MARKET-BASED INVESTMENT IN ELECTRICITY TRANSMISSION NETWORKS: CONTROLLABLE FLOW

Gert Brunekreeft
Department of Applied Economics
University of Cambridge,
Sidgwick Avenue
CB3 9DE Cambridge UK
g.brunekreeft@econ.cam.ac.uk
http://www.econ.cam.ac.uk/dae/index.htm

28 July 2003

Abstract: This paper discusses unregulated market-based electricity transmission investment by third parties as opposed to regulated investment by designated transmission system operators. Market-based transmission investment derives revenues from price differences between the two ends of the line. The paper focuses on four main areas of potential problems: economies of scale, market power, detrimental investment and risks. The analysis suggests to restrict market-based investment to controllable flow (DC or FACTS) only. This is in line with what seems to settle in practice, strikes a balance between pros and cons of market based investment and draws a sharp line between regulated and unregulated investments.

Keywords: electricity, transmission, merchant, investment

JEL classification: L1, L43, L94.

1. Introduction

The investment level of high-voltage electricity transmission capacity, in particular for the interconnectors between different networks, is concerning low. For instance, the North American Electric Reliability Council [NERC, 2002] estimates a mere 5% capacity growth until 2011. Underinvestment may have a number of reasons, among which in particular regulatory uncertainty and vertical integration of transmission and generation. A recent policy is to rely on the incentives of unregulated third-party investors to resolve the interconnector constraints. This is known as market-based transmission investment (also known as merchant transmission investment). In the USA, the federal electricity regulator, FERC, currently discusses merchant investment in the proposal for Standard Market Design, which states that:2 “LMP and Congestion Revenue Rights will provide price signals to indicate where new investment is needed; however, the price signals alone may not guarantee sufficient investment. We also propose to require a regional transmission planning and expansion process to provide a backstop process for ensuring that needed transmission construction is undertaken.” FERC relies on a dual system: merchant investment in first instance and presumably the independent transmission provider as the investor in last resort. The current proposal for an EU regulation on cross-border exchange [EC, 2003, art. 7] allows an

1 The author would like to thank Katja Keller, Karsten Neuhoff, David Newbery and Hans-Joerg Weiss for useful comments.
2 SMD Notice of proposed rulemaking 2002, p. 66, FERC, Docket No. RM01-12-000.
exemption from regulation for merchant investors provided some conditions are met. Meanwhile, merchant transmission projects are in operation in Australia and planned in the USA and in Europe.

Whereas market-based transmission investment may mitigate the problem of under-investment, it is unlikely to suffice alone and thus regulated projects by the designated transmission system operator (TSO) remain necessary. The inevitable mix of regulated and unregulated systems requires a sharp distinction. This paper will explore the prospects of restricting market-based transmission investment to controllable flow, as defined as direct current (DC) and flexible alternating current transmission systems (FACTS) only. Controllable means that the flow on a line can be controlled directly, rather than being determined by Kirchhof’s laws as in a meshed alternating current (AC) system; as a result, the loopflow effects are substantially reduced. There are two reasons for drawing this line. First, the distinction between controllable versus non-controllable flow is sharp and workable. Second, the inefficiencies of market driven decentralized investment in controllable-flow lines are far less than in meshed AC networks and may well be offset by the advantages of merchant projects.

Section 2 discusses the background and the literature, while section 3 summarises the principles underlying market-based transmission investment. Section 4 is the core of the paper and discusses four main areas of problems with market-based transmission investment and their relative severity for controllable versus non-controllable flow. In all, this paper will argue that it may be advantageous to allow market-based transmission investment for controllable flow and reserve the right to invest in systems with non-controllable flow to the designated transmission system operator. The distinction appears to settle in practice and is for instance requirements in EU legislation [EC, 2003, art. 7], is part of the Australian “safe harbours” [cf. ACCC, 2001, pp.126 ff.] and the US control area PJM recently received the federal regulator’s permission to broaden the scope beyond DC lines to stand-alone AC lines and AC network upgrades [FERC, 2003]. Section 5 concludes.

2. Background

The critical step of market-based transmission investment is that investment in the transmission grid no longer is the exclusive and statutory right of the designated (and as a rule, regulated) transmission system operator (TSO). Market-based transmission investment

---

3 Loopflows are explained in section 3. For a technical analysis of controllable flow, the interested reader may be referred to for instance Gyugi [1999] and Arrillaga [1998].

4 Art. 7 basically restricts merchant investment to DC lines but allows exceptions for stand-alone AC lines if the costs of DC lines would be very high relative to AC.
can be defined as transmission investment [ACCC, 2001, p. 126] “operating between two connection points assigned to different regional reference nodes, [...] supported by the revenue stream generated by trading electricity between the two interconnected regions, [and] not eligible to earn regulated revenue.” This definition applies in more or less the same terms for the US and Europe. Table 1 classifies different types of market-driven investment.

<table>
<thead>
<tr>
<th>Best option test</th>
<th>Eligible to Regulated Revenues</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>regulated</td>
</tr>
<tr>
<td>NO</td>
<td>hybrid / tender</td>
</tr>
<tr>
<td></td>
<td>(conversion)</td>
</tr>
<tr>
<td></td>
<td>unregulated</td>
</tr>
</tbody>
</table>

*Table 1: Options for market-based transmission investment.*

Table 1 stresses two points. First, market based investment is in principle unregulated, which implies two sides of the same coin. On the one hand the revenues derived from trading between the two electricity markets are not capped by regulation, while on the other hand the investor is not eligible to the pool of regulated revenues from the network connection charges. The idea reflects link-based interconnection of two different systems, where the line is financed from arbitrage profits between the two spot prices in each system.\(^5\) Second, the authorities may require to pass a best-option test to show that the investment is economically useful. This may or may not be combined with regulation of revenues. A commission might organise a tender for construction and operating the line, while the revenues are the line rentals; the criterion would be capacity size. In Australia a best-option test has been introduced with the regulatory test. The test requires that “a *new interconnector or an augmentation option* satisfies this test if it maximises the *net present value* of the *market benefit* having regard to a number of alternative projects, timings and market developments scenarios; ...” [ACCC, 2003]. This is mainly meant to be for those cases where a third party requests to be eligible for the regulated revenues; in those cases, a commission will have to check whether the investment is actually an investment worth to be built and subsequently rewarded. The remainder of this paper examines the unregulated option: the investor decided on timing, location and size of the investment while the unregulated revenues are derived from spot market arbitrage.

