Deutsches Institut für Wirtschaftsforschung

DIW BERLIN

Discussion Papers

1025

William Hogan • Juan Rosellón • Ingo Vogelsang

Toward a Combined Merchant-Regulatory Mechanism for Electricity Transmission Expansion

Berlin, June 2010

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Available for free downloading from the DIW Berlin website.

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN.
Papers can be downloaded free of charge from the following websites:

http://www.diw.de/de/diw_01.c.100406.de/publikationen_veranstaltungen/publikationen/diskussionspapiere/diskussionspapiere.html
http://ideas.repec.org/s/diw/diwwpp.html
http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=1079991

Toward a Combined Merchant-Regulatory Mechanism for Electricity Transmission Expansion

William Hogan

Kennedy School of Government, Harvard University

Juan Rosellón*

Centro de Investigación y Docencia Económicas (CIDE) and German Institute for Economic Research (DIW Berlin)

Ingo Vogelsang**

Department of Economics, Boston University

June 2010

* Contact details in Mexico: División de Economía, CIDE, Carret. México-Toluca 3655, Lomas de Santa Fé, C.P. 01210, México D.F. Mexico. Email juan.rosellon@cide.edu. Contact details in Germany: German Institute for Economic Research (DIW Berlin), Mohrenstr. 58, 10117 Berlin, Germany, email: jrosellon@diw.de. This author acknowledges support of the Programa Interinstitucional de Estudios sobre la Región de América del Norte (PIERAN) at El Colegio de México, the Alexander von Humboldt Foundation and Conacyt (p. 60334).

** This author's research on this paper was supported by NSF Grant ECS-0323620 to Carnegie-Mellon University on "EPNES: Dynamic Transmission Provision and Pricing for Electric Power Systems."

Abstract

Overview

Electricity transmission pricing and transmission grid expansion have received increasing regulatory and analytical attention in recent years. Since electricity transmission is a very special service with unusual characteristics, such as loop flows, the approaches have been largely tailor-made and not simply taken from the general economic literature or from the more specific but still general incentive regulation literature. An exception has been Vogelsang (2001), who postulated transmission cost and demand functions with fairly general properties and then adapted known regulatory adjustment processes to the electricity transmission problem. A concern with this approach has been that the properties of transmission cost and demand functions are little known but are suspected to differ from conventional functional forms. The assumed cost and demand properties in Vogelsang (2001) may actually not hold for transmission companies (Transcos). Loop-flows imply that certain investments in transmission upgrades cause negative network effects on other transmission links, so that capacity is multidimensional. Total network capacity might even decrease due to the addition of new capacity in certain transmission links. The transmission capacity cost function can be discontinuous. There are two disparate approaches to transmission investment: one employs the theory based on long-run financial rights (LTFTR) to transmission (merchant approach), while the other is based on the incentive-regulation hypothesis (regulatory approach). An independent system operator (ISO) could handle the actual dispatch and operational pricing. The transmission firm is regulated through benchmark or price regulation to provide long-term investment incentives while avoiding congestion. In this paper we consider the elements that could combine the merchant and regulatory approaches in a setting with price-taking electricity generators and loads.

Methods

Based on LTFTRs, merchant mechanisms are easiest to understand for incrementally small expansions in meshed networks under an ISO environment. The price-cap method seeks to regulate a monopoly Transco. The regulatory goal in this paper is an extension of Vogelsang (2001) for meshed projects. Transmission output is redefined in terms of incremental LTFTRs (or total LTFTRs, if a long period is assumed) so as to be able to apply the Vogelsang's incentive mechanism to a meshed network. For lumpy and large transmission projects a fixed part of the tariff plays the role of a complementary charge. The variable part of the tariff is based on nodal prices; pricing for the different cost components of transmission is such that they do not conflict with each other (fixed costs are allocated so that the variable charges are able to reflect nodal prices); variations in fixed charges over time partially counteract the variability of nodal prices giving some price insurance to the market participants.

Results

We consider two types of price index weights: chained Laspeyres weights and idealized weights. Laspeyres weights are easily calculated and have shown good economic properties under well-behaved and stable cost and demand conditions. Idealized weights correspond to perfectly predicted quantities and posses strong efficiency properties. In our model, idealized weights provide incentives for marginal cost pricing.

Regarding transmission cost functions, we explore a series of simplified cases to argue that in a variety of circumstances the cost functions could have reasonable economic properties. The results suggest directions for further research to explore the properties of the cost functions and implications for design of practical incentive mechanisms and the integration with merchant investment in organized markets with LTFTRs.

Conclusions

This paper addresses institutional frameworks, transmission cost and demand functions It is a step in a continuing research agenda to extend incentive regulation while maintaining compatibility with operation of electricity markets.

Keywords: Electricity transmission, Incentive regulation, Financial transmission rights, Loop-flow problem.

