separately, which it does by running a separate dispatch/auction process for each area. Because Area 1 has more than enough

[22] The market's settlement system may keep track of bilateral contract quantities so that only uncontracted quantities are priced and settled at the spot price, or may ignore bilateral contracts so that all energy is priced and settled at the spot price. In the latter case, bilateral contracts take the form of "contracts for differences" (CfDs). Under a CfD, the buyer and the seller each do whatever they choose in the spot market and then the buyer pays to (collects from) the seller any positive (negative) difference between the contract price and the spot price on the contract quantity. The net effect of these spot market/CfD transactions is that the contract quantity is traded at the contract price but both the buyer and the seller are exposed to the spot price on marginal quantities produced and consumed.

low-cost generation to meet its own demand but Area 2 must use higher-cost generation, the marginal generation cost/spot price will be lower in Area 1 than in Area 2.

Because the spot price is higher in Area 2 than in Area 1 (and contract prices will reflect average or expected spot prices), consumers in Area 2 would like to buy their energy from Area 1, while generators in Area 1 would like to sell their energy in Area 2. But as long as there is no way for energy generated in Area 1 to get to consumers in Area 2, the consumers in Area 2 and generators in Area 1 cannot get what they would like to have and the price differential will persist. This is demonstrated in the top part of Figure 4, where a fixed demand in each region is shown relative to the increasing marginal cost of generation at different levels of output.

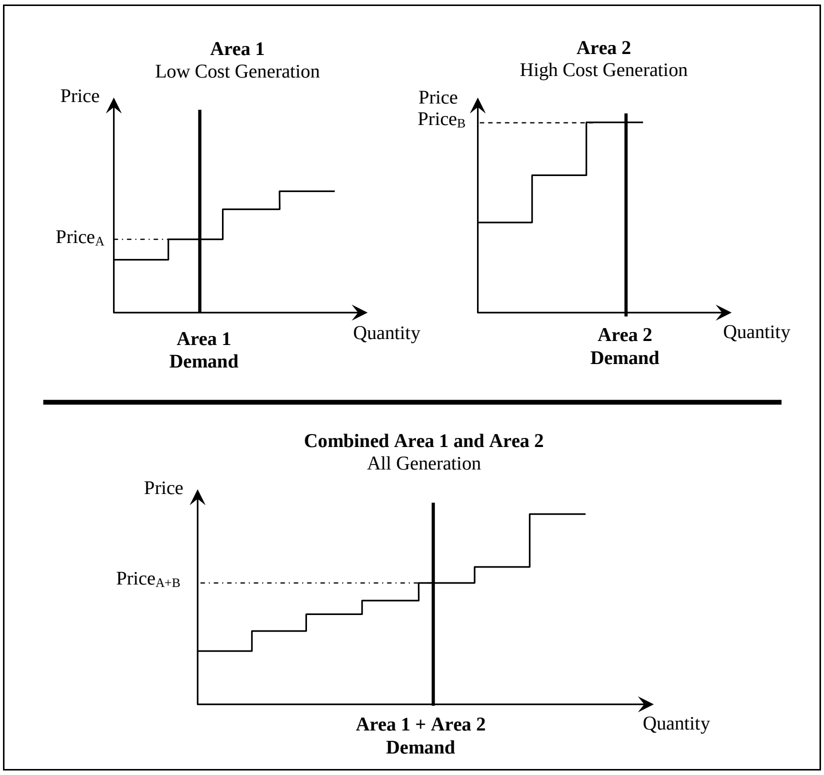

Figure 4: The impact on price and schedules of allowing trade between regions.

Now consider the case where the two areas are not separate, but are combined so that low-cost energy from Area 1 can flow into Area 2 with no losses or effective limit. The generators submit their offers to the single system/market operator as before, but the operator can now conduct a single auction/dispatch that matches supply to demand for the two areas as a whole. Compared to the market outcome when the areas are separate, the unified auction/dispatch will clear more lower-marginal-cost generation in Area 1 and less higher-cost generation in Area 2, and will have a single marginal cost/spot price for the combined area that is somewhere between the Area 1 and the Area 2 prices that applied when the areas were separate. This is shown in the lower part of Figure 4.