Why allow unregulated third-party transmission investors in the first place? After all, transmission is considered to be the domain of monopolies where regulated, designated operators are dominant. Three reasons are convincing. First, vertically integrated (generator

---

\(^5\) This formulation applies in particular to Australia and Europe, where zonal pricing is the predominant pricing scheme. In the USA, nodal pricing allows for a more sophisticated approach which will be explained further below in sections 3 and 4. This paper will focus on the simple link-based representation.
and transmission) incumbents may have poor incentives to invest in interconnectors. Transmission constraints tend to isolate parts of the networks and thereby increase generation market power within the isolated area. Vertical integration still forms an obstacle to a competitive playing field in parts of the USA [cf. Joskow, 2003, p. 13] and Europe [EC, 2002]. The second reason, which is actually more general than transmission investment, follows from a regulatory problem with risky significant new investment. The problem has been extensively discussed in Australia [cf. Gans & King, 2003]. Suppose that the investment has to made under uncertainty about the ex-post state of the world which is either good or bad. Suppose that the rate of return of the risky investment in case of a bad state of the world is 6%, while 14% if good. If both states have equal probability the risk-equivalent expected return thus is 10%. The argument is that a regulator will do nothing if the state of the world is bad, while the regulator will be tempted to strengthen rate regulation if the world turns out to be good. Assume that the regulator may ex post reduce the rate of return in the good state to 10%. Anticipating this, the expected rate of return is 8% rather than the required 10%. It is straightforward that this may lead to underinvestment or even abolishment of the project. The underlying argument is that a regulator cannot credibly commit to refrain from intervening ex post in the good state. It is argued that credibility to refrain from intervening is increased by not regulating the new investment at all (for a predetermined number of years): a “regulation holiday”. A third reason relies on a public-choice argument. Interconnecting a low priced area with a high priced area will normally imply that the electricity price in the low priced area increases, meaning the consumers in this area actually lose form the new line. If authorities of both sides of the new line have to give permission for the new line, the authorities on the losing sides may hesitate to go along. This problem may be mitigated if permission (on economic grounds) is not required, which would be the case under market-based investment.

The literature on merchant investment is divided. Littlechild [2003] points to the drawbacks of regulation and expresses quite strong sympathy for market-based investment, relying on Australian experience with the regulatory alternative. The point is illustrative. Murraylink was a genuine unregulated market-based transmission investor interconnecting Victoria and South Australia (SA). While Murraylink was being constructed, another project called SNI requested access to the regulated connection charges and passed the regulatory test in December 2001. SNI connects New South Wales (NSW) and South Australia (SA), which is largely parallel to Murraylink and hence competitive. In the regulatory test two options were considered. First, the bundled SNI, building the line plus some upgrading of especially the

---

6 As a result, a vertically integrated firm faces a trade off: increased interconnector capacity enhances trading opportunities but also increases competition from other areas. Which effect dominates is an empirical matter.

7 The connection points for SNI in NSW and Murraylink in Victoria are close.
grid in NSW. Second, the unbundled SNI which meant only upgrading especially the NSW grid, without building the line itself. Cost-benefit analysis revealed that the unbundled SNI had a substantially higher value to society. Nevertheless, the bundled SNI was approved, based on the argument that the unbundled SNI was not commercially feasible faced with risk of stranded assets: upgrading the grid without building the line would leave the investment exposed to the market power of Murraylink. Without commercial feasibility, the unbundled SNI could not be considered to be an alternative and thus the bundled SNI was approved instead. The arguments in the case centred around the question of the degree of market power of Murraylink with respect to the assets of unbundled SNI. As a result of the permission to build the bundled SNI, Murraylink expected its unregulated line-based revenues to fall and requested for conversion to a regulated operator, which was recently approved. The case has been quite controversial and leads Littlechild [2003, p. 28] to conclude that: “an implication of the Australian experience to date is that there may be more danger of excessive than thwarted regulatory investment. Even with reform, merchant transmission could remain vulnerable.” Although it depends on details, the case illustrates how regulated projects can crowd out unregulated projects.

Much on the same line as Littlechild [2003] is Hogan [2003] who argues, although with reservations, in favour of merchant transmission investment. More precisely, Hogan argues to draw a clear line between regulated and merchant investment, to avoid the ‘slippery slope’ that the regulated options crowd out the merchant options. Hogan’s [2003, pp. 22/23] approach is that: “regulated transmission investment would be limited to those cases where the investment is inherently large relative to the size of the relevant market and inherently lumpy in a sense that the only reasonable implementation would be a single project like a tunnel under a river. [...] Everything else would be left to the market.” ‘Large’ basically is defined as commercially unprofitable. The decision rule might thus be that a project is socially beneficial but not commercially feasible: in that case, the costs of line would partly be financed from the pool of regulated network connection charges. Hogan [2003, p. 23] continues that this decision rule avoids doing a cost-benefit analysis. Although this would be appealing, there appears to be a flaw: how could it be determined that a project is not only commercially unprofitable, but would also have a negative net benefit for society? In other words, perhaps the line should not be built at all.