JEL codes: D24, L51, L94

1 Introduction

The topic of long-term electricity transmission expansion has received limited attention in the economics literature. The analysis of electricity markets often assumes that transmission capacity is fixed in contrast with its dynamic nature and interdependence with other electricity subsectors. Analysis of incentives for expanding the transmission network is challenging in part because equilibrium in the transmission market has to be coordinated with equilibrium in other markets such as the electricity spot market, bilateral contracts, and related ancillary services such as capacity reserves markets (see Stoft and Graves 2000; Wilson, 2002). In addition, loop-flows imply that certain investments in transmission upgrades cause negative network effects on other transmission links, so that capacity is multidimensional.[2] Moreover, the transmission capacity function can be discontinuous.

Electricity transmission pricing and transmission grid expansion have received increasing regulatory and analytical attention in recent years. For overviews of alternative approaches and debates, see Brunekreeft et al. (2005) and Stoft (2006). Since electricity transmission is a very special service with unusual characteristics, the approaches have been largely tailor-made and not simply taken from the general economic literature or from the more specific but still general incentive regulation literature. An exception has been Vogelsang (2001), who postulated transmission cost and demand functions with fairly general properties and then adapted known regulatory adjustment processes to the electricity transmission problem. Vogelsang (2001) discusses a concern with this approach that the properties of transmission cost and demand functions are little known but may differ materially from conventional functional forms. Hence the assumed cost and demand properties in Vogelsang (2001) may actually not hold for transmission companies (Transcos).

[2] Total network capacity might even decrease due to the addition of new capacity in certain transmission links. For an illustration, see Hogan (2002a).

The electricity transmission network has attracted additional attention due to power outages such as the one of August 14, 2003, in Northeast US which affected more than 20 million consumers and six control areas (Ontario, Quebec, Midwest, PJM, New England, and New York), and shut down 61,000 MW of generation capacity. Problems with coordination and capacity of transmission grids were related to this outage. Similar events in other parts of the world such as UK, Italy, Norway, Sweden, Brazil, Argentina, Chile and New Zealand have also awakened the interest in the factors that determine investment to assure the reliability of transmission grids.

Hogan (1992, 2002b) applies nodal prices from the power flow model as well as financial transmission rights (FTRs) in order to hedge consumers from variations in such prices. Short run congestion of the transmission grid is then priced through the use of differences in nodal prices. The FTRs provide a workable system of property rights. However, there is debate --both in theory and practice-- over the way to promote the expansion of the network in the long run. Incentive structures proposed to attract investment to the grid range from "merchant" to "regulated". In practice, regulation has typically been used in the UK and Scandinavia, while mixed regulatory and merchant mechanisms have been considered in the organized markets such as PJM, New York State, New England, and California.[3] A mixture of regulation and merchant incentives has been tried in Australia (see Littlechild, 2003) and Argentina (Littlechild and Skerk, 2004).

In practice, most transmission investment occurs under a planning regime with traditional cost of serve regulation which contains a variety of challenges and incentive problems. Analysis of alternative incentive structures for transmission investment is mainly divided into two approaches: the long-run financial-transmission-right theory, and the incentive-regulation hypothesis.[4] The first approach is based on long-term FTR (LTFTR) managed through a variety of allocations and auctions by an independent system operator (ISO). Participation of economic agents in auctions is

[3] The merchant mechanisms used in Northeast US are a combination of long-term FTRs and planning (see Pope, 2002, and Harvey, 2002). No restructured electricity industry in the world has adopted a pure merchant approach.
[4] There is a third approach based on the market-power perspective, which derives optimal transmission expansion from the market-power structure of generators, and takes into account the conjectures of each generator regarding other generators' marginal costs due to the expansion (Sheffrin and Wolak, 2001, Wolak, 2000, and California ISO and London Economics International, 2003).

voluntary and therefore this approach is also known as a merchant mechanism. This method deals with loop-flow externalities. For instance, to proceed with line expansions protects all assigned and some unassigned rights while maintaining simultaneous feasibility of the system for new and existing FTRs (see Hogan, 2002a, and Kristiansen and Rosellón, 2006).[5] Under the approach, "merchants" could invest in new transmission capacity and finance their investments through the sale of LTFTRs.

The second approach to transmission expansion relies on regulatory mechanisms for a Transco. The transmission firm is regulated through benchmark regulation or price regulation to provide long-term investment incentives, while avoiding congestion. Léautier (2000), Grande and Wangesteen (2000), and Joskow and Tirole (2002) discuss mechanisms that compare the Transco performance with a measure of welfare loss. Another regulatory alternative is a two-part tariff cap proposed by Vogelsang (2001) where incentives for investment in expanding the grid derive from the rebalancing of the fixed and the variable parts of the tariff. While Vogelsang leaves the definition of the output for transmission open, Bushnell and Stoft (1997), and Hogan (2002a, 2002b) argue that this task is difficult since the physical flow through a meshed transmission network is complex and highly interdependent among transactions. However, Vogelsang suggests that bilaterals contracts or point-to-point transmission rights may provide for an appropriate definition of output.