Compared to the market outcome with two separate areas, the outcome for the combined areas has lower total generation costs, which is presumably a good thing because it conserves scarce resources for the combined areas. Furthermore, the increase in competition among generators will increase pressure on generators to reduce costs and will reduce the market power some generators may have had, particularly in the higher-cost Area 2. But the increase in the Area 1 spot price, while good for generators there, is not good for Area 1 consumers. And the decrease in the Area 2 spot price, while good for Area 2 consumers, is not good for Area 2 generators.[23]

In the long run, introducing trade between the two market areas will affect investment and location decisions, with effects similar to the short-run effects. If generation is inherently less costly in Area 1 – because of, for example, better access to fuel or fewer environmental constraints – demand growth in the market will be met by building new generation in Area 1 rather than in Area 2, and existing expensive generators in Area 2 will tend to be retired earlier and replaced with more new, lower-cost generation in Area 1. In the long run, this will tend to lower total generation costs and reduce the (common) market price back towards (but probably not to) the price that would have prevailed in Area 1 had the areas not been combined. The levelling of the prices across the market will also reduce the advantage that Area 1 would otherwise have had for energy-intensive industries, perhaps causing some electricity-intensive industries to locate or expand in Area 2 rather than Area 1 if Area 2 has other advantages (e.g., proximity to markets or suppliers).

This simple example illustrates a fundamental point about the economic costs and benefits of expanding trade, whether specifically in electricity or more generally through, for example, globalization: Increasing trade has positive economic benefits overall, but any specific measure to increase trade will create losers as well as winners. In principle, the winners from any specific trade expansion should be able to compensate the losers so that everybody wins, and in specific cases some such compensation may be possible; but as a practical matter there will always be some losers from any expansion of trade – or from any other economic change, no matter how beneficial overall. The objective of policy should be to expand trade (and other beneficial economic activities) broadly, so that most individuals win more than they lose from the resulting aggregate economic growth, while maintaining social safety nets for the minority that does less well.

A.2.2. The Effects of Transmission

The example above ignored the transmission network in order to focus on some basic economic concepts involved in regional electricity trading. But the transmission network is so central to an electricity system, and creates so many challenges not encountered in other network industries, that it must be discussed early in any analysis of regional trade in electricity.

An electricity transmission network (or power grid) comprises multiple transmission lines – basically, metallic cables that conduct electricity – that connect generators at multiple locations to the various distribution networks that supply energy to businesses and households. A power grid has three important technical characteristics that affect electricity trade: (1) power lines and other network components, particularly transformers,

---

[23] This analysis of winners and losers is simplistic as there are mitigating factors. For example, consumers that have a share in the ownership of generators, whether via direct share ownership or government ownership, may have higher or lower purchase costs offset by lower or higher generation asset value.

heat up and lose some energy as power flows through them; (2) there are limits to how much power can be transmitted safely on a line (or, in some cases, on several lines crossing a system "interface"); and (3) power flows around the network according to the laws of physics rather than as specified in contracts or desired by traders.

The fact that energy is lost as power is transmitted on an AC power grid[24] means that more energy must be generated than is consumed in any short period, and that the marginal cost of energy will be different at different locations on the grid. For example, suppose that Area 1 contracts to deliver 500 MW to Area 2 and that 50 MW will be lost as heat on the line(s) between Areas 1 and 2. The 10 percent (average) loss factor means not only that Area 1 must export 550 MW in order to deliver 500 MW to Area 1 but that the price of energy must be at least 10 percent higher in Area 2 than in Area 1 to economically justify the 500 MW transfer with its 10 percent loss. In fact, thermal losses increase with the square of power flow, so the marginal loss on the 500th MW transmitted is 20 percent, implying that it is not economic for Area 1 to deliver the full 500 MW to Area 2 unless the price of energy is at least 20 percent higher in Area 2 than in Area 1.[25] In practice, it is difficult to determine marginal transmission losses accurately, and charging marginal losses would produce a surplus for the settlement system, so the energy prices in most electricity markets only approximately reflect the true economic effects of losses. But losses and their pricing effects can become significant, particularly for regional trading, which often involves longer distances and higher losses.

