DAE Working Paper WP 0214

UNIVERSITY OF
CAMBRIDGE
**Department of
Applied Economics**

Optimal congestion treatment for bilateral electricity trading

Karsten Neuhoff

The
Cambridge-MIT
Institute

*Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research*

CMI Working Paper 05

DAE Working Paper Series

UNIVERSITY OF
CAMBRIDGE
Department of
Applied Economics

not to be quoted without permisson

The
Cambridge-MIT
Institute

Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research

CMI Working Paper Series

Optimal congestion treatment for bilateral electricity trading

Karsten Neuhoff[1]
University of Cambridge

April 26, 2002

[1] I am deeply indepted to David Newbery, Tanga McDaniel, Michael Pollitt, Paul Joskow, Ralph Turvey, Alex Henney, Richard Court and Min Zhu for comments and suggestions. Research support from Gottlieb-Daimler Foundation and Economic and Social Research Council (award R42200034307) is gratefully acknowledged. (correspondence address: Karsten.Neuhoff@econ.cam.ac.uk)

Abstract

Recently liberalised electricity markets frequently focus on bilateral trading rather than on a centralised pool. Most prominent is the recent shift from a pool-based system to bilateral contracts in England and Wales. Still to be addressed is the question of how to treat transmission constraints in a bilateral trade setting in meshed electricity networks. Approaches can be classified into three categories: First, constraint resolution by the system operator and socialisation of incurred costs; second, physical transmission contracts; and third, locational charging with the option of financial hedging.

The paper shows that constraints resolved by the system operator with socialisation of costs can results in inefficient dispatch and incorrect incentives for the location of new generation plants. Most transmission access designs are optimal in a very simplified model world, however the paper shows that transaction costs, illiquid markets and uncertainty about demand create significant differences between the concepts. Physical transmission contracts are best designed as zonal access rights, but have to be centrally administered to be efficient. However, only locational charging can cope with the uncertainty and volatility of electricity demand in an efficient and non-discriminatory manner. If required it can be complemented with simple financial contracts to allow hedging on an annual basis.

The paper gives qualitative arguments and illustrative examples, which suffice to rank explicit constraint treatment mechanisms. However, network specific analysis is required to balance efficiency gains from explicit constraint treatment with additional transaction costs.

1 Introduction

In England and Wales, the new electricity trading arrangement (NETA) replaced the Pool on March 27, 2001. Now, generators, electricity suppliers and large consumers trade bilaterally and notify the SO of energy injections and withdrawals. The system operator (SO) has to accept all notifications. The SO then accepts bids and offers from generators and demand side in the balancing market $3\frac{1}{2}$ hours before dispatch to resolve transmission constraints as well as to resolve energy imbalances.[1] Due to low liquidity in the imbalance market the SO (National Grid) started contracting with generators outside of the balancing market to resolve anticipated constraints and imbalances. Costs for resolving transmission constraints are socialised among all market participants whilst imbalance charges are attributed to responsible parties. In this paper, I will refer to this arrangement as implicit congestion treatment. This system allows generators in export constrained zones to obtain monopoly profits. It results in inefficient dispatch and incorrect incentives for the location of new power plants.

In a May 2001 consultation document, the electricity regulator Ofgem suggested an explicit treatment of congestion to complement NETA[15]. Ofgem intends to create tradable property rights for access to the transmission network. I show that the proposed approach is close to an optimal implementation of property rights in a simplified theoretical analysis and only requires a few modifications to correspond to a theoretical optimal design. However, any implementation of property rights has severe shortcomings which result in inefficient dispatch, discrimination against small and intermittent generators and consumers, and high transaction costs.

I suggest an alternative approach without physical property rights which is based on the underlying idea of nodal pricing, as initial presented by Bohn et.al.[2] with the option to use financial transmission contracts to hedge against price risk as presented e.g. by Hogen [10] and further examined by Joskow and Tirol [12]. As already currently implemented under NETA the SO continues to resolve congestion in the balancing market with the crucial difference that costs are no longer socialises. Instead, generators in export constrained zones pay the marginal costs of congestion resolution whilst generators in import constrained zones obtain such payments.[2] The approach would reduce market power in export constrained zones and incentivise new generators to locate in import constrained zones. The risk of uncertain congestion charges can be hedged against with tradable congestion contracts issued by the SO. The price of the contracts conveys information about the expected costs of redispatch. Generators anticipate the planned generation pattern and theoretically remove - and, practically, at least reduce - the amount of redispatch required to eliminate constraints.

One motivation for the implementation of explicit constraint treatment in the England and Wales is to provide efficient signals for the location of new generation plants. Implicit constraint

[1] $3\ \frac{1}{2}$ hours prior to dispatch is referred to as "gate-closure" as it represents the end of the bilateral markets and the beginning of the balancing market (called balancing mechanism).

[2] Subsequent discussion assumes a single buy and sell price in such a balancing market, but could equally be applied to the current approach of pay-as-bid in the balancing market.

treatment can strongly distort such signals in a heavily congested transmission system. The open question not answered in this paper is whether the low level of constraints justifies the transaction costs associated with any explicit constraint treatment mechanism or whether other locational signals, e.g currently applied locationally differentiated annual access charges, are more appropriate.

The second section of the paper illustrates the shortcomings of the implicit treatment of congestion. The third section looks at how property rights can best be implemented. The fourth section discusses the shortcomings inherent even in an optimal implementation of property rights. The fifth section is an examination of how the current regime with a SO resolving constraints can be improved upon by charging costs to liable parties. The advantages of financial transmission contracts to hedge risk and convey information are also discussed.

2 Shortcomings of implicit congestion treatment

Currently, generators and consumers trade bilaterally and make their final announcements of where they insert and withdraw energy at gate closure. If the resulting transmission schedule violates physical constraints, the SO uses the balancing market to create counter-flows. Costs incurred in the balancing market are smeared over the users and therefore are not location-specific.

In section 2.1, the effects of implicit congestion treatment in a competitive environment are illustrated. Generators in export constrained zones retain transmission rents which would otherwise be collected by the SO and available to cover i.e. fixed costs. Furthermore, cross-subsidising demand in the importing zone results in inefficient allocation and welfare losses.

Section 2.2 assumes the presence of regional market power. Initial assessment for a two node case suggests that market power is enhanced by implicit congestion treatment.[3]

2.1 Welfare losses in a competitive market

Two nodes A and B with competitive generation and competitive demand and one capacity constrained interconnector are used to compare explicit and implicit congestion treatment. Three methods of explicit constraint treatment result in the same efficient allocation: nodal pricing, physical transmission contracts and charging the parties, which contribute to the congestion, at marginal cost. Under implicit congestion treatment competitive generators can capture transmission rent and are weakly better off than under explicit congestion treatment. Consumer losses and lost transmission rent are greater than the gains of generators; therefore, implicit congestion treatment reduces total welfare relative to explicit congestion treatment.

[3] This result is supported by a recent paper comparing nodal and zonal pricing in the presence of market power. Borenstein shows that nodal pricing, which would correspond to explicit constraint treatment, is superior to zonal pricing, which contains an implicit constraint management component [3].

I first calculate the welfare gains from an interconnection using explicit congestion treatment relative to a situation without interconnection, then the welfare gains from an interconnection using implicit congestion treatment relative to a situation without interconnection. Then both results are compared.

2.1.1 Welfare gains due to interconnection with explicit constraint treatment

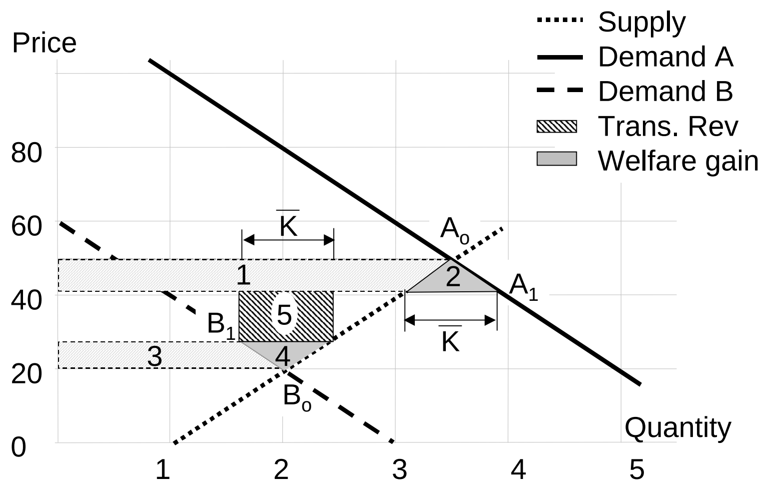


Figure 1: Welfare gains from interconnecting two competitive nodes with a capacity constraint interconnector with explicit congestion management.