---

8 “Further, ‘large’ would be defined as large enough to have such an impact on market prices that the ex post value of incremental FTRs and other explicit transmission products could not justify the investment.” [Hogan, 2003, p. 23]. Regulated here means that the revenues comes from a pool of regulated connection charges. Details will be clarified further below.
More critical to the prospects of market-based transmission investment are Joskow & Tirole [2003] who forcefully point out a number of problems. A first argument is lumpiness in transmission investment, which implies that rewards based on marginal prices do not suffice and hence lead to underinvestment [Joskow & Tirole, 2003, p. 21 ff.]. This type of argument is basically in the same group as economies of scale as discussed in section 4 below. Further, generation market power at one end of the line will distort the prices and thereby the line investment decision [Joskow & Tirole, 2003, p. 17 ff.]. This may lead to over- or underinvestment depending on which the node the market power is. A quite special problem is what Joskow & Tirole [2003, p. 25] call “state-contingent rights and diversification”; the problem relies on the difficulty to determine the line capacity (as an operational capacity), which depends on the flows in the connecting networks, which in turn depends on for instance demand. If usage of the line is sold off by long-term rights, then it is not clear what is to be sold if capacity is not determined. The theoretical answer is to sell state-contingent rights, which however are not well developed as yet. As the authors [2003, p. 25] note, this problem is typical for AC meshed networks. Another set of problems rely on governance problem associated with the split between the transmission owner (TO) and system operator (SO) which inherent in merchant investment. The problem is who gets paid for what. The details are beyond the scope here but it seems that as above the problem is less severe in the DC case because the flow and thereby the “output quantity” can be controlled by the line owner. A problem which receives in-depth attention in Joskow & Tirole [2003, pp. 39 ff.] is the problem that the new line may be detrimental to the system (due to loopflow effects). The problem has been discussed in for instance Bushnell & Stoft [1996 JRE] and will receive detailed attention in section 4.3 below. A last point to be mentioned is regulatory risk [Joskow & Tirole, 2003, p. 57]. It is suggested that regulatory uncertainty may make funding of the merchant projects infeasible. Whereas this is a strong argument, it should be noted that regulatory uncertainty was at least in Australia the predominant reason to grant regulation holidays and rely on unregulated merchants in the first place. In all, Joskow & Tirole point out a set of possible inefficiencies, which overall appear to be severe only for meshed AC networks, implying that there may be good prospects (or at least less problems) for lines with controllable flow (DC and FACTS).

The difficulty identified in these papers is the regulatory mix of the unavoidable co-existence of regulated and unregulated lines. The difficulty is to find a stable and workable borderline where crowding out of commercially viable projects by regulated projects (or reverse) is unlikely. Drawing the line between controllable flow projects as merchant projects (especially the interconnectors) and AC projects in meshed networks to be reserved area for designated TSO or by authorised tenders may achieve a sharp distinct.
3. Locational marginal pricing, line rentals and investment

The principle underlying merchant investment is called locational marginal pricing (LMP), or nodal spot pricing, which was developed for congestion management by Bohn, Caramanis & Schweppe [1984] and was worked out and applied in New Zealand [cf. Read & Sell, 1989; and Read, 1989]. An important formalisation and modification came with the contribution of Hogan [1992], who extended the basic model by a set of financial hedges, so-called Transmission Congestion Contracts (TCCs). Meanwhile, the LMP-TCC approach has been (or will soon be) implemented in some variation in several states in the USA (e.g. PJM, New York, New England, Texas and California) and is a cornerstone of FERC’s currently debated Standard Market Design. Europe and Australia do not have nodal spot pricing but at best zonal pricing. Basically this means that for an area like for instance the Netherlands there is only one spot price and the network is not further differentiated (as if the network is unconstrained). Nodal spot pricing would define a set of different nodes within the Dutch network, which, depending on congestion within the network, might have different prices. In the nodal pricing scheme a new line connects two different nodes (quite possibly within one network), whereas in Europe and Australia a new line is more likely to interconnect different networks and trade between the associated zonal prices. As a result, the European and Australian interconnectors can roughly be considered as two-node interconnectors.

The basic idea is straightforward. Consider a two-node network with a transmission line between the two nodes. Suppose that at each node a spot price reflects the marginal costs of electricity at that node at that moment. As long as the spot prices at the two nodes differ, then the difference must reflect the opportunity costs of transmission, otherwise traders buy power at the cheap node and sell at the expensive node until the price differential is zero. The opportunity transmission costs to be reflected rely on energy losses and congestion (also called, constraints), which can be seen as the limiting case of energy losses. If the load increases up to the line’s capacity the line will be congested. At that point the TSO will have to secure a dispatch such that the load on this particular line is not further increased. In other words, congestion in the lines affects the dispatch of the generation units, such that a price differential between the nodes remains.

9 Also known as Congestion Revenue Rights or Financial Transmission Rights.
10 The reader may note the equivalence with congestion charging in road pricing as developed by Mohring & Harwitz [1962]. Since energy losses are a squared function of the line load, the optimal transmission charge is twice the energy loss. If the system-dispatcher (TSO) minimizes the production costs (given demand), then the nodal prices will exactly reflect this. Half of the revenues from the transmission charges would cover energy losses (which are real costs) and the other half is a surplus, similar to the Pigouvian tax.
Some situations in electricity networks can usefully be represented by a two-node network. Typically, however, an electricity network consists of many different interconnected nodes, creating a meshed network. A network with alternating current (AC) creates so-called loopflows. Electric power in an AC-network follows Kirchhof’s law, meaning that a power flow divides itself over the network proportional the inverse of the line impedances. The idea is illustrated in figures 1a and 1b. In these figures, representing a three-node network, nodes G1 and G2 are generation nodes and node D3 is a load node. Line Lij is the line between nodes i and j. The three lines are assumed to have the same physical impedance and are equally long. Hence the route from node 2 to node 3 over L12 and L13 is twice as long as over L23, and thus the impedance on the short route is half the impedance of the long route.

Figures 1a and 1b: Loopflows in a three-node AC network.