The second and usually more important effect of transmission is that the amount of power that can safely flow from one place to another on the grid is limited, sometimes in very complex, dynamic and counterintuitive ways. The simplest form of constraint is due to thermal effects: As power flow increases, individual transmission lines get hot and sag dangerously, or associated equipment such as transformers heat up to their design limits. The resulting limits on power flows are relatively simple to understand, but they change with the weather – i.e., air temperature and wind – and conditions such as the height of trees.

There are some much more complex constraints on power flow. Transmitting power over long distances causes line voltage to drop – or, more accurately on an AC system, causes voltage and current to get out of phase – which can damage transmission and customer equipment. A transmission system must be operated so that it can continue operating

---

[24] The term "energy" refers to a unit of work or heat (usually measured in megawatt-hours or MWh for electricity), while the term "power" refers to the rate at which energy is produced or flows on a transmission line (measured in MWh/hour, or megawatts). An AC (alternating current) transmission system incurs some thermal energy losses even when no power is being transmitted, as electrons vibrate in place without actually going anywhere.

[25] If total losses on a power flow of P are $L = a\times P^2$, average losses – total losses as a fraction of the power flow – are $L/P = a\times P$, while marginal losses – the derivative of total losses with respect to P, or the incremental change in losses with an incremental change in flow, are $\partial L/\partial P = 2\times a\times P$. Thus, marginal losses are two times average losses. If 550 MW must be exported from Area 1 in order to deliver 500 MW to Area 2, exports from Area 1 would have to increase from 550 MW to 551.2 MW in order to increase deliveries to Area 2 from 500 MW to 501 MW.

even if a large generator or a transmission line fails suddenly; such "contingency constraints" can limit the total power that can flow over several lines at once even when none of the lines is anywhere near its individual thermal constraints. These non-thermal constraints often reflect the judgments of experienced system operators rather than specific design limits, and can change quickly depending on conditions such as which generators are operating or the proximity of thunderstorms.

Whatever the cause of the transmission constraints, they affect system operations and, in a market-based system, prices. For example, consider the simple market discussed above, and suppose the two areas are connected by a transmission line that puts a "hard" limit on the amount of power that can flow from Area 1 to Area 2. This is shown in Figure 2.

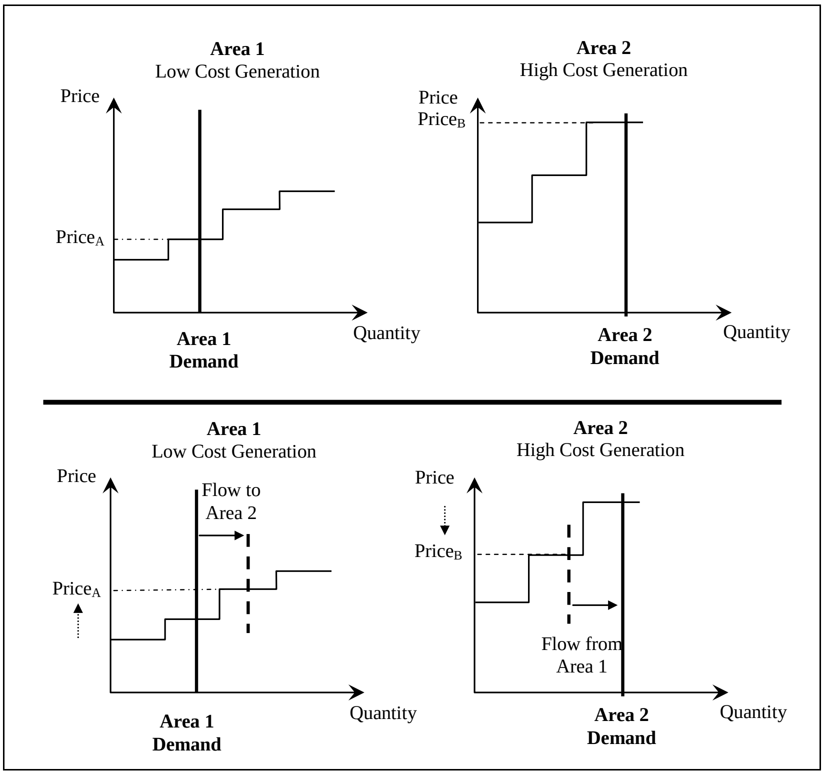

Figure 5: The impact of a transmission constraint on inter-area trade.