Figure 1 illustrates the basic model used in the following discussion. Two nodes with competitive demand and generation are interconnected with a transmission line with capacity $\overline{K}$. Transmission losses are ignored. The dotted line through A_0 and B_0 represents marginal generation costs, which are identical at both nodes. If more capacity is required, more expensive generators have to be scheduled; therefore marginal costs are increasing with output. The continuous line represents demand in node A and the dashed line demand at node B. Demand decreases if prices rise. If the nodes are not interconnected, then the competitive equilibrium is given by the intersection of demand curve and marginal costs of generation at point A_0 for node A and B_0 for node B.

Prices in node A are higher; therefore, electricity will be sold to node A if an interconnection is installed. The transmission line has a limited capacity $\overline{K}$, therefore additional energy imported to node A can only reduce prices and increase demand to the point A_1, whilst increased prices reduce demand in node B to point B_1.

Reduced prices reduce profits for generators at node A. The size of area 1 represent these lost profits. The area equals the product of price change and quantity of output sold plus lost profits of generators which reduced their output. At the same time, reduced prices increase consumer surplus in node A. The combined size of areas 1 and 2 equals consumer surplus: the product of quantity consumed times price change plus additional surplus due to electricity consumption which was prevented at higher prices. The difference between consumer surplus and producers' losses equals welfare gains in node A, represented by area 2. Applying a similar argumentation in node B shows that generators' profit increases more due to price increases than consumers' surplus is reduced, resulting in welfare gains in node B (area 4).

A third component of welfare gains is due to the congested transmission line. More access to the transmission line is required than is available. Therefore scarce access rights are auctioned in order to ensure non-discriminatory access for all parties. The auction is expected to produce transmission revenue (area 5) corresponding to the volume of transmission rights times the value of each right, which is given by the price difference between node A and B.

Auction revenues can be significant in liberalised markets. Transmission rights from Germany to the Netherlands were auctioned for 10 Euro/MWh in May 2001, roughly half of German electricity prices. The transmission operator can collect profits corresponding to area five from the interconnection if prices in both markets are well-defined in a competitive setting

2.1.2 Welfare gains due to interconnection with implicit constraint treatment

Figure 2 illustrates the components of welfare gain from applying implicit congestion treatment in the previous setting consisting of welfare gains in the energy market and welfare losses in the balancing market.

In the initial energy market (graph on the left side), generators and consumers can contract irrespective of transmission constraint. The result is a homogeneous market clearing price for both nodes A and B. Assuming the required transmission from node B to node A could be performed, the two triangles formed by the black and grey areas represent the total welfare gains. In the exporting zone additional generators' profits outweigh reduced consumer surplus, whereas, in the importing zone, additional consumer surplus outweighs reduced generators' profits. The grey area corresponds to welfare gain which was already achieved in the previous section (generators' profits, consumer surplus and transmission rent). The black area represents additional welfare gains if the transmission link is assumed not to be capacity constraint.

The graph on the right side shows, how the SO interacts in the balancing market to resolve the congestion. The SO has to reduce generation at the exporting node B until the difference between generation and demand equals the transmission capacity $\overline{K}$. Simultaneously, the SO has to increase generation at the importing node A by the same amount.[4]

The SO must pay generators at node B not to produce. The payment itself represents a redistribution and is therefore welfare neutral. However, usually, the required funds have to be collected with a tax on all consumers, thereby distorting efficient allocation. For the sake of simplicity, we ignore this additional disadvantage of implicit congestion treatment in the subsequent analysis.

[4] Demand side response in the balancing market is possible but rare. For reasons of simplicity it is ignored, but would not alter the qualitative result.

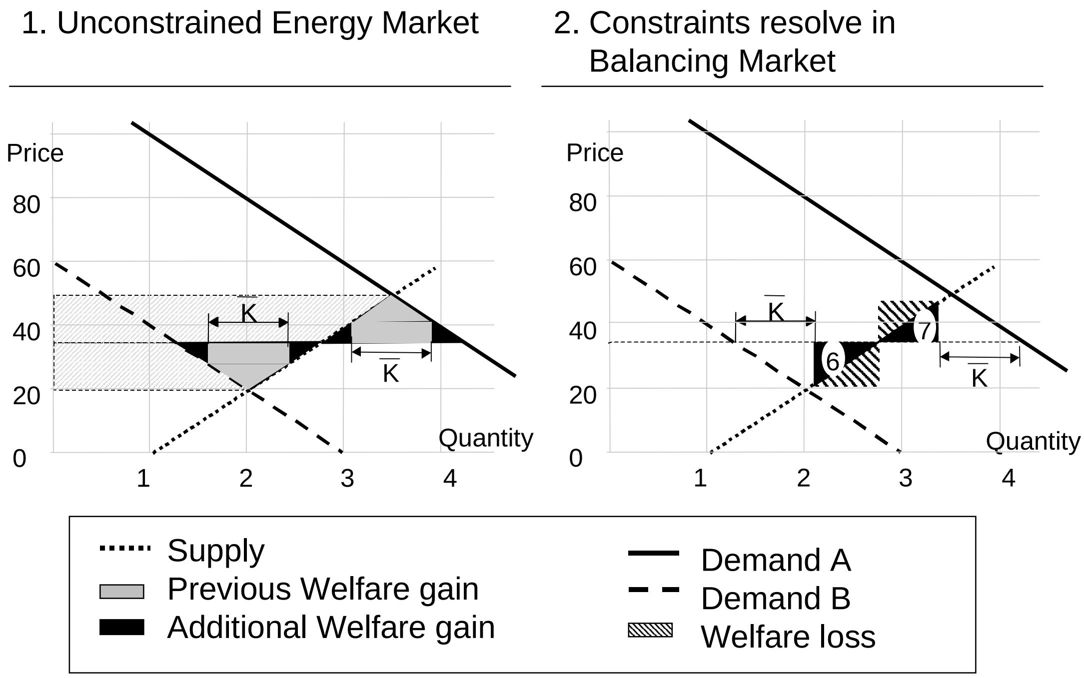


Figure 2: Total welfare gains from interconnecting two competitive nodes with a capacity constraint interconnector consist of gains in the energy market and losses in the subsequent balancing market. The initial energy market determines an allocation corresponding to an unconstrained transmission link. Then the SO reschedules generation in the balancing market to resolve constraints.

However, the rescheduling has a welfare component as well. Generators at node B which are paid for in the balancing market not to produce would have produce at lower marginal costs than the generators at node A which are paid to produce instead. Area 6 represents the welfare losses of generators not producing, which are more efficient than the marginal generator in the energy market. Area 7 shows the welfare loss due to additional costs of generators which are more expensive than the marginal generator in the energy market.

2.1.3 Comparison of explicit and implicit congestion treatment under perfect competition

Implicit congestion treatment appears to allow more welfare gains in the initial energy market than explicit treatment. However, the SO must subsequently resolve constraints. Welfare losses in the balancing market (areas 6 and 7) are twice as high as the additional welfare gains from implicit congestion treatment in the energy market relative to an explicit congestion treatment (black area in figure 2-1). This means implicit congestion treatment creates less welfare gains than explicit congestion treatment.[5]

Explicit constraint management allows the SO to collect transmission revenue (area 5 in Figure 1). The subsequent utilisation of revenue can be determined by the regulator. Under implicit congestion treatment, the revenue is split between generators in the exporting zone and the consumers in the importing zone, and can only be collected with the help of distorting taxes. Such taxes are required to finance the rescheduling which must be ordered by the SO for constraint resolution.