In figure 1a, it is assumed that there are no line constraints and the load of 900 MW is completely generated by G2. The power flows according to the inverse of the line impedances and thus 600 MW flows over L23 and 300 MW over L12 and then L13. In figure 1b it is assumed that the dispatch is 600 MW from G2 and 300 MW from G1. The 600 MW from G2 divides 400 MW on line L23 and 200 MW on L12 and then L13. The 300 MW from G1 flows 200 MW over L13 and 100 MW over L12 and then L23. In total thus, the flow over L23 is 500 MW, L13 is 400 MW and L12 100 MW. The flow on L12 is determined by subtracting the opposing flows: 200 MW - 100 MW is 100 MW. The dispatch in figure 1b may be the resulting dispatch if the lines are constrained. Suppose that G2 has lower production costs than G1 such that the dispatch in figure 1a would be the desired dispatch. If the capacity of L12 is constrained to 100 MW then the dispatch of figure 1b would be the constrained

11 In technical terms, the impedance is the “sum” of the line’s resistance and reactance.
optimum. The unconstrained dispatch of figure 1a would not be feasible, because L_{12} cannot handle 300 MW.\textsuperscript{12}

Figure 1b also depicts nodal spot prices. The spot prices at the generation nodes are derived from the marginal production costs at these nodes (for this dispatch), which is straightforward.\textsuperscript{13} The price at the demand node is derived as the marginal production costs of one additional demand unit. In this case, 1 MW additional demand would (have to) be produced 0.5 MW from each generation node. \((0.5 \cdot 30\text{€}) + (0.5 \cdot 50\text{€})\) makes 40\text{€}. And so a complete set of nodal spot prices can be calculated. The transmission charges (denoted by \(t_{ij}\)) immediately follow: \(t_{21} = 20\text{€}, t_{23} = 10\text{€},\) and \(t_{13} = -10\text{€}.\) Multiplying with the flows on the subsequent lines gives a surplus of 3000\text{€}. Note that the transmission charge on L_{13} is actually negative because the flow is from a high price node to a low priced nodes.\textsuperscript{14}

The LMP system has the drawback that the spot prices are volatile and the spot price on one node depends implicitly on all other nodes in the network. In other words, an LMP system involves (short and long term) risk for the users. Transmission Congestion Contracts (TCCs) have been developed to hedge these risks. A TCC is defined as a contract for a line between nodes i and j with a strike quantity \(R_{ij}\) paying out to the owner of the TCC the difference between the nodal spot prices \(p_j\) and \(p_i\) times the strike quantity \(R_{ij}\). Hence, the payment of a TCC can be denoted by \(T_{ij} = R_{ij}(p_j - p_i)\). The TSO, being the collector of the transmission charges, may be the counterparty and the TCCs may be allocated by an auction.\textsuperscript{15} Suppose that a trader trading quantity \(q_{ij}\) to prices \(p_i\) and \(p_j\) actually pays the transmission charges \(q_{ij}(p_j - p_i)\) to the TSO and hedges this risk with the contract \(R_{ij}(p_j - p_i)\). It can then quickly be seen that if the strike quantity \(R_{ij}\) approximately matches the real quantity \(q_{ij}\), and given that initial payment for the TCC as such, the price differences cancel out and the risk is hedged.

The LMP concept has been controversially discussed [cf. Wu et. al., 1996 and Oren et. al., 1995]. This debate produced the following important result of Chao & Peck [1996], who contrast the LMP-TCC approach of financial transmission rights with a \textit{flowgate} approach,

\textsuperscript{12} Note that for instance \(G1 = G2 = 450\text{ MW}\) would also be feasible, but not (constrained) optimal. Production costs would be higher than under \(G1 = 300\text{ MW}\) and \(G2 = 600\text{ MW}\) per assumption. In case \(G1 = G2 = 450\text{ MW}\) the power flow on L_{12} = 0.

\textsuperscript{13} Note that the cost functions are not given here.

\textsuperscript{14} In this interpretation, the transmission charge is actually paid by the owner of a trading contract between nodes i and j. Another way to think of transmission charges is that load pays 40\text{€} at node 3 and the G1 receives 50\text{€} at node 1 and G2 receives 30\text{€} at node 1. Multiplied by subsequent quantities gives a surplus of 3000\text{€}.

\textsuperscript{15} There are variations. For a combination with Contracts for Differences, see Bushnell & Stoft [1996 PWP-034]. The allocation of the revenues of the auction (or in other words, line rentals) is again a different issue. It seems natural to allocate the revenue to the line owner as a contribution to the line’s costs. This issue appears to be rather controversial. See for instance Read [2002] for the discussion in New Zealand.
which relies on physical transmission rights. In the LMP approach the “transmission rights” follow from the dispatch, while the TCCs are merely financial instruments and do not provide physical transmission rights. In contrast, physical transmission rights would be allocated prior to production and hence dispatch follows transmission rights rather than reverse (at least, the dispatch should take the transmission rights into account as binding constraints). The flowgate approach applies powerflow distribution factors (PDFs) to calculate which nodes claim how much from the capacity of which line, for congested lines only. Chao & Peck [1996] show that under certain conditions the flowgate model gives the same results as the LMP approach. This is an extremely useful result, which allows to restrict attention here to the LMP analysis; with caution the analysis below may thus be carried over to a flowgate approach. In other words, with respect to interconnector investment, the LMP analysis in the USA may be carried over to an auction setting as in Europe.

A system of spot prices, be it as refined a nodal LMP system or as crude as two different zones, implicitly defines a pricing rule according to which investment in interconnector capacity can be paid: the price differential between different nodes. This can be interpreted as a high-level regulatory rule: the rule-maker has decided that market-based line investment will be paid according to this rule. That is what the definition of the ACCC states explicitly. Section 4 will now explore the problems which may arise with unregulated market-based investment paid by the price differentials.

4. Problems and prospects of market-driven investment

This section discusses four main areas of potential inefficiencies associated with market based transmission investment and examines these on the difference between controllable and non-controllable flow. The four areas are: economies of scale and cost-recovery, market power and the size of capacity, detrimental investment and risk. To recall, this section concerns the unregulated, decentralised option of table 1 only.