The top part of Figure 5 shows the earlier situation with no flow, with a low price in Area 1 and a high price in Area 2. At low flow levels, marginal losses are very low, so it is always economic for some power to flow from low-cost Area 1 to high-cost Area 2. As the power flow increases, the spot price increases in Area 1 and decreases in area 2 and the cost of marginal losses increases. If the falling spot price differential meets the increasing cost of marginal losses before the transmission constraint becomes binding, no further flows from Area 1 to Area 2 are economic and the market clears with the spot price higher in Area 2 than in Area 1 by just the value of marginal losses.[26] But if the price differential remains above the cost of marginal losses when the hard transmission constraint is reached, no further flow is possible, even though the spot price differential between the areas may be large. Thus in the lower part of Figure 5 we see that the price in

[26] If there had been no losses then the solution would be that in the lower part of Figure 1, with the two prices the same.

Area 1 has risen, and that in Area 2 has fallen, but the limitation on flow has prevented the prices from aligning.

When the transmission constraint is creating a significant price differential between the areas, generators in Area 1 are selling all their energy at the low Area 1 spot price but some of that energy is flowing into Area 2 where it is being sold for a significantly higher spot price, creating a surplus in the market's settlement system. There are several ways to interpret this result. One interpretation is that the settlement system is buying cheap energy in Area 1 and selling it for a high price in Area 2. Another is that generators in Area 1 are selling energy to consumers in Area 2 at the high price there but are paying the settlement system a transmission charge equal to the price differential between Area 2 and Area 1. Either way, the spot price differential is essentially a transmission charge for moving energy from Area 1 to Area 2. But neither of these interpretations explains why the settlement system should be entitled to collect this transmission charge or what it should do with it, which has been a significant issue in every electricity market.

In the early days of electricity markets, transmission owners tended to take the view that the settlement surplus resulting from transmission constraint charges belongs to them, or should at least be used to pay for new transmission investment. But the constraint charge is, in economic terms, a "rent" on scarce transmission capacity, and letting transmission owners keep these scarcity rents would create incentives to restrict rather than expand transmission; it would also give transmission owners a source of income that has little or nothing to do with their costs. Most markets have decided that transmission should be a monopoly business that is paid a reasonable, cost-based fee by system users, not a risk-taking owner of transmission constraints that it might be able to manipulate.

These transmission profits ideally would fund new transmission investment. If there were no transmission investment then the costs created by transmission constraints would rise over time, raising the costs of energy trade. Generators and consumers would therefore have an incentive to fund new transmission capacity so as to relieve the constraints and lower the cost of trade. This ideal model fails in reality for a range of reasons. In the context of the current discussion a fundamental reason is that a high proportion of transmission investment costs are related to securing access to land and building transmission towers or installing under-ground cabling and not the actual capacity of the transmission line. This means that the average cost of increasing transmission capacity far exceeds the avoided cost due to relieving transmission constraints.

A more practical and logical use of transmission scarcity rents is to create hedging instruments that allow market participants to contract with one another across potentially constrained boundaries. As discussed above, most market participants do not want to take spot market risk and hence most electricity is traded under contracts. In our simple two-area market, a consumer and a generator in the same area both face the same spot price, so a contract price near the expected value of that common spot price is fair and reduces market risk for both of them. But when a significant amount of Area 2 demand is met by Area 1 generation, there will have to be contracts between some generators in Area 1 and some consumers in Area 2. A generator in Area 1 will be happy with a contract price near the expected value of the Area 1 spot price, and a consumer in Area 2 will be happy with a contract price equal to the expected value of the Area 2 spot price, so the parties would be happy paying a fixed transmission charge equal to the *expected value* of the Area 1-

Area 2 price differential. The problem is that when a generator in Area 1 delivers energy to a consumer in Area 2, the transmission charge is equal to the spot price differential between the two areas *at that time*, which is highly uncertain and volatile, making a contract across the boundary highly risky for one or both parties.[27]