In this paper, we consider the merchant and regulatory approaches in the context of price-taking generators and loads. We would like to extract the best properties of these two mechanisms, and allow them to cohabitate in the electricity market in a way that addresses some of the special problems of transmission networks without creating a regulatory framework that undermines a policy goal of relying more on market choices. Based on LTFTRs, merchant mechanisms are most easily understood for incrementally small expansions in meshed networks. A price-cap method

[5] Bushnell and Stoft (1997) address this by having the agents responsible for externalities pay back for them. They show that when FTRs exactly match dispatch, welfare cannot be reduced through the gaming of certain agents.

seeks to regulate a monopoly Transco.[6] The particular regulatory model pursued in this paper is an extension of Vogelsang (2001) for large and lumpy *meshed* projects. It is designed for Transcos but – as in the Vogelsang (2001) discussion – it could also be applied under an ISO institutional setting. Transmission output is redefined in terms of incremental LTFTRs (or total LTFTRs, if a long period is assumed) so as to be able to apply the Vogelsang's incentive mechanism to a meshed network. Constructing the output measure and property rights model in terms of FTRs provides the regulatory model with a connection to the merchant investment theory.

As discussed in Vogelsang (2001), the regulatory incentive approach could be difficult to apply to meshed networks due to the loop-flow problem.[7] The transmission cost function may not be well-behaved in the sense that marginal costs could increase and decrease as transmission output increases. Such a problem might be theoretically addressed with free disposal, but this is difficult in the case of electricity.

Pérez-Arriaga et al (1995) show that revenues from efficient nodal prices that define short-run marginal opportunity costs of transmission recover only approximately 25% of total costs of a representative transmission grid.. Therefore, Rubio-Odériz and Pérez-Arriaga (2000) propose that revenues from FTRs be complemented with a charge (*complementary charge*) to recuperate the remaining 75% (fixed) costs.[8] With both merchant and regulated investments for lumpy and large transmission projects, the fixed part of the tariff plays the role of a complementary charge. Our model has desirable properties of transmission pricing too: a) The variable part of the tariff is related to the efficient nodal prices; b) Pricing for the different cost components of transmission is such that they do not conflict with each other: fixed costs are allocated so that the variable charges are able to reflect nodal prices, and c) Variations in fixed charges over time partially counteract the variability of nodal prices, giving some aggregate price insurance to the market participants.

[6] See Vogelsang (2001), and Rosellón (2007).
[7] In fact, as argued by Joskow and Tirole (2005), this problem remains with FTRs and point-to-point transmissions in general. Ways to deal with the loop-flow problem under an FTR approach are proposed in Hogan (2002a), and Kristiansen and Rosellón (2006).
[8] The "complementary charge" is equivalent to the "license access charge" frequently used in the terminology of the United States electricity industry.

The plan of the paper is as follows. Section 2 sketches and institutional set up, while section 3 characterizes transmission output based on point to point transactions or LTFTRs. The further two sections successively examine simplified transmission cost and demand functions for such an output. Section 6 summarizes a possible sequencing of moves in a model for combined merchant-regulatory mechanism. We end up in the last section with discussion onthe implications for a continuing research agenda.

2 The institutional setup

The institutional setup we have in mind for a combined merchant-regulatory model is based on ownership of and investment in the transmission grid by a Transco that is regulated under a type of price cap scheme. In the first instance, focus on the regulated transmission company and turn later to consider the role of merchant investment. The Transco sells in regular long-term intervals LTFTRs to interested parties. The sale guarantees market clearing through an approprriate auction. The LTFTRs entitle their owners to receive congestion revenues from the difference in nodal prices in real-time markets run by an ISO. Thus, the Transco does not receive any congestion revenues directly. However, the buyers of LTFTRs will base their LTFTR purchases on expected congestion revenues so that the Transco will receive an amount related to the expected value of congestion revenues. In addition, the Transco receives fixed access fees from the loads as users of the grid. Thus, if loads pay the fixed fee they end up paying the price of electricity at their nodes plus the fixed fee, while the generators receive the electricity price at their nodes. To simplify, we assume no fixed fees charged to generators.[9] Generators and loads can bid their respective electricity supplies and demands, from which real time nodal prices would be determined. Alternatively, they can sign bilaterals based on expected transmission congestion charges or based on the acquisition of the necessary FTRs.

[9] The problem of two-part tariffs under competition between the buyers was first analyzed by Ordover and Panzar (1982).

The use of FTRs separates the Transco decision from the volatile short term prices and congestion rents. Assuming efficient operation provided by an ISO, the incentives for the Transco turn to the incentives for investment.

There are many approaches to incentive regulation. The essence of the Vogelsang model is to provide a regulated firm with a price-cap type constraints with a "fixed" and "variable" component. Given a demand *q* for the product at price *p*, and a lump sum payment of *F*. The firm faces costs *c(q)*. Assuming stable costs and demands, the regulator sets weights q^w and the regulated firm faces a price index constraint:

$$p^t q^w + F^t \leq p^{t-1} q^w + F^{t-1}.$$

The regulated firm's profit maximization problem becomes:

$$\begin{aligned}
&\underset{p^t, F^t}{Max} \quad p^t q\left(p^t\right) + F^t - c\left(q\left(p^t\right)\right) \\
&s.t. \\
&p^t q^w + F^t \leq p^{t-1} q^w + F^{t-1}
\end{aligned}$$

The task of the regulatory design is to specify the output and the weights. Vogelsang (2001) explains how this simple price index constraint can provide good incentives with minimal information requirements for the regulator. The focus on output in terms of FTR provides the natural connection to the merchant investment model.