The third, and potentially most severe, disadvantage of implicit constraint management is that generators which are scheduled in the energy market obtain the same revenue at both locations. The decision for location of new generation capacity is not influenced by the costs of constraint resolution brought about by the investment. For some generators, it is even more profitable to locate at the export constraint node. Let us suppose a generator at node A with generation costs X between 20 (money/MWh) and 32 first contracts in the energy market at a price of 33. His revenues exceed variable costs by 33-X. Subsequently, in the balancing market, he buys energy at the market clearing price of 20, saving variable costs of X-20. Total revenue is 33-20=13 without incurring variable costs for energy. The same generator at node B would have to generate and receive revenue 33. Variable costs X are between 20 and 32, leaving the generator between 1 and 13 revenue after paying variable costs. Under explicit congestion management, these generators would only be contracted for at node B and therefore would locate at the efficient location.

2.2 Effect under market power

The subsequent analysis only represents an initial assessment. Generators in constrained zones can profitably abuse the system. Imagine an oligopoly in an export constrained zone. Oligopolists will sell more electricity than can be transmitted on the interconnection. The SO realises that the transmission constraint is violated and resolves it in the balancing market: oligopolists in the constraint zone reduce generation relative to the announced schedule and generators outside the constraint zone increase generation. Oligopolists profit by obtaining general market prices for initially contracted electricity and then using locational market power in the balancing market to pay a lower price to the SO for the fraction of the electricity they do not produce themselves. To increase profits, oligopolists will over-contract in the energy market in order to increase the volume to be resolved in the balancing market. In the balancing market, generators have local market power and can thereby reduce the price at which they sell their energy back. Harvey and Hogan present a similar argument in a discussion of zonal vs. nodal pricing [9]. It is shown that nodal pricing is a better response to local market power than zonal pricing, with its subsequent socialising of rescheduling costs.

These profits from abusing transmission constraints to exercise market power are additional

[5] Total welfare of explicit congestion treatment only exceeds total welfare of implicit congestion treatment if demand elasticity is above zero.

to profits to be made from regional market power under any (implicit or explicit) treatment of market power. Even if transmission constraints are treated explicitly, the oligopoly at the exporting node will reduce output to increase price and capture the transmission rent (area 5 in Figure 1). This result is supported by experiments [1].

2.3 Conclusion on implicit congestion treatment

The previous analysis for the competitive case shows that implicit congestion management results in inefficient scheduling of generation and cross-subsidises demand at the import constrained node. National Grid Company calculated that if 100MW of energy are inserted in West Lancashire located in the North instead of the South Coast, total system losses are increased by 8 MW.[6]

Generators at the export constrained node receive transmission rent which otherwise would be collected by the SO and could be used to reduce transmission tariffs for all consumers. Implicit transmission management does not only fail to confront generators with the cost of constraint resolution but even incentivises some generators to locate at the export-constrained nodes. Newbery has estimated that incorrect locational signals for investment could increase generation costs by 12% [13].

Initial analysis of the situation with market power suggests that these effects are only enhanced. The result is supported by Borenstein [3]. He concludes from a theoretical analysis and comparison between historical and present flow patterns between the western and eastern portion of the PJM electricity control area in the North-East of the USA that due to the exercise of zonal market power transmission links can be more frequently congested after liberalisation.

Even without explicit constraint management two further methods are available and are applied to support efficient allocation and scheduling. First, generators can be charged for marginal losses due to their transmission, and second, connection charges can be location-differentiated. Marginal loss charging over-recovers energy losses by the factor two and therefore makes money available for capital or operating expenditure. If grid usage were not peaky and investment not lumpy, then over-recovery from marginal loss pricing would suffice to recover fixed costs. In reality, grid usage varies over the day. A company locating a new generation facility for peak supply based on gas in the North would pay marginal loss factors for transmission to the South for only a fraction of the year, insufficient to recover investment costs for additional transmission. The same company, locating in the South, would have to obtain firm transmission rights for gas which are priced corresponding to usage at peak times. Thus the company might decide to locate in the North, even though total costs for transmission investment are lower for gas than for electricity. This is because the company would have to pay for the peak capacity in gas, but under marginal pricing rules not in electricity.

[6] This effect can be incorporated by marginal pricing of transmission losses - but will not result in efficient dispatch if revenues from locational market power are higher than additional costs for transmission losses.

The problem can be either solved with an explicit constraint treatment or with deep connection charges. These charges imply that a generator newly connected to the grid not only has to pay for the direct link from his facility to the next high voltage line, but also has to bear the costs of any required grid investment to handle the additional load. As it seems difficult to assign such costs properly, deep connection charges have not been applied so far. A thorough analysis is advisable, but beyond the scope of this paper.

I have neither attempted to quantify welfare losses due to implicit congestion management nor the costs of explicit congestion management[7]. Therefore a final judgement about the necessity of explicit congestion management is not offered here.

3 Property rights for transmission

In a May consultation document the electricity regulator Ofgem suggested a new transmission access regime to complement NETA (Ofgem2000a). Ofgem intends to create tradable property rights to regulate usage of the transmission network.[8]. The Coase theorem states that if trade in the externality can occur, then costless bargaining will lead to an efficient outcome, no matter how property rights are allocated [6].

In electricity networks, the capacity of every transmission line and transformer is limited. Depending on the load and generation pattern, many of these constraints can be binding at some time of the year. Too many contract types would be required to cover all relevant constraints; therefore Ofgem proposed to bundle property rights to several constraints in order to create entry/exit rights.[9] A company which wants to transmit electric energy only has to obtain one entry right at the starting point and one exit right for the end point of the transmission. These two rights encompass access rights to all constraints involved in the transmission. Section 3.2 shows that Ofgem's approach represents the best bundling strategy for property rights.

Although starting from a good bundling strategy, the current Ofgem proposal creates inefficiencies in two respects. First Ofgem suggests that entry/exit rights with a positive price are options. Section 3.3 argues that, in meshed networks, all rights must be obligations. A second cause of inefficiencies is that the SO can not effectively reconfigure entry/exit rights. Section 3.4 illustrates that in the presence of uncertainty, efficient usage of scarce capacity requires the SO to simultaneously buy back several entry/exit rights and to rebundle the underlying access rights and then sell different types of entry/exit rights.

[7] See Richard Green [8].

[8] The consultation document furthermore suggests a change in the treatment of transmission losses towards a per MWh charging of time-averaged marginal losses. I support this approach because it gives proper locational signals for investment and dispatch.

[9] If demand can be predicted with sufficient accuracy by the system operator and is assumed to be price inelastic, then it suffices to create only entry rights.

3.1 Criticism of Flow-gate rights

Flow-gates rights are an attempt to directly match property rights with the scarce transmission capacity. Transmission lines which are expected to be congested are defined as flow-gates. Any energy transmission between two points of the network affects all flow-gates because electric flows cannot be directed in a meshed AC network. The energy flow is divided amongst all possible paths proportionally to the inverse of the resistance along the path. This phenomenon leads to so called loop flows. For any energy transmission between two zones of the network, one would therefore have to obtain property rights for all flow-gates.

Pre-defined constant flow-gates offer little scope to adapt to dynamic developments within a network. Constant proportionality factors have to be assumed in order to allow companies to hedge risk with flow-gate rights. However, in a real network, the flows are divided between different paths, depending on line outages due to maintenance, reactive power flows due to different generation patterns and the increasing usage of phase shifting devices (Flexible AC transmission networks). Furthermore, the maximum capacity of a flow-gate is usually not determined by the thermal constraint of a line but by the (n-1) security constraint. This implies that the capacity of a flow-gate changes dependent on the network situation. Changing proportionality factors and changing flow-gate capacities imply that the flow-gate approach does not allow for an efficient use of the network [17].

Hogan documents the difficulties of flow-gates [11]. He assumed in a counterfactual study that the PJM system would be based on flow gate rights instead of nodal pricing. In the period January 1998 to April 2000, there were 161 unique constraints that produced congestion in PJM. Hogan concludes that a complete flow-gate model would require purchase of 161 flow-gate rights to secure a single point-to-point transaction. One might expect this number to drop if minor constraints were ignored - but even if only 80% of the constraint hours are to be covered, 50 flow-gates would be required, still implying significant transaction costs.