Thus, the existence of transmission constraint rentals raises two problems: The settlement system collects an uncertain surplus that it does not deserve; and market participants entering into contracts across the constraint must pay uncertain transmission charges. But the uncertain settlement surplus and the uncertain transmission charges are (in this simple example) exactly the same, making obvious the solution to both problems: Let market participants buy rights to the settlement surplus. If a generator in Area 1 contracts to sell 100 MWh/hour to a consumer in Area 2, the generator can purchase from the settlement system a "financial transmission right" (FTR) that pays the generator the spot transmission charge on 100 MW in every settlement period. Because this is exactly the same transmission charge the generator will have to pay for its contract delivers, it is no longer exposed to the spot transmission charge. Because the market value of this FTR should equal the expected value of the spot transmission charges, if the contract price paid by the consumer in Area 2 is just the expected value of the Area 2 spot price, the net price received by the generator in Area 1 will be the expected value of the Area 1 spot price. The contract is fair and risk-reducing for both parties.

This solution does not fully solve the problem of what to do with the settlement surplus resulting from transmission constraints, because selling FTRs simply converts the uncertain settlement surplus into a fixed amount. In most markets, the settlement surplus plus the proceeds from FTR sales are used to reduce the payments market participants make to cover transmission system and market overhead costs.

The final significant impact of transmission flows is that power flows are governed by rather esoteric concepts as voltage and impedance. Power flows from regions of high voltage to low voltage in much the same way as wind flows from high pressure to low pressure zones. The practical consequence of this is that where there is more than one pathway by which power can flow between two regions then some of the power will flow one way, the rest the other way. More particularly, a relationship exists between the proportions of power that flow on each path.[28] This means that if one of two available pathways is constrained by a thermal limit, then the flow on both is effectively

[27] In a market that keeps track of contracts for settlement purposes, the settlement system applies the transmission charge to the contract quantities in each settlement period and charges this amount to the party designated by the contract. In a market that does not consider contracts, the Area 2 consumer pays the higher Area 2 price for the quantity it takes and the Area 1 generator is paid the Area 1 price for the quantity it generates, so the parties automatically pay the price differential on the contract quantity.

[28] Power flows along the paths of least resistance (or more strictly, least impedance), which in practice means that flows will split between the pathways so as to minimise total transmission losses. Since transmission losses on each transmission line rise with the square of flow it is always better to have a small flow on each pathway rather than having a large flow on just one.

constrained. If flow increases on one pathway, then it will increase on the other and we cannot avoid violating the thermal limit. This effect is known as "loop flow".

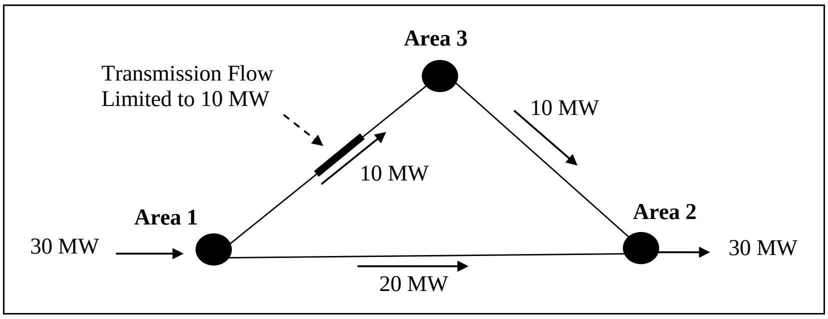

Figure 6: The impact of a transmission constraint on inter-area trade.

Figure 6 illustrates loop flow, with the trade between Area 1 and Area 2 being complicated by the introduction of Area 3. To simplify matters, we ignore transmission losses. The three transmission lines have equal impedance, which means that the impedance of the path from Area 1 to Area 2 via Area 3 is double that of the path directly from Area 1 to Area 2. The laws of physics require that 1/3 of the 30 MW entering from Area 1 flows to Area 2 via Area 3, while 2/3 flows directly from Area 1 to Area 2. If a generator in Area 1 were to try to send 33 MW from Area 1 to Area 2 then the flow from Area 1 to Area 3 would increase to 11 MW which violates the transmission limit on that path. This means that the constraint between Area 1 and Area 3 actually constrains flow from Area 1 to Area 2.