The fixed fees are determined, based on the prices for FTRs in such a way that the average of fixed fees and FTR does not exceed a pre-specified level. Weights for calculating the averages could be (1) last period's quantities for the types of FTRs and the number of customers for the fixed fees (Laspeyres weights), or (2) they could be projected optimal quantities (idealized weights). The general framework implicitly seeks to inherit the efficiency properties of the model in Vogelsang (2001): When Laspeyres weights are used, and well-behaved cost and demand functions are assumed stationary, transmission capacity converges to optimal capacity and two-part Ramsey prices after many periods. The use of idealized weights in Vogelsang (2001, pp. 147-151) grants immediate optimality. A combined Laspeyres/Paasche weight performs better than Laspeyres weight under a linear or concave demand function.

In order to assess the restrictiveness of these assumptions we illustrate general properties of transmission cost functions and demand functions.

3 Characterization of transmission outputs

In order to characterize cost and demand functions one needs to define the relevant output variable, for which costs and demands shall be determined. In a vertically separated setting, where transmission is provided by a stand-alone Transco, the users of the transmission grid are electricity generators, who want to deliver electricity to load-serving entities (LSEs), and loads, who want to buy electricity from generators (with or without the help of intermediaries). Transmission makes these transactions possible so that the main service of Transcos is the delivery of electricity between generation nodes and consumption nodes.

The literature on price cap regulation of Transcos often considers the electricity transmission activity as an output (or throughput) process (Vogelsang, 2001). It seeks to derive cost and production functions for transmission services, which are abstract throughput constructs that are difficult to make operational in a meshed network (Vogelsang, 2006). The generic price-cap model assumes that well-behaved transmission demand functions are differentiable and downward sloping, and that transmission marginal costs curves cut demands only once. These assumptions are unrealistic with loop flows in meshed networks (Bushnell and Stoft, 1997, Hogan, 2002a, and Hogan, 2002b). These objections do not demonstrate that the Vogelsang's mechanism totally fails in the presence of loop flows. Rather, the price-cap incentive behavior has not been analyzed.

The FTR literature does not consider the electricity transmission activity as a throughput process. It rather concentrates on a simultaneously feasible set of "point-to-point" (PTP) financial transactions based on rights, obligations and options (Hogan, 2002b). Physical transmission rights are also discussed in the FTR literature.The benefits FTRs over physical rights has been addressed in Joskow and Tirole, 2000.[10]

In this paper, we consider transmission output as LTFTRs (obligations) that are defined between nodes. An LTFTR q_{ij} represents the right to collect or the obligation to pay the net revenues equivalent to injecting electricity in the amount of q at node i and taking delivery of the same amount at node j. The FTR does not specify the path taken between i and j. The Vogelsang price index model depends on the definition of output (here FTRs) with the corresponding cost and demand functions for the transmission output.

4 Transmission cost functions

In the following, we illustrate stylized transmission cost functions to develop some insight about the underlying economic properties. To simplify, the topology of all nodes and links is given, and only the capacity of lines can be changed. By a network topology, we mean a set of nodes with their locations and a set of lines with associated impedances between these nodes. This is poor approximation of reality but it is sufficient to illustrate some the considerations that must be addressed in real electricity networks.

Many of the potential problems of transmission cost functions alluded to earlier derive from loop flows that could lead to negative marginal costs and discontinuities in costs. For practical purposes we distinguish between generation nodes, consumption nodes and intermediate nodes. Generation nodes and consumption nodes are naturally given by the set of transmission outputs (LTFTRs) The network topology is described by the network incidence matrix (Léautier, 2000, p. 83). Given the topology there is a set of power transfer distribution factors (PTDF) that govern the flows on the individual lines. For a given network topology we assume that the line capacity is

[10] PTP forward obligations have been the primary financial instrument in practice, compared to PTP options and flowgate rights. PTP-FTR obligations can be either "balanced" or "unbalanced". A perfect hedge is achieved through a balanced PTP-FTR, while an unbalanced PTP-FTR obligation can be seen as a forward sale of energy. See Hogan (2002b)

variable so that, at a cost, it can be changed between 0 and ∞. There may be a fixed cost at zero capacity.