3.2 Zone-to-zone contracts are inferior to zonal entry and exit rights

Obligatory zone-to-zone contracts and firm entry and exit rights are compared. The section is based on the simplification, that zones are small enough, such that they can be represented by nodes. However, the results comparing different designs directly translate to larger zones.[10] Appendix 1 contains the proof that obligatory zonal access rights and zone-to-zone contracts represent the same physical constraints. This is due to the fact, that electricity flows are independent of the pattern of generation and consumption. Additional electricity flows superimpose upon existing flows. Two flows in opposing directions cancel each other. Therefore, the marginal cost of transmitting the first flow is exactly the negative of the cost of the second flow. In a system with N zones, $N*(N-1)/2$ zone-to-zone contracts are defined, whereas the same system

[10] To allow for an homogenous representation we will continue to refer to zones rather than nodes, see section 4.1 for further discussion.

only requires N zonal access contract types.[11] Obligatory contracts can be held in negative and positive quantities, therefore one type of contract suffices for electricity withdrawal and insertion at each node.

As both designs represent the same underlying world the higher number of contracts in the zone-to-zone design does not reveal additional information. Given the same amount of trade interactions, the liquidity of contracts in the zone-to-zone design will be lower than in the other design. More liquidity tends to reduce the volatility of prices. Transaction costs are reduced in a less volatile market because the market interactions can be automated when individual trades do not significantly alter prices. The second advantage of more liquidity is that the SO can more easily reconfigure different types of transmission rights - this aspect will be illustrated in section 3. It can be concluded that obligatory zonal access rights are superior to obligatory zone-to-zone contracts.

3.3 Entry/exit rights should be obligations

A consequence of bundled property rights (entry/exit rights, point-to-point contracts) is that all transmission contracts must be obligations to result in efficient network utilisation. This is not the case with unbundled rights (i.e. flow gate rights). These may be treated as options if they have a positive value, because no transmission constraints will be violated if the option is not executed. Unbundled property rights are only obligations if they have a negative value. If the obligation is not executed, then congestion is not relieved and transmission constraints may be violated. Entry/exit rights are bundled property rights and usually consist of a combination of options and obligations. If the value of the underlying options is greater than that of the underlying obligations, then the bundle has a positive value. It still cannot be treated as an option, because it contains obligations which have to be executed to keep the network stable. Therefore, entry/exit rights must be treated as obligations, even if they have a positive value[12].

If, as currently proposed by Ofgem, entry/exit rights are treated as options, then the SO cannot rely on the congestion relieving contribution of the components ([15], page 79). Whereas financial

[11] One can argue whether the number of contracts in the zone-to-zone design will in reality be smaller than $N*(N-1)/2$. Why should one require transmission contracts for transmissions between zones if the transmission does not contribute to any congestion. However, if prices for transmissions between such zones would be zero, whilst prices from these zones to a third zone differ, any two firms could make use of the arbitrage opportunity and start a profitable triangle trade. Triangle trade between European countries has caused near incidents in Belgium in 1997 and 1999 [16]. As it is difficult to legally inhibit triangle trade, the best solution is to issue and require transmission rights for all links.

[12] Ofgem suggests that the system operator should anticipate what proportion of the options will be executed and possibly create additional entry/exit rights. Such an evaluation seems difficult to perform given the multiple bilateral negotiations for electricity (German electricity trading volume in 2000 was twice the total generation, Günter Marquis, VDEW President, Hamburg 30.5.01)

If the system operator has to anticipate what proportion of property rights will be executed then the purpose of the entire exercise is eliminated: to inform the system operator and market participants of the location and value of scarce capacity.

options are only executed when their price is above zero, transmission options will frequently be executed even if the value of the corresponding option is zero. This has the consequence that the SO cannot infer anything about the usage of the option from a zero price. The SO will issue fewer entry/exit contracts than if he could anticipate the utilisation of the contracts. The result is lower utilisation of the network and therefore welfare losses and more exercise of market power.[13]

A second advantage of firm transmission contracts is that they allow for an increased trading volume. Non-firm transmission rights can only be issued up to the capacity of the network and thereby limit the trading volume. However, high trading volumes do not imply large flows of electric energy. If trades go in the opposite direction, the resulting flows cancel each other. This is mirrored by the simultaneous creation of firm contracts corresponding to transmissions in both directions.

A further advantage of firm transmission rights is that they allow formalising currently existing bilateral contracts between SO and 'must-run' generators located in import constraint zones. The 'must-run' generator obtains obligatory entry-rights, potentially along with a financial payment, and contracts independently for the energy delivery based on his minimum output requirement. The increased transparency should increase the threat of entry at these locations. The increased threat of entry reduces prices [14]. A further advantage could be that increased transparency simplifies detection of abuse of market power.

The disadvantage of firm transmission rights is that companies either have to incur additional transaction costs adapting their transmission rights portfolio to intended transactions or to incur penalty charges for deviations from the permitted flows. This discriminates against small companies, because the fewer transactions a company procures the less likely it is that imbalances will cancel each other. Furthermore, all renewable energy sources, such as wind and solar energy, are at a disadvantage because they are weather-dependent, reducing predictability of future generation. Combined heat and power production (CHP) is also at a disadvantage, as it is usually driven by the heat demand which is a function of weather.

3.4 System operator has to rebundle property rights

In meshed networks, a high number of independent constraints are usually active at some time of the year. If individual property rights were to be assigned to each constraint, the number of required contract types would be excessive. Therefore the SO bundles individual property rights into entry and exit rights.

The disadvantage of bundling property rights is that it can inhibit optimal allocation of unbundled property rights. A theoretical optimal design would therefore require the SO to continually

[13] Borenstein shows that increased interconnection between two zones reduces the prices in both zones due to increased competition[4].

rebundle property rights to eliminate this effect. In the current Ofgem proposal, the SO is empowered to "buy-back access rights from participants and sell additional rights in order to resolve transmission constraints. The SO (National Grid Company) would also be able to sell any rights made available but not purchased in the initial auction."([15],p. 14). Furthermore, "Use it or lose it" provision would be required to prevent the hoarding of access rights which could distort the operation of the market in firm access rights ([15], section 3.2). However, the current proposal does not go far enough to allow the SO to trade entry/exit rights.[14] In appendix B, a example for a three node network illustrates the welfare loss if the SO cannot buy back entry/exit rights, rebundle the underlying rights for scarce capacity and sell different entry/exit rights. The Ofgem proposal could be improved upon by enabling the SO to trade entry/exit rights[15] - but only if the corresponding transaction costs were lower than the efficiency gain.

Not allowing the SO to trade entry/exit rights furthermore discriminates against intermittent energy sources for the following reason. The SO will initially bundle transmission rights in order to match expected flow patterns. If, for example, wind generators are located at the west and east cost, then access rights to links that are needed for both transmissions would be split between both locations and hence scarce and expensive. Wind generators only realise their actual production close to real time and can then sell rights. Entry/exit rights for transmission from the west coast can be bought back and reconfigured to entry/exit rights for the east coast, thereby decreasing balancing costs.

3.5 Conclusion

Property rights are theoretically an optimal way to internalise constraints on transmission into the electricity market. In order to increase liquidity, property rights to individual transmission links can be easily bundled into entry/exit rights. If these rights are freely tradable, and different types can be traded by the SO, then entry/exit rights are an optimal implementation of property rights - because they require the lowest number of contract types. However, the number of required contract types seems to be too high to allow for sufficient liquidity and transaction costs incurred when matching property rights to energy transmissions might outweight any benefits.

The Ofgem proposal falls short of this theoretical optimal implementation of property rights in two respects. First, the current proposal does not allow the SO to trade and reconfigure transmission rights, thus inhibiting efficient utilisation of the network. Second, entry/exit rights with a positive value are considered to be options. However, as entry/exit rights represent a bundle of property rights to scarce transmission capacity, some of the underlying property rights

[14] A first step towards swapping entry/exit rights would be for the system operator to keep back some of the rights as long as possible. In the final allocation, these rights could be bundled such that revenues are maximised and the rights are used most efficiently.

[15] The initial analysis suggests a conflict between economically efficient solutions in meshed networks and the NETA 'spirit' of reducing the role of the system operator in favour of entirely decentralised mechanisms.

might have negative value, implying that the corresponding entry/exit right is an obligation. No analysis of additional transaction costs due to these two improvements is available.