4.1 Economies of scale and cost recovery

Economies of scale in the construction of transmission lines are substantial. Footnote 10 pointed to the similarity with the theory on road congestion charging as in Mohring & Harwitz [1962]. The long-run effects are well known from this literature and are directly applicable here. The congestion charge is a surplus over energy losses which can contribute the fixed costs of the infrastructure. The surplus depends on demand relative to capacity. In the long run in which capacity is variable the following result holds: if long-run marginal costs (i.e. capacity expansion costs) are decreasing in capacity, the surplus resulting for optimal capacity will be less than the fixed costs. Hence with economies of scale in the
construction of new transmission lines, the transmission charges relying on the price differentials will not entirely recover fixed costs with optimal capacity size. As a result we can conclude that market-based transmission investment either is not profitable (in which case it will not take place) or capacity is smaller than optimal.

Figure 2: Economies of scale in transmission infrastructure

Figure 2 indicates the relevance of economies of scale based on real construction costs [cf. Fuldner, 1998]. Figure 2 plots average construction costs (US$ per MW per mile) in relation to the line’s capacity, as the least-cost envelop of different technologies. Similar indications come from for example and Read [2002] and Perez-Arriaga et. al. [1995], suggesting that not more than 30% of total costs could be recovered by LMP differentials if capacity is optimal. It appears that DC interconnectors are used for bulk power transactions. As a result the scale economies may be exhausted at some point; figure 2 suggests that beyond 750 MW long-run marginal costs are near constant. The typical merchant situation is DC interconnection of different networks aiming at transmission of bulk power. For this scenario the problem of economies of scale appears less severe than for the typical AC line in a meshed network. The projects in the USA are typically around 1000 MW, which is also the order of magnitude for European projects (for instance, UK-Norway or UK-NL interconnectors). Curiously, however, the interconnectors in Australia are around only 200 MW.

---

16 These are line construction costs only and exclude AC/DC converters. Investment costs may include right-of-way charges, which may be high. If these are high and charged per capacity unit they may dominate the construction costs and may cause capacity expansion costs to have constant or even decreasing scale effects.

17 The argument is in the same direction as Hogan’s position [2003], who, as mentioned above, proposes to restrict regulated lines to those which are large relative to the market, which seems stronger applicable to AC lines.
A first approach to the problem is second-best pricing. Imagine that an implicit tender (say, competition for the field) determines the line owner. In a competitive bidding, the result would be a second-best capacity of the line where the line rentals would exactly recover costs with mark-ups on marginal costs. With relatively inelastic demand, the resulting deviation from the optimum may actually be rather small. The relative deviation from the first-best solution gets smaller the larger the line. This argument is appealing but can be criticized on two accounts. First, a second-best solution would be inferior if a first-best solution with two-part pricing is feasible; theoretically, the first-best solution with two-part pricing can be achieved by the designated TSO and hence there is a trade-off involved. Second, an implicit auction for the right to build the line is not entirely compatible with a decentralized scenario. 18

A second way to proceed is user-specific two-part pricing, although this is not as obvious as it might seem. Apart from an LMP based variable charge a fixed use-of-system charge may contribute to the remaining costs. It is principally possible, but cumbersome and theoretically on weak legs. The idea is to develop an algorithm which allocates the costs of the line in some relation to the usage of the line, for which two methods are used: the area-of-influence method (also called marginal participation) and tracing (also called average participation). 19 Tracing has the economic advantage of relying on the Shapley value [Kattuman et.al., 2003]. The allocation of the cost of the line is irrelevant for the sunk costs of existing lines, but is important for cost-recovery of new lines and hence is important for investment decisions. Roughly speaking, the more meshed the network is, the more difficult it gets to identify users in an economically useful way. Concluding, allocationg the (fixed) costs of a (new) line can be done, although not without difficulty and hence user-specific charging is possible.

The argument has a theoretically flaw. Demand for the interconnector is derived demand from the arbitrage possibilities between the interconnected spot markets. Assume for sake of the argument that the users of the line are traders who arbitrage between two spot markets. The traders generate revenues by buying “kWhs” at the cheap node and sell at the expensive node; in other words, their revenue is expressed in variable terms (per kWh). The underlying cost structure for using the transmission line will be passed through (if at all) as a variable charge by the traders. As long as the traders’ revenues with which the line should be paid are variable, the final result will always be second best. Stronger even, if competition among the traders is fierce, they would compete each other down to variable costs and would not be able to recover the fixed charge. The problem of under-recovery of the costs would simply be

18 In table 1 this would approximate the “hybrid / tender” solution. It is beyond the scope here to go into detail. 19 It is beyond the scope of this paper to go into detail and instead the reader may consult Vazquez et.al. [2002] and Kattuman et.al. [2003]. The method of area of influence is applied in Argentina where it works reasonably well, because of the radial network into Buenos Aires [cf. Woolf, 2003, p. 265].
passed on. If this is the result then the line itself might have been charged with a uniform mark-up in the first place.

A third aspect to be considered is whether all costs and benefits are in fact included in the LMP-based line revenues and hence whether they are internalised in the investment decision. Three issues are relevant. First, new lines will in general have an impact on the reliability of the system. A new line may increase reliability in the network by increasing capacity in which case the TSO might compensate the line owner. Moreover, controllable flow lines increase the system’s transfer capability and add to the system’s stability by being controllable [cf. Gyugi, 1999, p. 31; and Arrillaga, 1998, p. 8]. On the other hand, especially in the face of loopflows the new line might decrease reliability and even require upgrading the network. In that case, the line should be charged “deeply” for the costs of upgrading. Second, capacity payment. In for instance PJM, the authorities installed a market for generation capacity in which capacity contracts are traded. The capacity prices differ according to relative scarcity between different areas. A new line connecting two areas with different capacity prices can arbitrage the price difference. Line revenues would then consist of energy price differences and capacity price differences. Third, environmental effects should be taken into account. New transmission lines will in general cause environmental cost, but these costs may be less than alternatives. For instance, subsea and underground cables are far less perceived as environmentally damaging than overground cables.