4.1 Example: Derivation of total costs and marginal costs

We analyze grid expansion costs in the seven-bus network example in Hogan (2000, pp. 7-17), based on the conventional DC load approximation. Ten lines are connected as shown in figure 1, where each line is assumed to have the same impedance (to simplify the illustration). Figure 1 shows the flows of 100 MWh/h from bus 1 to bus 7, and the implied table of distribution factors. Figure 2 illustrates a similar calculation for a transfer of 100 MWh/h from bus 3 to bus 7. What do costs look like for this example? Define:

$X = [q_{ij}]$ matrix of balanced point-to-point FTRs

x = vector of net injections

k = vector of line capacities

$f_i(k_i)$ = cost of building line i with capacity k_i

i = a vector of ones

H=PTDF matrix

$-Hx$=vector of line flows

There are no losses. The balanced FTRs ensure that total net injections are zero. The cost function is defined by the minimum costs necessary to produce each level of output, subject to feasibility constraints and the relationship between net injections and output. Since FTRs are the outputs, we have

Figure 1

Power Flow and Distribution Factors
(Bus 1 to Bus 7)

Line	Distribution Factor
1->2	1/2
1->3	1/2
2->4	1/6
2->6	1/3
3->4	1/6
3->5	1/3
4->5	1/6
4->6	1/6
5->7	1/2
6->7	1/2

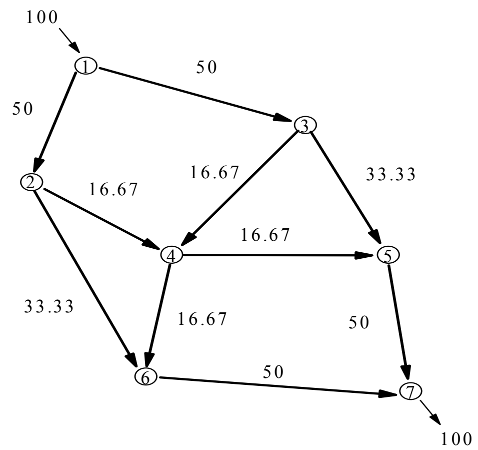
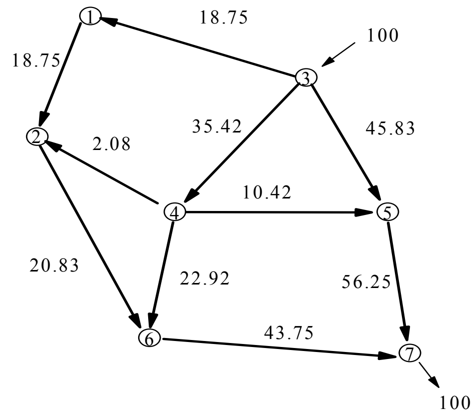


Figure 2

Power Flow and Distribution Factors
(Bus 3 to Bus 7)

Line	Distribution Factor
1->2	3/16
1->3	- 3/16
2->4	- 1/48
2->6	5/24
3->4	17/48
3->5	11/24
4->5	5/48
4->6	11/48
5->7	9/16
6->7	7/16



$$C(X) = \min_{k_i} \sum_{i=1}^{10} f_i(k_i) \qquad \text{s.t.} \tag{1}$$

$$-Hx \le k \tag{1a}$$

$$x = Xi \tag{1b}$$

An assumption here is that the costs of lines are separate from each other. This simplification can be relaxed. We will argue below that this is not a specific issue of transmission but would be similar in any type of network, such as natural gas transportation or telecommunications networks. Based on Figures 1 and 2 we take H from Hogan (2000, p. 13):

$$H = \begin{matrix} -1/2 & 3/16 & -3/16 & 0 & -1/16 & 1/16 & 0 \\ -1/2 & -3/16 & 3/16 & 0 & 1/16 & -1/16 & 0 \\ -1/6 & -17/48 & 1/48 & 1/6 & 1/16 & -1/16 & 0 \\ -1/3 & -11/24 & -5/24 & -1/6 & -1/8 & 1/8 & 0 \\ -1/6 & 1/48 & -17/48 & 1/6 & -1/16 & 1/16 & 0 \\ -1/3 & -5/24 & -11/24 & -1/6 & 1/8 & -1/8 & 0 \\ -1/6 & -11/48 & -5/48 & -1/3 & 3/16 & -3/16 & 0 \\ -1/6 & -5/48 & -11/48 & -1/3 & -3/16 & 3/16 & 0 \\ -1/2 & -7/16 & -9/16 & -1/2 & -11/16 & -5/16 & 0 \\ -1/2 & -9/16 & -7/16 & -1/2 & -5/16 & -11/16 & 0 \end{matrix}$$

Assuming separability of the investment function for each line, minimizing w.r.t. k_i on the right side of equation 1 implies that constraint (1a) is binding. Thus,

$$C(X) = \min\{\sum_{i=1}^{10} f_i(-H_i Xi)\} =$$

$$f_1(\frac{1}{2}\sum_i q_{1i} - \frac{3}{16}\sum_i q_{2i} + \frac{3}{16}\sum_i q_{3i} + \frac{1}{16}\sum_i q_{5i} - \frac{1}{16}\sum_i q_{6i}) + f_2(...) + ..f_{10}(...)\} \tag{2}$$