4 Shortcomings of physical property rights

A challenge for any system of locational charges is how small the regions must be. Section 4.1 shows that England and Wales would need to be divided into more than 30 zones to resolve sufficient constraints. However, the more the market is divided, the lower will be liquidity in the market for entry/exit rights - a dilemma illustrated in section 4.2. Low liquidity reduces the efficiency of the market because contractual positions can be matched with demand only approximately and because interactions cannot be automated if individual transactions have a large impact on prices. In the market for transmission rights, missing liquidity has three more particular disadvantages. First, it is common in any meshed network that one transmission constraint is binding. Therefore the SO can no longer reconfigure the system efficiently by buying one type of transmission contract, swapping it to a different type according to the contribution of each type of transmission contract to the constraint, and selling the second type. Instead, the SO has to retain several transmission contracts to allow for an efficient swap. (section 4.3). Second, if the market for entry/exit rights is not liquid enough, then intermittent and small generators cannot adjust the amount of entry/exit rights to their generation, but have to pay penalty fees for the SOs' constraint resolution in the balancing market. Third, low liquidity in the market for physical transmission contracts implies that large transactions can change the price significantly. This inhibits companies from signing bilateral contracts before securing the transmission with physical rights (Section 4.4).[16]

4.1 Zonal vs nodal definition of entry/exit rights

The discussion of nodal vs. zonal pricing repeats the argument in section 2 of explicit treatment of transmission constraints versus socialisation of constraint costs. Zones are defined as a collection of nodes. Within a zone, all nodes are treated equally and entry/exit rights can be used by generators/demand at any node within the zone. The SO has less control over the usage of entry/exit rights. Therefore, he either issues fewer rights to prevent violations of constraints or uses the balancing market to resolve constraints. Costs are socialised. Therefore, companies with market power can once again abuse the system by over-contracting energy and subsequently profiting in the balancing market for adapting generation. The advantage of zonal pricing is that fewer contracts are required, so that liquidity is increased. Some sources falsely assume

[16] A further disadvantage of physical transmission rights is that they can be strategically withheld to exercise market power. Use-it-or-loose-it rules are implemented in auctions for transmission capacity between European countries, e.g. Germany-Netherlands, to prevent such abuse, but require strict monitoring. It is uncertain whether late returning of rights can still influence prices in earlier markets or even prevents re-scheduling of inflexible generation assets. (See [12] for further discussion)

that higher liquidity in the zonal pricing regime reduces market power. However, the network interaction in the nodal pricing regime implies that neighbouring generators compete against each other, but have to pay for constraints they induce.

The transmission SO National Grid Company (NGC) has been critical about the effects of zonal access rights. In a simulation with 12 zones for the English and Welsh market, only half of the total volume of transmission constraints could be resolved prior to Gate Closure. ([15], Appendix 1.30, page 110). Increasing the number to 24 zones, still only 75% of constraints can be resolved, whilst 31 zones allowed the resolution of 88% of constraints. Ofgem concluded that "further consideration needs to be given as to whether some form of nodal definition of rights, along with other mechanisms to facilitate trading, should be employed." ([15], page 113) because "NGC expected its [market based arrangement around firm access rights] effectiveness to be limited if the access rights were zonal rather than nodal." ([15] page 59)

4.2 Liquidity

In the last section, we quoted NGC studies that 31 zones would be required to cover 88% of transmission constraints. This implies that at least 31 different property rights must be defined. As demand patterns change fast, most electricity markets are oriented on intervals of half an hour or shorter. Defining 31 property rights for every half hour of the year requires 543 000 different contracts. The fundamental criticism of physical property rights is that there will be not sufficient liquidity for continuous trade of 543 000 contracts - failing the Coase requirement for trading of externalities to occur.

Ofgem presumably realised that such a multitude of property rights might be difficult to deal with. To relieve the SO of any difficulties, Ofgem therefore proposed combining the initial auction of property rights for several periods, for example, access rights in one zone at peak times. Ofgem suggests that the property rights might be subsequently traded separately on the market. However, if Ofgem suggests that a central SO incurs difficulties with such a multitude of property rights, it is difficult to believe that decentralised markets with even lower liquidity could facilitate efficient trade. Even the optimally defined physical transmission contracts result in illiquid markets and subsequently inefficient allocation of scarce transmission capacity.

Physical transmission rights function at well defined bottlenecks between markets, i.e. England and France. Residual demand then can always be satisfied within the market. Property rights for transmission between the markets therefore do not have to be adjusted very often and can, furthermore, be aggregated over several hours. Furthermore, if property rights are defined for one bottleneck, as between England and France, then they do not need to be reconfigured. Property rights to such an interconnection allow for efficient usage and capture the transmission rent. However, note that this is a special case of one-line interconnection with nicely one way flows.

4.3 Efficient rebundling requires SO to retain transmission rights

The four-node network in figure 3 illustrates why it does not suffice if the SO buys one entry right in order to subsequently issue a second type. Efficient utilisation of the network is only feasible if the SO has access to several types of transmission rights simultaneously. Assume a lossless DC

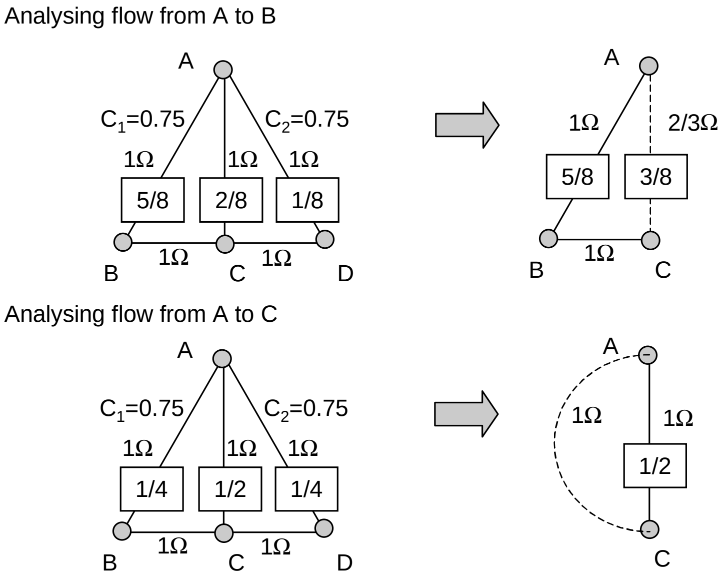


Figure 3: Four-node network with two capacity constraint links

network where all links have identical resistance and two links between node A and B and node A and D are capacity constraint. The other links are assumed to be unconstrained. Node A is a withdrawal node and nodes B, C and D are injection nodes. It suffices to define entry rights for nodes B,C and D because the amount of energy withdrawn at node A follows from the amount of energy injected at the other nodes. Due to loop-flows energy transmitted from node B to A does not only use the direct link, but part of the energy is transmitted on alternative routes via node C and D. Ohm's law allows us to calculate the proportion of the energy transmitted on different routes (See appendix B for a more detailed description). For example, for energy inserted at node B and withdrawn at node A, 5/8 is transmitted via the direct link, 2/8 transmitted via the link between C and A and 1/8 transmitted via the link between D and A.

For its internal calculation of the SO defines flow gate rights for the constraint links $B - A$ and $D - A$.

An entry right to insert one MW at node B corresponds to a transmission of one MW from node B to A. The SO associates 5/8 MW of flow gate rights to link $B - A$ and 1/8 MW of flow gate rights to link $D - A$.

Assume the SO initially assigns the entire transmission capacity to create entry rights for nodes B and D. A market participant requests an entry right for node C, but one MW of the right increases flows on both constrained links $B - A$ and $D - A$ by 1/4 MW, thus violating the transmission constraints. Before issuing entry rights to node C the SO therefore has to buy back some entry rights to nodes B and/or D.

If we assume that the market for entry rights is illiquid, then the SO might only be able to obtain one type of rights, lets assume one MW of entry rights at node B. In the internal calculation of the SO this rights corresponds to 5/8 MW of flow gate rights on link $B - A$ and 1/8 MW on link $D - A$. The SO subsequently sells 1/2 MW of entry rights to node C, increasing the load on both links $B - A$ and $D - A$ by 1/8 MW. The constraint on link $D - A$ is binding again, restraining the SO from issuing more rights, whilst the constraint on link $B - A$ is slack.