4.2 Market power and size of capacity

Profit-driven investors will have an incentive to maximize profits rather than welfare. New transmission capacity between two nodes will usually lower the price difference between the two nodes and hence lower the line rentals. In analogy with normal monopoly type behaviour, investors will seek to restrict capacity below optimal capacity. Faced with scale economies in line-construction, it is to be feared that decentralized transmission investment may result in inefficient monopoly-type investments.21

Apart from the direct distortion, there may be indirect effects. An important benefit of additional interconnector capacity is that it enlarges the relevant markets of the generators; in

20 The effect on total revenue is not a priori clear.
21 Depending on demand and the magnitude of scale economies, the underinvestment may partly be offset by preemptive investment (similar to limit pricing). In a world with firm transmission rights the line owner may then decide not to use all capacity to restrict availability of the line. In a slightly different setting, the argument reminds of the argument put forth in Gilbert, Neuhoff & Newbery [2002] and Joskow & Tirole [2000] that a dominant importing generator has an incentive to acquire (and then restrict the use of) physical transmission rights in order to retain its market power. Restrictions on capacity withholding would relieve the problem partly, but might on the other hand have adverse effects for the level of investment.
other words, depending on whether market power is on the exporting or the importing node it may mitigate market power on the generator side. Assume market power in generation at the import node. The direct effect is that additional capacity is the same as an additional competitor (say, Cournot-like competition with more firms) and the indirect effect is that increased total capacity reduces the margin between (peak) demand and total capacity and hence will decrease the probability of a pivotal firm (i.e. a change in the nature of competition). Thus a new line may increase competitiveness, but if market power induces the investor to keep capacity inefficiently small, the effect on generator competitiveness will be inefficiently small as well.

There are several ways to approach the “monopoly” problem. The straightforward approaches are to either require passing the best option test which thereby gives a check on capacity or organize a tender after the project has been indentified by a commission (i.e. the hybrid/tender option in table 1). Ideally, both cases would result in the second-best solution which may be highly preferable to the monopoly outcome. Especially in combination with arguments put forth below this is appealing but has the drawback that it inevitably reintroduces an element of monopolised decision making.

An alternatively approach might take the view that the monopoly problem is primarily a problem of the AC network and less so for controllable interconnectors. In as far as “parallel” lines are feasible at all, controllable flow actually allows a competitive choice. In a non-controllable system “parallel lines” would still be “monopoly”, because the parallel lines are “bundled”. In a DC-system two parallel lines can actually compete in capacity, while in a non-controllable system this is technically not possible. Moreover, in a non-controllable system, the capacity of line A determines the capacity of “parallel” line B. It follows that regarding the capacity decisions, parallel controllable lines are strategic substitutes and “parallel” non-controllable lines strategic complements. From this it is then straightforward that -if at all- the competitive pressure among controllable lines will be stronger.

A related but slightly different “monopoly” problem is pre-emptive investment, which basically means strategic investment to deter others. The arguments are much in line with the limit pricing approach developed by Bain, Sylos-Labini and Modigliani.22 If due to economies of scale and/or lumpiness entrants can only profitably enter at some minimum efficient scale, the incumbent can invest pre-emptively so as just to deter the entrant. The result is that the capacity of the investment is the minimum of the monopoly capacity and the minimum

capacity which just deters entry: call this the limit capacity. The pessimistic view is that the limit capacity is less than the optimal capacity, which is normally correct, but may be the wrong benchmark. The optimistic view may emphasise that given the monopoly problem, the limit capacity is at least as big as the monopoly capacity and thus pre-emptive investment mitigates this problem. In all, the argument stresses that there may be some pressure from potential new investors.

Following the line of argument on limit pricing approaches implies that if demand is large as compared to the minimum efficient scale, a point will be reached where it is no longer profitable to deter entry. The limit capacity would have to be too large and it might actually be more profitable to accommodate new entry. There are assumptions underlying this rather theoretical result, but the main lesson seems to hold throughout. If interconnecting DC lines are typically used for bulk transaction over long distances, the size of the market may be large relative to optimal line sizes. Consequently, the required pre-emptive investment may be sufficiently big such that entry accommodation is more attractive.

4.3 “Profitable expansion can be bad”

The principle of rewarding investment according to the price difference between the two nodes which are interconnected by the new line (link based) is flawed, because it ignores network effects. In the debate on the usefulness of LMP, Wu et.al. [1996] and Oren et.al. [1995] pointed out that under a regime with link-based LMP-TCCs, profitable market-based transmission investment can actually be detrimental to the system and hence be inefficient. Consider figures 3a and 3b, which are closely related to figures 1a and 1b.

In figure 3a, there is no transmission line between G1 and G2 and the resulting dispatch then is that G2 produces 900 MW and G1 0 MW and the power flows are straightforward. In absence of constraints the prices are 30€ at all nodes, corresponding to marginal costs at node G2. Now assume that a merchant invests in a 100 MW line between nodes G1 and G2. The corresponding dispatch then becomes as in figure 3b, which corresponds to figure 1b. The noticeable change is that the power flows cause the new line to be constrained which then alters the dispatch such that G1 produces 300 MW at as can be seen relatively high costs. The resulting prices are as given. Assuming link-based payment, it follows immediately that the investment is profitable if the investment costs are lower than 2000€. Welfare has decreased

23 This appears to be the point of view of Joskow & Tirole [2003, p. 23] with respect to pre-emptive investment.
24 With lumpiness, pre-emptive investment may even result in overinvestment: i.e. larger than optimal capacity.
25 Link-based line rentals on G1-G2 are 100·(50-30) = 2000.
because the production costs have increased while output did not change. Hence, a bad modification can be profitable.