Note that the variables in parentheses are the k_i. For example,

$$k_1 = \frac{1}{2}\sum_i q_{1i} - \frac{3}{16}\sum_i q_{2i} + \frac{3}{16}\sum_i q_{3i} + \frac{1}{16}\sum_i q_{5i} - \frac{1}{16}\sum_i q_{6i}$$

Now, for a given network architecture of active nodes and lines (where we assume that only line capacities can be adjusted, but nodes, lines and impedances cannot be changed), H is given and thus the equation in (2) is unique and represents the minimum costs.[11]

Assume, as in Hogan (2000, pp. 7-17) that only nodes 1 and 3 have generation facilities and only node 7 has consumers. Thus, the only valuable FTRs are q_{17} and q_{37}. This implies

$$C(q_{17}) = f_1(\frac{1}{2}q_{17}) + f_2(\frac{1}{2}q_{17}) + f_3(\frac{1}{6}q_{17}) + f_4(\frac{1}{3}q_{17}) + f_5(\frac{1}{6}q_{17}) + f_6(\frac{1}{3}q_{17}) + f_7(\frac{1}{6}q_{17})$$

$$+ f_8\frac{1}{6}(q_{17}) + f_9(\frac{1}{2}q_{17}) + f_{10}(\frac{1}{2}q_{17})$$

$$C(q_{37}) = f_1(\frac{3}{16}q_{37}) + f_2(-\frac{3}{16}q_{37}) + f_3(-\frac{1}{48}q_{37}) + f_4(\frac{5}{24}q_{37}) + f_5(\frac{17}{48}q_{37}) + f_6(\frac{11}{24}q_{37})$$

$$+ f_7(\frac{5}{48}q_{37}) + f_8(\frac{11}{48}q_{37}) + f_9(\frac{9}{16}q_{37}) + f_{10}(\frac{7}{16}q_{37})$$

Because a negative quantity is the same as a positive quantity in the other direction, it makes sense to assume that for total flow q in a positive direction on a line we have $f_k(q) = f_k(-q) = f_k(|q|)$.

$$C(q_{17}, q_{37}) = f_1(\frac{1}{2}q_{17} + \frac{3}{16}q_{37}) + f_2(\left|\frac{1}{2}q_{17} - \frac{3}{16}q_{37}\right|) + f_3(\frac{1}{6}q_{17} - \frac{1}{48}q_{37}) + f_4(\frac{1}{3}q_{17} + \frac{5}{24}q_{37}) + f_5(\frac{1}{6}q_{17} + \frac{17}{48}q_{37}).....$$

Since we have assumed that the line cost functions are independent of each other and therefore simply add up, we can derive most of the properties of the transmission grid cost function by looking at an individual line cost function. If $f_l(k_l)$ were an affine linear function it would also be an affine linear function of all the q_{ij}. Thus, if $f_m(k) = a_m + b_m|k|$, then a segment of the piecewise cost function would be:

$$C(q_{17}, q_{37}) = a_1 + b_1(\frac{1}{2}q_{17} + \frac{3}{16}q_{37}) + a_2 + b_2(\frac{1}{2}q_{17} - \frac{3}{16}q_{37}) + a_3 + b_3(\left|\frac{1}{6}q_{17} - \frac{1}{48}q_{37}\right|)$$

[11] Thus, the optimal line capacity is a linear combination of the FTRs that we have postulated as outputs. The cost of line capacity will likely not exhibit constant returns to scale but rather have some setup costs for planning, etc. that are independent of scale so that there will be some scale economies. Scale economies can also be expected from the cost of land, poles and other items that increase less than proportionally with capacity (by the two-thirds rule). On the other hand, there will be physical limitations to line capacity that might lead to diseconomies of scale from some capacity onwards. As a result, the line cost function will be nonlinear and have both convex and concave parts.

$$+a_4+b_4(\frac{1}{3}q_{17}+\frac{5}{24}q_{37})+a_5+b_5(\frac{1}{6}q_{17}+\frac{17}{48}q_{37})+.....=A+B_{17}q_{17}+B_{37}q_{37}\ for\ \frac{1}{2}q_{27}>\frac{3}{16}q_{37}.$$

Even if there are no fixed costs for lines, there can be interaction in costs between different FTRs. This holds, when the quantities of q_{17} and q_{37} change in such a way that a sign change occurs for $\frac{1}{2}q_{17}-\frac{3}{16}q_{37}$ and/or $\frac{1}{6}q_{17}-\frac{1}{48}q_{37}$. For example, if only q_{37} increases in the neighbourhood of $\frac{1}{2}q_{17}=\frac{3}{16}q_{37}$ the cost curve of line 2 first decreases to zero and then increases, while if both FTRs increase in proportion $\frac{1}{2}q_{17}\Big/\frac{3}{16}q_{37}$ costs of line 2 stay flat at zero. Thus, counterflows can introduce nonlinearities in $C(X)$ even if all $f_m(k_m)$ are linear.