The alternative approach, which the SO might pursue if there is a liquid market in access rights, is to simultaneously buy entry rights at B and D. In exchange for half a unit of entry right B and half a unit of entry right D the SO can sell 3/2 units of entry right C.

The second approach is more efficient. In order to facilitate exports of one MW from node C, in the first approach combined exports from node B and D are reduced by 2 MW, whereas in the second approach the combined exports are reduced by 2/3 MW.

In either case only the SO can reconfigure entry rights, because he has to ensure that they correspond to the underlying structure of flow gate rights. If the market is not liquid, the resulting allocation might not be efficient; if the market is liquid, the resulting allocation may be efficient. In reality the number of binding constraints is significantly higher, implying that the SO would have to simultaneously trade a large number of entry rights to ensure an efficient allocation.

4.4 Interaction with energy market

So far, little research has been done on the question of interaction between the energy market and the transmission rights market. Companies will be reluctant to sign contracts for delivery of energy before they know the price of the corresponding transmission rights. Illiquid markets for transmission rights imply that any transaction can have significant impact on the price.

This is illustrated by the English and Welsh situation. If the market is divided into 31 zones, any zone represents, on average, 1200 MW entry/exit rights. Contracting for the output of a nuclear generator with, for example, 700 MW can change the scarcity of the corresponding transmission rights and would, in that case, change the price of transmission rights. The price itself does not provide information about the elasticity of demand and supply which the company will be facing when trying to obtain 700 MW worth of contracts.

The situation is slightly improved if the transmission operator can trade transmission contracts thus increasing the total amount of available transmission contracts.

Given the problems with the transmission access markets, one last point seems of interest. Entry/exit rights have an advantage over node-to-node transmission contracts in that generators can obtain entry/exit rights for the location of their generation, and the consumers can obtain entry/exit rights for their withdrawal point, before ever contracting with each other. However, such a behaviour would imply that all companies obtain entry/exit rights before ever contracting in the energy market. Not all companies can do so - because the amount of available entry/exit rights is not sufficient. Only some of the generators would obtain entry/exit rights and can subsequently be expected to sell the corresponding energy. Due to missing liquidity, this initial allocation would prevail and therefore determine which generators will sell and generate electric energy.

4.5 Conclusion

Physical property rights for meshed electricity networks connecting illiquid zones result in inefficient dispatch, high transaction costs and discrimination against small and intermittent demand and supply side. Dispatch is inefficient because illiquid markets for property rights prevent the arbitrage required for optimisation. Transaction costs are high because of the huge number of contracts to be traded continually. Small parties are discriminated against because big parties can net of different residual demands, reducing the need for adjustment of transmission contracts or payment of balancing fees.[17] Intermittent parties are discriminated against because the illiquid market does not provide the flexibility to adjust transmission positions fast enough to changing weather conditions. Coase promises that property rights are an efficient solution to the problem of externalities. The complexity of the electricity network and the need for continual balance of generation, demand and transmission constraints violate the conditions required for Coase's theorem.

5 Locational congestion charges with financial hedging

Section 2 illustrated the potential need for active constraint management, whilst section 4 showed that physical property rights are an inappropriate solution. Instead, I suggest building on the current approach of using the SO to resolve constraints. Differing from the current mechanism the costs are to be attributed to parties contributing to violations of transmission constraints. Parties that have scheduled transmissions that relieve constraints receive payments. To allow for hedging against future congestion charges the SO issues Tradable Congestion Contracts (TCC). These contracts are proposed by [10] and have been successfully implemented in Pennsylvania, New Jersey and Maryland (PJM system) since 1999.[18]. TCC were originally proposed for systems with nodal pricing and can be equally applied to congestion charges determined in the balancing market.

[17] Netting off entry and exit imbalances shall be prohibited, "since it could unduly discriminate in favour of vertically integrated participants." ([15]). Balancing between different generators of one company or customers of one supplier is allowed.

[18] In the PJM system, TCC are referred to as Firm Transmission Contracts(FTC). See www.pjm.com

Transmission contracts can mitigate or amplify market power, as shown by Joskow and Tirol [12]. Entry/exit rights should be defined such that they minimise the amplification of market power without loosing arbitrage opportunities or chances to mitigate market power (see Gilbert/Neuhoff/Neuhoff mimeo, DAE). Therefore financial entry/exit might be optimal from the perspective of alleviating market power.

5.1 What are location specific energy charges?

Location-specific energy charges can be best understood in comparison to physical transmission contracts and nodal pricing. In a theoretical model without uncertainty, illiquidity, market power and transmission losses, all three approaches can be used to achieve welfare-maximising generation and demand scheduling. All three schedules will be identical.[19]

The first approach, physical transmission contracts, requires the SO to issue physical transmission contracts corresponding to the amount of available capacity on each link. Coase's theorem states that the resulting market equilibrium, is a welfare maximum. Problems with bundling of property rights, and the combination of illiquidity of the market with uncertainty about future schedules inhibit the efficient functioning of the market for property rights. Therefore, the following two approaches do not create an explicit market for property rights, but assume that the SO simulates the market.

The second approach, nodal pricing, requires a central dispatching SO. The SO collects bids from all generators and consumers. Based on these bids, the SO simulates a market with physical transmission contracts.[2][20] The resulting equilibrium prices for all nodes are published. Generators who bid lower than the equilibrium price at their node will be scheduled and receive the equilibrium price; the same goes for demand-side bids that are higher than the equilibrium price at the given node. The PJM system, which has about the same load as England and Wales, has successfully implemented nodal pricing.

Location specific congestion charging can be seen as an intermediate approach. Generators trade bilaterally and inform the SO about planned transmissions. The SO asks generator and consumers for bids in the balancing market. Based on these bids, the SO simulates a market with physical transmission contracts. He assumes outstanding transmission contracts corresponding to the transmissions he has been informed about. If constraints are violated because the net amount of contracts issued for one link exceeds the link capacity, then the SO auctions contracts for flows in the opposite direction to ease the constraint. Market clearing prices in the auction may be positive or negative. Subsequently, the SO announces the simulated balancing market clearing prices for all nodes. Bids for additional energy are accepted, if the bid price is below

[19] Chao and Peck used Coase to show the identity between nodal pricing and complete physical transmission contracts [5]. Locational-specific energy charges are a different approach to nodal pricing.

[20] Chao shows that the optimal scheduling can be achieved in a market given the convexity existing in a simplified physical network [5] and Fang provides an algorithm to determine the solution [7].

the market clearing price at a node. The bidder is paid the market clearing price. The same procedure applies to bids for energy withdrawal. The balancing market clearing prices at a node is used to determine the entry/exit charge. In section 3.2, I argued that physical entry/exit rights are required for all zones but the reference zone. The same argument can be used to show that entry/exit congestion charges are required for all nodes apart from the reference node. The entry charge equals the balancing market clearing price at the reference node minus the entry node. The exit charge is the balancing market clearing price at the exit node minus the price at the reference node.

Physical transmission contracts differ from nodal pricing in whether the contracts are explicitly traded or whether the trade is simulated by the SO. A second difference is that physical transmission contracts are complemented by bilateral energy trade, which is performed as pay-as-bid, whereas nodal pricing is priced at market clearing price. Both are identical in a world without uncertainty and with full information, because all companies bid the market clearing price.

Location specific energy charges combine both approaches. The SO simulates the market for transmission rights based on the announced schedule and resolves constraints in the balancing market at the market clearing price. Therefore bilateral trades are possible. As all market participants are exposed to locational charges they face incentives to minimise their costs by participating in the balancing market to reduce congestion. This should increase liquidity in the balancing market. This will introduce nodal pricing through the backdor while allowing for existing and future bilateral energy contracts.[21]

5.2 How to hedge for uncertainty in congestion charges

The previous section illustrated how location-specific energy charges are calculated. The charges are determined in the balancing market, which is performed a day ahead and iterated closer to dispatch. Uncertainty about future generation and demand schedule implies that future congestion charges cannot be predicted with certainty, either. To hedge against the financial risk, tradable financial congestion contracts may be used.