The fact that there is a constraint in the loop means that to meet an increased demand in Area 2 the system operator must reconfigure the pattern of generation around the loop so that as to supply that increased demand while simultaneously keeping the transmission constraint within its limit. Thus if demand were to rise to 31 MW at Area 2 a possible solution might be to have a generator costing \$10/MWh in Area 1 raise its output by 0.5 MW to 30.5 MW, with 1/3 of this power flowing to Area 2 via Area 3 and 2/3 flowing directly, while a (more expensive) \$40/MWh generator in Area 3 increases its output from 0 MW to 0.5 MW, with 1/3 of this power flow to Area 2 via Area 1 and 2/3 flowing directly.

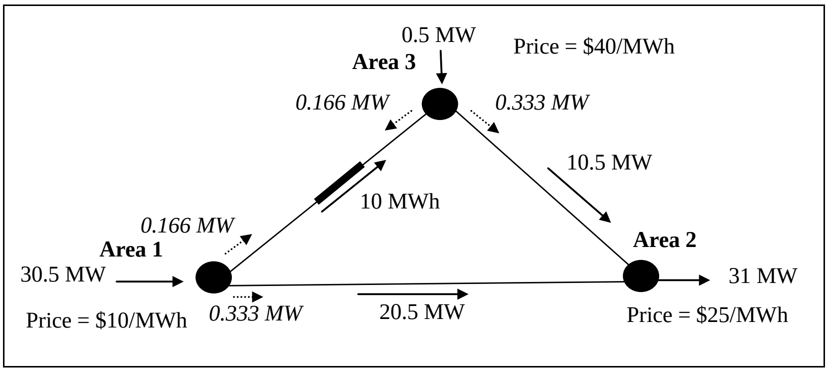

Figure 7: The impact of a transmission constraint on inter-area trade.

Figure 7 shows that this solution can be accommodated by the network. Note that italicised text indicates the incremental impact on flows due to the extra 0.5 MW of generation from each of Area 1 and Area 2. Because 1/3 of Area 1's extra flow is 0.1666

MW and is towards Area 3 and 1/3 of Area 3's extra flow is 0.1666 MW and is towards Area 1 then these two flows cancel and keep the net flow to 10 MW, which is within the transmission limit. However the contributions of Area 1 and Area 2 are additive on each of the transmission lines connecting Area 1 with Area 2 and Area 3 with Area 2 so that an additional 1 MW of power arrives at Area 2.

Loop flow has significant implications for electricity prices. Before flows rose to a level where the transmission constraint has an impact, the $10/MWh generation in Area 1 could provide energy to all locations and, ignoring losses, prices everywhere would be $10/MWh. But once transmission flow levels rise to a point where the transmission limit applies, prices diverge in complicated ways. The price in Area 1 remains $10/MWh as generation there would still provide any increased demand in Area 1, while the price in Area 3 has to rise to $40/MWh to fund the generation there. The price in Area 2 is just the cost of the last MW delivered there, which must be half of the Area 1 price plus half of the Area 3 price or $25/MWh. Effectively the price rises from a low level on one side of the constraint to a high level on the other. This is often known as the "spring washer effect".[29] The importance of the spring washer effect is that the price in one area can be substantially driven by costs in multiple other areas. This has significant implications in the context of regional pricing.

A.2.3. Ancillary Services

The primary purpose of electricity markets is to facilitate the efficient trade of energy. However there are other "ancillary services" which must be procured and without which energy could not be reliably delivered. These services are used to preserve system local voltage levels, the operating frequency of the power system (e.g. 60 Hz), and to provide security in the event of generator or transmission line break-downs. Some ancillary services are of a nature which favours them being procured through a tendering process by the system operator. An example of one such service is "voltage support", involving installing specialised equipment around the network to help maintain system voltage levels. Other ancillary services are of a nature which allows them to be priced in a spot market just like energy.