Interaction in costs becomes more widespread if line costs are nonlinear in k_m, because then there will be interactive terms between the q_{ij}. However, except if q_{ij} have negatively signed coefficients, these interactive terms will generally be quite similar to interactive terms in other multi-product cost functions. In particular, if $f_m(k_m)$ is in a range where economies of scale prevail, then economies of scale will prevail for individual q_{ij} and economies of scope will prevail between different q_{ij}'s. For example, for a line cost function of the form $f_m(k_m) = a_mk - b_mk^2$ (in the range, where $f_m(..) > 0$ and $f_m' > 0$) we get for line 1 $f_1(q_{17},q_{37})=a_1(\frac{1}{2}q_{17}+\frac{3}{16}q_{37})-b_1(\frac{1}{2}q_{17}+\frac{3}{16}q_{37})^2$ and for line 2 $f_2(q_{17},q_{37})=a_2(\frac{1}{2}q_{17}-\frac{3}{16}q_{37})-b_2(\frac{1}{2}q_{17}-\frac{3}{16}q_{37})^2$. Thus, there is an interactive term for line 1 of the form $-\frac{3}{16}q_{17}q_{37}$, which produces economies of scope, and for line 2 of the form $+\frac{3}{16}q_{17}q_{37}$, which produces diseconomies of scope but cannot overcome the economies of scope from the squared terms.

Similarly, diseconomies of scale are associated with diseconomies of scope, unless the sign of the relevant coefficient of H is negative. For example, in case of quadratic line cost functions of the form $f_m(k_m) = a_mk^2$ we get for line 1 $f_1(q_{17},q_{37})=a_1(\frac{1}{2}q_{17}+\frac{3}{16}q_{37})^2$ and for line 2

$f_2(q_{17}, q_{37}) = a_2(\frac{1}{2}q_{17} - \frac{3}{16}q_{37})^2$. Thus, there is an interactive term for line 1 of the form $+\frac{3}{16}q_{17}q_{37}$, which produces diseconomies of scope, and for line 2 of the form $-\frac{3}{16}q_{17}q_{37}$, which produces economies of scope but cannot overcome the diseconomies of scope from the squared terms .

Equation 2 above implies marginal costs of the form

$$\frac{\partial C(X)}{\partial q_{17}} = \frac{1}{2}\frac{\partial f_1}{\partial k_1} + \frac{1}{2}\frac{\partial f_2}{\partial k_2} + \frac{1}{6}\frac{\partial f_3}{\partial k_3} + \frac{1}{3}\frac{\partial f_4}{\partial k_4} + \frac{1}{6}\frac{\partial f_5}{\partial k_5} + \frac{1}{3}\frac{\partial f_6}{\partial k_6} + \frac{1}{6}\frac{\partial f_7}{\partial k_7} + \frac{1}{6}\frac{\partial f_8}{\partial k_8} + \frac{1}{2}\frac{\partial f_9}{\partial k_9} + \frac{1}{2}\frac{\partial f_{10}}{\partial k_{10}} \quad \textbf{(3a)}$$

$$\frac{\partial C(X)}{\partial q_{37}} = \frac{3}{16}\frac{\partial f_1}{\partial k_1} - \frac{3}{16}\frac{\partial f_2}{\partial k_2} - \frac{1}{48}\frac{\partial f_3}{\partial k_3} + \frac{5}{24}\frac{\partial f_4}{\partial k_4} + \frac{17}{48}\frac{\partial f_5}{\partial k_5} + \frac{11}{24}\frac{\partial f_6}{\partial k_6} + \frac{5}{48}\frac{\partial f_7}{\partial k_7} + \frac{11}{48}\frac{\partial f_8}{\partial k_8} + \frac{9}{16}\frac{\partial f_9}{\partial k_9} + \frac{7}{16}\frac{\partial f_{10}}{\partial k_{10}} \quad \textbf{(3b)}$$

Note that the positive signs in (3a) and (3b) only hold, as long as there are no counterflows on those lines, while the negative signs in (3b) only holds if there are sufficient counterflows on lines 2 and 3 (e.g., at zero net injections).

Since the f_i(.) are the costs of individual lines, they are monotonically increasing functions (in the positive range, and decreasing functions in the negative range). The marginal costs of FTRs are piecewise linear combinations of the marginal costs of all lines. The weights of these linear combinations are constant. As a result, marginal costs of FTRs should be well-behaved over some range.

5 Transmission demand functions

The demand for transmission services is derived from the demand for electricity by loads and the supply of electricity by generators (and possibly by supply and demand by intermediaries). In the simplest case of a single line with perfectly competitive generators at one node and competitively purchasing loads at the other node the demand for transmission is simply the vertical difference between demand by the loads and supply by the generators. However, the simplicity vanishes if there are multiple generation and consumption nodes (and imperfect competition). For example, transmission demand functions could exhibit discontinuities if one allows for all possible price combinations. This is due to the fact that electricity is a homogeneous good so that consumers will move fully from one source (represented by a generation node) to another (represented by another generation node) if the second becomes cheaper than the first, including transmission fees. Thus, if one sufficiently reduces the transmission fees from the second node, keeping the transmission fees from the first node constant, one can always generate a discontinuity in the transmission demands from both nodes. Such discontinuities, however, need not bother the Transco, as long as the law of one price holds at all nodes. This means that transmission fees fulfill an arbitrage function so that electricity at each consumption node costs the same independent of the generation node it comes from; and each generation node receives the same for its electricity, independent of which consumption node it supplies. Nodal prices calculated ex post by an ISO could have the exact market-clearing property.