Analogous to the previous described entry/exit rights, the SO would auction financial entry/exit rights. Firms obtain an entry/exit right for zone A, corresponding to X units energy transmission from zone A to the reference zone or, alternatively interpreted, $-X$ units energy transmission from the reference zone to zone A. The transmission contract will pay the congestion charge which would be levied or paid out for such a transmission, irrespective of whether the owner actually performs the transmission. The market value of financial entry/exit rights corresponds to the expected congestion charge. Therefore, financial entry/exit rights only serve to hedge against uncertain constraint costs.

[21] In PJM only 10% to 15% of the energy actually is scheduled through the security constrained dispatch, based day-ahead, and real time markets used to set locational prices. The remaining energy is traded on longer term (bilateral) contracts.

The advantage of financial entry/exit rights over physical entry/exit rights is that they can be aggregated over time. For example, in the PJM financial transmission contracts are only differentiated in peak and one off-peak times.[22] Generation and transmission schedules of companies will always deviate slightly from the financial contract. However, companies are not punished for these deviations. They only face the risk of unexpected changes in congestion charges, either for transmissions they perform without financial entry/exit rights, or for financial entry/exit rights they own without performing the corresponding transmission. The advantage of financial entry/exit rights is that the amount of different transmission contracts to be traded would be reduced from 543.000 to 372 per year[23], assuming England and Wales are divided into 31 zones.

5.3 Financial entry/exit reduce balancing interaction

One might assume that a design based on tradable congestion contracts requires more contracting in the balancing market than physical transmission contracts, increasing costs as generation output must be changed more frequently and by less flexible generators. However, the price of tradable congestion contracts provides information about the congestion situation, which market participants already incorporate into their trade decisions. Thus, in a world without uncertainty and with market power, the SO would not have to interact at all in the balancing market. In the following argument, we ignore transmission charges for losses:

Companies develop expectations about future congestion levels, which are expressed in the price of financial entry/exit rights. Companies have an incentive to buy the amount of financial entry/exit rights corresponding to the transmission they perform. If company A under-contracts and obtains fewer entry/exit rights than required for a transmission, then company B can schedule an additional transmission with the remaining entry/exit rights. More transmissions are scheduled than financial entry/exit rights are issued, therefore the constraint will be violated. Marginal costs of constraint resolve are weakly monotonically increasing; therefore, resolving the constraint induces higher marginal costs than the expected costs of constraint resolution expressed in the financial entry/exit rights. Thus, entry/exit congestion charges will be above the price of the financial entry/exit contracts. Firm A therefore loses from under-contracting. The same argument applies if company A over-contracts instead of matching contracts with its transmission plan. The price of financial entry/exit rights will be higher than the marginal costs for subsequently removing a marginal violation.

Firms only schedule transmissions if the energy price difference between origin and destination is higher than the costs of obtaining the financial entry/exit rights. Complete markets and no uncertainty imply that the price of the financial entry/exit rights exactly matches the subsequent entry/exit congestion charges. As no unprofitable energy trades were performed, the financial entry/exit rights, and therefore equally the entry/exit congestion charges for all transmissions,

[22] http://www.pjm.com/energy/ftr/downloads/ftraucug.pdf

[23] Always assuming 31 zones, but reducing the amount of contracts from 48*365 to 2 (peak/off peak) times 12 months, assuming that weekends can be subsumed under off peak times.

are never above the price difference between origin and destination. Therefore, no bids of companies in the balancing market will be accepted and therefore the SO will not reschedule any transmission. Put it the other way around: if the SO were to reschedule in the balancing market, then market participants must have had incorrect expectations about entry/exit congestion charges, which is not possible in a world of perfect information, complete markets and no market power.

In reality, we face uncertainty. Furthermore, bundling of financial entry/exit rights over several time periods fails the complete markets assumption. Therefore, the SO will need to interact in the balancing market to correct for errors due to both effects. However, the amount of constraint resolution required in the balancing market is low as a result of the information provided in the financial market.

5.4 Market power considerations

A currently discussed topic is the effect of transmission contracts on the exercise of market power by electricity generators. While Stoft [18] titles an article "How financial transmission contracts curb market power [in a two node network]" Joskow and Tirol [12] give examples for situations when holding of transmission contracts enhance market power as well as for situations when contracts mitigate market power. Gilbert, Neuhoff and Newbery (DAE mimeo, March 2002) show that the allocation process of transmission contracts, irrespective of whether physical or financial, is crucial. In the two-node case if the generators with market power are located at an import constrained node, it is always undesirable to allow them to retain entry rights for the exporting node. It is also undesirable to allow them participate in the auction for entry rights at the exporting node where this is pay-bid, or where there is likely to be asymmetric information favouring the generators. Even under the ideal circumstances of perfectly informed arbitrageurs and a single-price auction, while there is no harm in allowing generators to bid, there is also no benefit. Nor do the generators benefit from bidding even in the pay-bid case, as all the additional distortionary revenue is secured by the transmission company. On the other hand, if the generators with market power are located at an export constrained node, their bids for entry rights allow them to precommit to additional output and reduce prices at the exporting node, to the benefit of consumers there.

In a three-node network (and with more nodes), transmission entry rights can enhance market power even when they correspond to own production, if loop-flows imply that output changes have a bigger impact on the price at the reference node than at the entry node. The implication of this analysis is, that the reference node for entry/exit rights (financial or physical) matters, and should be chosen such that it has the price least influenced by any generator's output decision. Generators should then be restricted to entry rights corresponding to their energy production. Consumers and their representatives would buy exit rights. Such a policy minimises the risk that transmission contracts enhance market power while ensuring that they provide risk hedging services and provide information for network expansion.

5.5 Conclusion

Location-specific congestion charges hedged with financial entry/exit solve the problems of implicit congestion treatment, namely inefficient dispatch, incorrect signals for location of new generation facilities and possibly enhancement of locational market power. Location-specific energy charges allow the SO to collect the same transmission rent as the auction of physical transmission contracts. In order to minimise their exposure to congestion costs market participants will participate in the balancing market to allow resolution of transmission constraints at least costs. This should allow for the minimal cost constraint resolution and will in a long term approach a nodal-pricing design. In order to minimise the risk that transmission contracts enhance market power, the financial entry/exit rights should be defined relative to the node which has a price least influenced by any generator's output decision.

The following problems of physical transmission contracts are resolved:

The amount of required contract types is reduced from 543.000 to 372 per year if the same zonal resolution is maintained. This allows for liquid markets. Liquid markets reduce transaction costs, allow for efficient arbitrage and remove the possibility of vertically-integrated or large generators profiting from aggregation. Not only is the liquidity of financial entry/exit rights higher than that of physical entry/exit rights, but less liquidity is required, because companies do not need to continuously adjust their energy positions with entry/exit rights. Whereas under physical contracts, companies have to pay fines for imbalance between energy transmission and physical contracts, under financial contracts, companies only face risk for imbalance between transmitted energy and financial entry/exit rights.

Physical transmission contracts discriminate in favour of large companies. Large companies can aggregate imbalances of their generation and demand-contracts such that they partly cancel. As a result of this aggregation, large companies have fewer interactions in the imbalance market, and therefore pay fewer fines for imbalances. Entry/exit congestion charges are symmetric such that the entry charge at a location is exactly the negative of the exit charge. The only advantage of aggregation in the presence of financial contracts is that risk is reduced. Therefore, financial entry/exit rights discriminate less than physical transmission contracts.

Physical transmission contracts have so far, to my understanding, not been implemented successfully in highly meshed networks whereas some financial transmission contracts based on nodal pricing have been successfully implemented in the PJM since 1999.

References

[1] S. R. Backerman, S. J. Rassenti, and V. L. Smith. Efficiency and income shares in high demand energy networks: Who receives the congestion rents when a line is constrained? Economic Science Laboratory, University of Arizona, 1997.

[2] R. E. Bohn, M. C. Caramanis, and F. C. Schweppe. Optimal pricing in electrical networks over space and time. *Rand Journal of Economics*, 15, No 3:360–376, 1984.

[3] S. Borenstein, J. Bushnell, and C. R. Knittel. Market power in electricity markets: Beyond concentration measures. *The Energy Journal*, 20, No. 4:65–88, 1999.