Operating reserve allows the power system to recover quickly from the sudden failure of generating or transmission equipment and can be provided by generators or by consumption that can be interrupted quickly (this is called "curtailable load"). There a various forms of operating reserve, such as "spinning reserve" (or "synchronized reserve") and "replacement reserve" (or "nonspinning reserve" or "nonsynchronized reserve"). A generator providing "spinning reserve" would be scheduled to provide energy, but would operate at lower level of output than might be expected given the prevailing energy price. A generator providing "replacement reserve" would not be operating at all. In the event of a generator failure in the power system, the power system frequency will start dropping below 60 Hz and generators providing spinning reserve will automatically increase their output. This increase keeps supply and demand in balance, and halts the fall in system frequency. Spinning reserve responses in North America occur in a time frame from seconds to 10 minutes. Over longer intervals, and up to several hours after the event,

[29] The spring washer effect is described more fully by Read and Ring (1996).

generators providing replacement reserve can be instructed to run (as required) by the system operator. Replacement reserve is primarily used to allow the spinning reserve providers to back off again so as to restore the original levels of available spinning reserve. Over even longer periods, replacement reserve generators will be backed off as the market processes find new efficient generator schedules to address the changed circumstances.

In addition to the spot price for energy, a spot price can be defined for spinning reserve. Consider a generator which makes energy available to the market at a cost of $50/MWh. If the market price were $80/MWh the operator of the generator would expect the generator to be fully scheduled, allowing it to earn $30/MWh profit on its output. Now suppose that the system operator wanted this generator to reduce its energy production by 10 MW so as to increase the spinning reserve available to the market. By reducing output by 10 MW the generator operator would avoid $50/MWh of production costs, but would also miss out on energy payments of $80/MWh. To induce the generator operator to want to provide this spinning reserve it would need to be paid at least $30/MWh, its forgone profit, on the 10 MW. Thus we might expect in this case that the energy price is $80/MWh and the spinning reserve price is $30/MWh.[30]

Regardless of how suppliers of ancillary service are paid, whether via contract or spot price, these services are purchased for the ultimate benefit of the market as a whole, so the cost of these services tend to be spread amongst generators and consumers outside of normal market processes under various types of cost allocation rules.

A.2.4. Planning and Scheduling Processes

Market participants, transmission operators, system operators and market operators must coordinate their plans and actions so as to ensure successful market clearing, feasible transmission schedules, and reliable supply. This process does not happen in real-time, but is a culmination of a planning process that commences a long time prior to real-time.

Years and months before the day of trade buyers and sellers enter into bilateral contracts for energy. These contracts provide security for the financing of new generation and demand-side projects to respond to demand growth. Transmission system operators planned augmentations to the existing transmission network over similar timeframes.

Weeks and months before the day of trade a maintenance coordination process must be under-taken, typically coordinated by the system operator to minimise the risk of involuntary load curtailment. The process coordinates the timing of generator and

---

[30] This is a relatively simple example which it ignores some complications. For example their may not be a one-to-one relationship between how much spare generating capacity a generating facility has and the amount of spinning reserve it can provide, causing the required price for spinning reserve to differ from the unit profit foregone by the generating operator. Also, actual market implementations often allow generators to specify a cost for the spinning reserve they offer to the market. The generator should only be scheduled if the price for spinning reserve exceeds this cost plus the profits forgone on the energy market. Finally, there may be more than one class of spinning reserve, with prices determined for each class.

transmission network maintenance to ensure that adequate equipment is in service to meet the likely most extreme load.

In the days prior to the trading day the main focus of generating operators is on planning which of their plant to operate. A typical generator operator may control multiple generator stations where each station may have multiple generating units which can be run independently. This “unit commitment” problem is important as significant costs can be incurred in starting up generators, it takes a long time to start some generators, and once shut down some of these generators cannot restart for many hours or even days. In some markets this unit commitment problem is solved centrally by the system operator given information on generating unit start up costs and limitations, while other markets leave it to the individual participants to schedule their own facilities based on expectations of price.

Some market operators facilitate a day-ahead forward market prior to the trading day. These forward markets use the most up to date data to schedule energy to match supply and demand through the trading day. A day-ahead forward market effectively is a centralised process by which market participants can adjust their contract position, but rather than the contracts being bilateral between the participants they are contracts between the participants and the market operator.[31] Many day-ahead forward markets incorporate unit commitment problems so as to ensure that unit commitment costs can be fully funded based on the day-ahead schedule.