**Figures 3a and 3b: “Bad” modification can be profitable.**
Source: Bushnell & Stoft [1996, PWP-034, p. 5].

The fundamental problem underlying this example is that link-based line rentals, defined as the difference between the prices at the two nodes connecting the line, do not reflect incremental network effects. Whereas the line transmission charge does reflect the opportunity costs in a two-node network, this is no so in a meshed network. The net benefits of the line investment should take account the impact elsewhere in the network (here the change in the line rentals between G1 - D3 and G2 - D3). This discrepancy between link-based line rentals and rewards for line investment casts doubts on the usefulness of LMPs as a guide for transmission investment if incremental network effects are significant. It may be recalled that LMPs aim in first instance at the (short-run) use of the network and investment signals to the users of the network (generators and load) and not investors in the network [cf. Hogan, 2003, p. 13].

A powerful solution to the problem has been developed by Bushnell & Stoft in series of articles [1996 (JRE), 1996 (PWP-030), 1996 (PWP-034), 1997]. The authors propose to modify the investment reward system such that new investment is rewarded by a “must-accept” set of TCCs, which in essence captures the incremental external effects of the new investment over and above the direct rewards of the invested line. It is important to realize that the TCC pays \( \Delta R_{ij} (p_j - p_i) \) to its owner; in the proposal, the merchant line investor is the owner. As above, \( p_i \) is the spot price at node i. \( \Delta R_{ij} \) is the TCC strike quantity of line ij, representing the difference between the dispatched flow after and before the line investment. Note that the value of a so-defined TCC can be negative. The investor would have to accept
the set of so-defined TCCs for all affected lines. Bushnell & Stoft [1996 (JRE) p. 73] show that if the consolidated set of contracts match the current dispatch “then no group of agents whose contracts match their dispatch will find it profitable to make detrimental alterations to the grid.”

The key modification is to capture the incremental network effects, which implies the step from link-based line rentals to network rentals on the one hand, and payment according to increments of flows ($\Delta R_{ij}$) rather than total flows ($R_{ij}$) on the other. In result, the impact of the new line on the “entire” system is captured. In the example above, the new set of TCCs would be: $\Delta R_{12} = -100 \cdot (30-50) = +2000$, $\Delta R_{13} = +400 \cdot (40-50) = -4000$, and $\Delta R_{23} = -400 \cdot (40-30) = -4000$, which in total sums to -6000.

The system is not without drawbacks. First, the system is path-dependent. It relies on changes in flows and thus always compares with a no longer existing situation. Payment and thereby incentives for new investment rely on the current network by method; hence if the current network is not efficient, then the inefficiency is likely to carry over. A second problem has been pointed out by Bushnell and Stoft [1996 (JRE), p. 77]. The requirements of matching are extremely unlikely to be met. A third problem is more fundamental. This type of reward for the investment requires assessment of both the old and the new dispatch which is controversial. A central institution will have to decide on the external incremental power flows as the basis for the must-accept contracts. Hence, whether or not the investment will be profitable depends to a large extent on a discretionary decision making power of a commission. This may be unavoidable but principally contradicts the idea of decentralized, unregulated merchant. The point has been well put by Joskow & Tirole [2003, p. 42, italics in original]: “It should be clear as well that in practice the merchant transmission model cannot operate “as if by an invisible hand”, since some de facto regulatory authority must have the ability accurately to simulate load flows on the network, apply contingency criteria, define feasible sets and changes in feasible sets associated with transmission investments, and ensure that rights allocations are consistent with feasibility under numerous contingencies.”

To sum up the argument, link-based LMP-TCC based rewards is the elegant way to reward decentralized new investment, but ignores the network effects and may thus be inefficient. Network-based LMP-TCCs capture the network effects but do not allow fully decentralized decision making and restricts market-based transmission investment. Presumably, this is the

---

26 Unclear is whether this is the same as “efficient investment”.
27 Note that the link-based LMP-TCC based price is part of the bigger scheme.
28 It seems that US control area PJM moves into this direction for AC lines with “Incremental Deliverability Rights” (IDRs).
most important flaw of the link-based LMP-TCC based system as a basis for rewarding market-based transmission investment. Once again, however, the problem is inherent to non-controllable flow and far less severe for controllable flow.

Consider the example in figure 3. The “bad” modification as exemplified in figure 3b is caused by the loopflow problem. Kirchhoff’s laws dictate that the power flow on $R_{23}$ is less than 900, because of the proportional split and the line constraint on $R_{12}$. This no longer holds if the new line $R_{12}$ is a controllable flow. If in the example in figure 3b the flows are controllable, the new line simply would not be used (i.e. the flow would be set at zero) and the dispatch could be as in figure 3a. It is thus rather unlikely that the line would be built in the first place. In result, controllable technology reduces the network effects and thereby strengthens the relation between link-based LMP based profitability and welfare effects.\(^\text{29}\)

### 4.4 Risk

The last problem with market-driven transmission investment to be put forth here relies on high risks caused by monetary spill-over effects.\(^\text{30}\) The precise extent and nature of the spill-over depend on how exactly the line investor is rewarded, but the result always is that revenue is uncertain. Within an LMP based scenario it appears quite difficult to find a perfect hedge. It follows that market-based transmission investment may be quite risky which will tend to suppress investment levels.

Suppose first that the line owner is paid according to link-based LMP line rentals. The profits would be:

\[
\pi_{ij} = q_{ij} \cdot (p_j - p_i) - K
\]

where $q_{ij}$ denotes the real flow in the line $ij$, $p_i$ and $p_j$ are the spot prices at nodes $i$ and $j$ and $K$ is the investment cost. In this setting the investor is extremely vulnerable to investments elsewhere in the network. Not only the spot prices may vary beyond its control, but due to loopflows the quantity as well, which is well illustrated by figures 3a and 3b. Importantly, once invested the returns are largely beyond control of the investor, which, given the interactions of meshed networks, makes it rather hazardous.