[4] S. Borenstein, J. Bushnell, and S. Stoft. The competitive effects of transmission capacity in a deregulated electricity industry. *RAND Journal of Economics*, 31, No. 2:294–325, 2000.

[5] H.-P. Chao and S. Peck. A market mechanism for electric power transmission. *Journal of Regulatory Economics*, 10(1):25–60, 1996.

[6] R. Coase. The problem of social cost. *The Journal of Law and Economics*, 1:1–44, 1960.

[7] R. Fang and A. David. Transmission congestion management in an electricity market. *IEEE Transactions on Power Systems*, 14(3):877–883, 1999.

[8] R. Green. Electricity transmission pricing: An international comparison. DAE, July 1997.

[9] S. M. Harvey and W. W. Hogan. Nodal and zonal congestion management and the exercise of market power. Harvard University, 2000.

[10] W. W. Hogan. Contract networks for electric power transmission. *Journal of Regulatory Economics*, (4):211–242, 1992.

[11] W. W. Hogan. Flowgate rights and wrongs. Center for Business and Government, Harvard, 2000.

[12] P. L. Joskow and J. Tirole. Transmission rights and market power on electric power networks. *RAND Journal of Economics*, 31 No. 3:450–487, 2000.

[13] D. Newbery. *Privatisation, Restructuring and Regulation of Network Utilities*. MIT Press, 2000.

[14] D. M. Newbery. Competition, contracts, and entry in the electricity spot market. *RAND Journal of Economics*, 29 (4):726–749, 1998.

[15] O. of Gas and E. Markets. Transmission access and losses under neta, a consultation document, May 2001.

[16] E. T. S. Operators. Evaluation of congestion management methods for cross-border transmission. Florence Regulators Meeting, 1999.

[17] L. E. Ruff. Flowgates, contingency-constrained dispatch, and transmission rights. *The Electricity Journal*, January/February:34–55, 2001.

[18] S. Stoft. Financial transmission rights meet cournot: How tccs curb market power. *Energy Journal*, 20(1), 1999.

A Comparison of firm point-to-point and entry/exit rights

In DC approximation, a constrained electricity network with point-to-point transmission contracts can be presented as follows. C_m gives the maximum amount of energy that can be transmitted on a link m. The proportionality factors $\alpha_{i,j}^m$ summarise the resistance of links and describe the proportion of the flow from node i to node j, passing over link m. The net amount of firm point-to-point contracts which are used for transmissions from node i to node j is $I_{i,j}$. The following equation describes constraints on the amount of transmission contracts which can be used:

$$-C_m \leq \sum_{i,j,i<j} \alpha_{i,j}^m * I_{i,j} \leq C_m \tag{1}$$

In a similar approximation for nodal entry rights, we use the same transmission constraints C_m.

The proportionality factors β_i^m are based on the resistance of all links in the network. They describe how net inflows F_i in node i influence the load on constrained link m.

$$-C_m \leq \sum_i \beta_i^m * F_i \leq C_m \tag{2}$$

Assuming zero transmission losses, the net amount of electric energy inserted into the transmission system is zero ($\sum_i F_i = 0$). One remaining net inflow F_i can be deduced if all other inflows are known (F_j $\forall j \neq i$). It would suffice to define nodal access rights for all but one node. We nevertheless propose to define nodal access rights for all nodes in order to treat all nodes equally. This is advantageous when implementing access right pricing in a federal environment and as an accounting device to determine whether all transactions have been closed. The redundant access right creates one degree of freedom allowing for a shifting overall price level for access rights. The price level can be fixed by setting the price for access rights to one arbitrarily-chosen reference node. A change of the reference node or price level changes the prices of all contracts, but the total price of a pair of contracts representing one energy transmission will stay constant, being the difference between the two entry/exit prices.

We first prove that point-to-point transmission rights represent the same underlying physical structure as nodal access rights. In figure 4, three transmissions of identical energy are assumed, from node i to node j, from node j to node k and from node k to node i. The flows entering and leaving each node cancel each other such that no electric energy is transmitted in the system and the net flows on all links are zero. This implies that load flows created by a transmission from node i to node j would be nullified by two identical transmissions from node j to k and k to i. The load flows resulting from the transmission from node i to j could also have been nullified by a identical transmission from node j to i. It can be concluded that a transmission from node j to i has the same effect on the network as two identical transmissions from node j to k and node k to i.

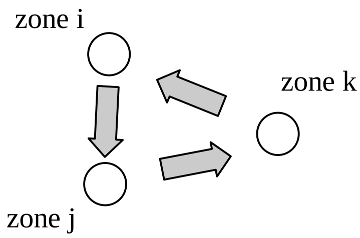


Figure 4: Three identical flows result in net zero flow in the network. A transmission from zone j to i has the same effect on the network as two identical transmissions from zone j to k and zone k to i.

From the previous argument follows that any transmission $I_{i,j}$ can be replaced by two transmissions $I_{i,k}$ and $I_{k,j}$. Adding together transmissions which originated in i and subtracting transmissions which end in i, we obtain the net nodal access in node i:

$$F_i = \sum_{j \neq i} I_{i,j} - \sum_{j \neq i} I_{j,i} \tag{3}$$

This proves that firm point-to-point contracts can be expressed as nodal access contracts.

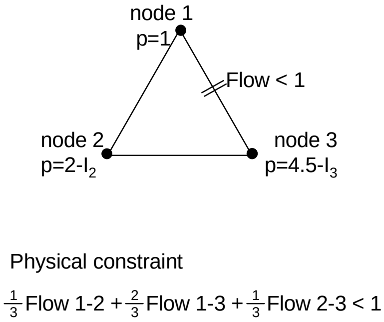
In a second, step nodal access rights are expressed as point-to-point contracts. For a energy transmission, access rights to insert energy into node j and access rights to withdraw energy from node i are obtained. The same flows would be observed if a transmission from node i to k and a transmission from node k to j were superimposed. Therefore, the SO could, instead of issuing access rights to node i and j, issue point-to-point contracts from node i to k and from node k to j. If we define an arbitrary reference node k, then nodal access rights to a node i correspond to a point-to-point contract from node i to k. Nodal access rights in the notation of point-to-point contracts are given in equation 4.

$$I_{i,k} = F_i \quad \forall i \neq j \qquad\qquad I_{i,j} = 0 \quad \forall i,j \;\; j \neq k \tag{4}$$

Each approach subsumes the other, therefore, nodal access rights and point-to-point contracts represent the same physical constraints. This observation contrasts with the usual economic perspective that transport costs are related to the route chosen for the transport.

B Trading SO vs. restraint SO

The three-node network in figure 5 illustrates that a trading SO can use the system more efficiently. In the initial auction of transmission rights, demand at nodes two and three is unknown; therefore, the transmission operator auctions one unit of transmission rights from generation at node one to demand at node two and three. In the example, demand at node three turns out to be higher than demand at node two. Market participants cannot recombine existing transmission rights to react to this situation. The binding transmission constraint prevents the SO



	Price			Net import			Δ Welfare		
Node	1	2	3	1	2	3	1	2	3
Initial Assignment	1	1	3.5	-2	1	1			
After SO swapped	1	2	3	-1.5	0	1.5	0	-.5	.625

	Price			Quantity			Rev
Contract	1-2	1-3	2-3	1-2	1-3	2-3	Tot
Initial Assignment	0	2.5	2.5	1	1	0	2.5
After SO swapped	1	2	1	.5	1	.5	3

Figure 5: Three node network with one transmission constraint, generation with constant costs at node one and demand at node two and three. The initial assignment of transmission rights can only be improved upon if the system operator can trade transmission rights.

from issuing additional rights for transmission to node three unless existing rights are returned. However, market participants will not return transmission rights from node one to node two, because they know that rights are associated with a positive value of one after reconfiguration. The situation can be resolved, if the SO buys rights for transmission from node one to node two at the expected price of rights for transmissions from node two to node three. Market participants will accept the deal and the SO makes no losses when subsequently selling the reconfigured rights. In the numerical example, consumer gains are increased by 0.125 units and 0.5 units of additional transmission revenue are collected, so total welfare change is $+0.625$.