By real-time most participants have sold or purchased the vast majority of their energy through long term contracts and the day ahead forward market. This limits their exposure to surprise events in real-time.

Power system operators oversee the operation of the power system in real-time, typically relying on periodic re-runs of market scheduling and pricing software to improve the efficiency of their schedules. Generators and dispatchable consumers follow instructions from the power system operator in real-time. This ensures that the power system can respond rapidly to unanticipated events without reliance on market processes alone.

Markets typically settle transactions in the days following real-time based on forward market positions, trading day schedules and prices, financial transmission rights, and meter data. Some markets facilitate the settling of bilateral contracts between participants, but where this is not the case the participants have to settle these contracts separately.

A.3. TRADE BETWEEN REGIONAL MARKETS

Power systems typically developed either as government owned industries or as regulated regional franchise based private companies which controlled all aspects of power production, transmission and sale within a region. As markets in electricity have evolved, so the boundaries of those markets have matched these regions.

[31] The market operator sells as much energy as it buys, so it has no net position.

The high value placed on power system security, and the attraction of trading energy, have meant that over time physical links called "interconnections" or "interties" have arisen between regional markets. In eastern North America a complex network has evolved that spans states, provinces, countries, power system control areas, and markets. Power flows across this network are governed by the laws of physics not commercial deals. How then do neighbouring power system/market operators manage trade with markets outside of their own region?

Even before modern electricity markets developed, system operators would allow trades to be scheduled with other regions either as imports to their system, exports from their system, or "wheeling" trades across their system. Arbitrary charging rates applied for these transactions. With the advent of markets it is now possible to schedule and price these transactions based on market prices in each of the sending and receiving markets. The problem encountered is that the transaction must be scheduled and priced in two different markets. The boundaries of these markets have become known as "seams" and the problems created by these seams are known as "seams issues".

Consider the situation where a supplier in the mid-west of the United States has surplus low cost power while a consumer in New York City has a need for more power and is willing to pay for it. In an idealised market the supplier could offer that energy into a market while the consumer in New York City could bid for that energy, and the market would match those trades and take care of all the transmission considerations automatically. Neither of the trading parties, or the operator of these idealised markets, would need to know anything about the other party.

Now let us consider the reality of a trade from the mid-west to New York City. The supplier might be in the market operated by the Mid West ISO (MISO) while the consumer is in the market operated by the New York ISO (NYISO). When power is injected into the network in the MISO market it does not flow directly to the customer in New Your City. Instead, loop flow comes into play. Some of the MISO injection will flow to New York market to the south of Lake Erie via the market of the Pennsylvania-Jersey-Maryland Interconnection (PJM) while some will flow north of Lake Erie (via Michigan and across the Canadian border) to the Ontario market and onward (across the Canadian border again) to the New York market.

This transaction must be scheduled in four different markets, all of which have subtle differences in their scheduling processes. It is not possible to have the transaction scheduled simultaneously in all those markets. Even once the transaction is scheduled, real-time events can mean that the scheduled transaction might be cancelled by one or more system operators, effecting the entire transaction. In completing this transaction some of the power is imported to Canada and then exported back to the United States, creating obligations in regard to international trade and taxes.

This simple example illustrates the nature of "seams" issues in the context of energy scheduling. It does not touch on the additional complications of forward markets, unit commitment, operating reserve trading, or access to and pricing of transmission networks. While opportunities for trade exist, and trade has become easier between deregulated markets, these seams issues create complexities and disconnections in the trade process that constrain the volume and efficiency of trade.

A.4. THE COSTS AND BENEFITS OF REGIONAL TRADE

The costs and benefits of regional trade are identical to those already described for area trade within a region. Generation will tend to increase in low cost regions as more energy is exported, with an associated increase in prices in those regions, while generation will tend to decrease in high cost regions as it is displaced by imported lower cost energy, causing prices there to fall. Over-time the efficiency of the electricity industry should improve with the average cost of production falling, at least relative to what it would have been otherwise, with associated flow on effects to energy users. Improvements in generation efficiency should also have benefits for the wider environment. While there is an overall societal benefit from regional trade there will be winners and losers within and between the regions.[32]

[32] Pierce et al (2007) argue in the Canadian context that greater integration between the North American markets should have tangible benefits.