Second, suppose that the line owner sells off TCCs to network users over and above the line rentals. Denote $A$ as the (auction) revenue of the sold TCCs. The investor’s profit is:

\[
\pi_{ij} = q_{ij} \cdot (p_j - p_i) - R_{ij} \cdot (p_j - p_i) - K + A
\]

\(^{29}\) The argument has larger application than merely part of a meshed network if the three nodes are considered to be for instance France, UK and Benelux and the interconnectors are AC or DC lines.

\(^{30}\) It must be emphasized that these should be distinguished from real external effects.
If the real flows \((q_{ij})\) and the TCC’s strike quantity \((R_{ij})\) match, the investor is insulated against changes in the spot prices. However, the investor is vulnerable against the quantity effect: any new investment (or demand) will affect the real power flows. It quickly follows that profit decreases if \(q_{ij}\) decreases. As above, with non-controllable flow the power flows are largely beyond control of the line owner and thus despite hedging considerable risks remain.

Third, suppose that the line investor is rewarded with TCCs (as opposed to line rentals). The investor’s profit then is as follows:

\[
\pi_{ij} = R_{ij} \cdot (p_j - p_i) - K
\]

The investor is insulated against quantity risk. Instead, it is now vulnerable to the spot prices. If the differential decreases, profit decreases. This is likely to happen, if for instance a new power plant is built in the vicinity of the high price node.\(^{31}\) It may be recalled that high nodal spot prices signals new investment opportunities. It may be noted that the risks involved in this third scenario readily translate to the Bushnell & Stoft network-based payment as characterized above.

A fourth option allocates TCCs to the line owner, who then auctions off the TCCs to the users. Both the congestion charges as well as payment to the TCC owners are taken care off by the system operator and beyond the line owner. The line owner’s profit would be:

\[
\pi_{ij} = A - K
\]

The ex-post risks would be shifted completely to the users. The line owner would only have the ex-ante risk of the auction revenue, which depends on the definition of the TCCs. Moreover the TCCs prices derived from the auctions presumably reflect a risk premium, which in turn depends on the level of uncertainty.

It is clear from the examples that the risk (-allocation) depends on the type of reward, which in turn depends on the institutions. The effect of the risk will be to require a high risk premium and hence increase cost of capital, or make isolated projects unprofitable all together. Whether the risks are prohibitively high or manageable is an empirical matter. Overall the difference between controllable and non-controllable seems decisive. The loopflows in the non-controllable system make the actual (future) flows rather difficult to predict; the risks are amplified by loopflows. In contrast, the quantity in the controllable line can be determined and the problem reduces to an economic problem.

\(^{31}\) In fact, Directlink, one of the merchant projects in Australia (connecting Queensland and New South Wales), faced this problem.
5. Concluding remarks

Despite drawbacks market-based transmission investment may well have sufficient prospects to offer support for close examination. First, the (monopolised and regulated) alternatives do have well-known drawbacks as well, among which under-investment. Second, market-based transmission investment takes place in practice. Third, legislators and regulators are developing regulatory frameworks to approach the situation [cf. e.g. Newbery, Von der Fehr & Van Damme, 2003]. Whether these are permanent developments is yet an open question, but they do provide arguments for giving the issue proper attention. This paper draws explicit attention to controllable flow and argues that market-based transmission investment may be allowed for controllable flow while prohibited for non-controllable flow. Thereby, it seems that new high-voltage direct current (HVDC) interconnectors can well be market based.

Four main problem areas of market-based transmission investment have been examined on their difference between controllable and non-controllable flow. A first problem is economies of scale. At least theoretically the argument can be made that market-based investment will be smaller than optimal. The severity of the problem depends on the size of the line relative to the market. Typically DC lines are used to interconnect different network areas between which potential power flows can be expected to be large. The problem will thus be more severe for situations where AC lines would be economical. Moreover, although hardly free of problems, controllable flow seems to offer some prospects for user-specific two-part charging.

A second problem is that market-based investment may actually be a monopoly investment. The severity reduces if the market is large as compared to the lines, because at some point the market may allow competing lines. More importantly, parallel DC lines being controllable can actually be competitive. In contrast, on AC “parallel” lines, the non-controllable flows over the lines would effectively be bundled and could not compete. Hence, competitive potential between lines, if at all, requires controllable flow and will thus reduce the severity of the monopoly problem for controllable lines.

A third problem is that the LMP-based reward for new transmission investment either may be inefficient or require a modification of the rule, which inevitably is a move away from decentralised decisions. In the face of loopflows, a reward system based on the spot price on the two ends of the line only (“link based”) may well be inefficient, because impacts elsewhere in the network are not reflected in the revenues. The way out is to modify the rule by creating a set of incremental payments (“network based”), basically reflecting the impact of the line in other parts of the network. The main problem with this is that this has to be estimated by a centralised agency and is open to controversy and legal challenges; thereby,
the major advantage of market-based transmission investment would vanish. The problem is inherently related to loopflows and is thus typically the problem of meshed non-controllable lines. Since both the link-based and the network-based approach have disadvantages, there seems to be a rather strong argument to restrict merchant investment to controllable lines, because it reduces the problem in the first place.

A fourth problem is risk. Ultimately, market-based lines are rewarded by the revenues coming from flows and prices determined in the market. Provided liquid markets for these financial instruments exist, it is possible to hedge these risks but the hedging will never be complete. Insulation against price volatility can be achieved, but hedging quantity volatility as well seems more difficult. Quantity is the main point which cannot be controlled on a non-controllable line and hence as above the problem seems to be more severe for non-controllable lines.

Further research might focus on the following two aspects. First, the distinction between controllable and non-controllable flow might distort investment decisions between different types of technology. Whether this is an empirically relevant effect needs further examination. Second, unregulated as used in this paper means that the revenues are not regulated. At the same time, a regulator or legislator might well require other regulatory provisions, for instance concerning third-party access to the line. The approaches differ quite strongly between the USA and Australia and the framework in Europe still has to be settled.

References