How does the Transco get the demand information necessary to optimize investments in transmission capacity? Since transmission demands are derived from the electricity demands at consumption nodes and the electricity supplies at the generation nodes, their calculation principally involves a simultaneous estimation of market equilibria in all electricity markets involved. This process can be simplified if generators and loads act perfectly competitively. The Transco would have to know the supply functions at the generation nodes and the demand functions at the consumption nodes. Assume there are L generation nodes indexed by subscript '*l*' and M demand nodes indexed by subscript '*m*'. Also assume that, at each generation node, generators supply electricity competitively and, at each demand node, the ultimate buyers demand electricity competitively.[12] The demand and supply functions are further all assumed to be independent of each other. The Transco knows these supply and demand functions. The Transco can now find the set of transmission demand functions for point-to-point transmissions by maximizing total surplus net of transmission charges.

Given an arbitrary set of transmission prices between locations, from *i* to *j* ($\tau = \left(\tau_{ij}\right)$), choose

[12] Even though loads aggregate consumers, they remain price takers for electricity generation and transmission services.

Max $W(\{q_{lm}\}) = \Sigma_m CS_m(q_m)$ - $\Sigma_l C_l(q_l) - \Sigma_l \Sigma_m \tau_{lm} q_{lm}$ s.t. $\Sigma_l q_{lm} = q_m$ and $\Sigma_m q_{lm} = q_l$ **(4)**

Here $CS_m(q_m)$ is consumer surplus for electricity at node m and $C_l(q_l)$ is the area under the supply curve for electricity at node l. Maximizing w.r.t. all q_{lm} gives lxm first order conditions of the form $p_m - p_l = \tau_{lm}$ (where p_m and p_l are the prices at nodes m and l, respectively) By substituting the electricity supply and demand functions for the p's provides lxm equations in lxm unknowns that can normally be solved for the q_{lm}'s as functions of the τ_{lm}'s. This yields the vector net demand function $q(\tau)$. This formulation characterizes the demand curve, even though (4) would net be a practical computational method with arbitrary prices. In order to obtain an equilibrium solution it would be necessary to impose the network constraints in the dispatch framework. In most markets this would be unrealistic. In electricity, it is necessary and therefore common practice.

A solution of (4) could be obtained efficiently through the same method as the dispatch. Given the supply offers and demand bids, we would replace the fixed transmission prices with the characterization of the transmission constraints. This would be of the form of the standard FTR auction. The resulting differences in nodal prices would be the associated transmission prices.

5.1 Examples

Following the above example from Hogan (2000, pp. 7-17), we are only interested in the transmission demands for the LTFTRs q_{17} and q_{37} with $q_7 = q_{17} + q_{37}$.

Example 1: Assume that the (inverse) electricity demand at node 7 is $p_7 = a - bq_7$ and that (inverse) electricity supply at node 1 is $p_1 = c_1 + d_1 q_{17}$, and at node 3 $p_3 = c_3 + d_3 q_{37}$. We require $q_{17} \geq 0$ and $q_{37} \geq 0$. We now want to maximize

$$W(q_{17}, q_{37}) = \int_0^{\tilde{q}_7} (a - bq_7) dq_7 - \int_0^{\tilde{q}_{17}} (c_1 + d_1 q_{17}) dq_{17} - \int_0^{\tilde{q}_{37}} (c_3 + d_3 q_{37}) dq_{37}$$

$$- \tau_{17} q_{17} - \tau_{37} q_{37} - \lambda (q_7 - q_{17} - q_{37}) - \mu_{17} q_{17} - \mu_{37} q_{37} \quad \textbf{(5)}$$

After eliminating the term with λ by imposing the equality condition $q_7 = q_{17} + q_{37}$, the first order conditions are

$$\frac{\partial L}{\partial q_{17}} = a - c_1 - (b + d_1)q_{17} - bq_{37} - \mu_{17} - \tau_{17} = 0 \qquad \textbf{(6a)}$$

$$\frac{\partial L}{\partial q_{37}} = a - c_3 - (b + d_3)q_{37} - bq_{17} - \mu_{37} - \tau_{37} = 0 \qquad \textbf{(6b)}$$

$$\frac{\partial L}{\partial \mu_{17}} = q_{17} \geq 0 \text{ and } \frac{\partial L}{\partial \mu_{37}} = q_{37} \geq 0 \qquad \textbf{(6c)}$$

Within the regions, where the non-negativity constraints are not binding, the first two f.o.c's give us the two demand functions for FTRs q_{17} and q_{37}. Note that the two FTRs are substitutes in demand. The demand functions are most easily expressed by differentiating between the ranges where quantities are all strictly positive, and where at least one of them is zero.

In the range where q_{17} and q_{37} are both positive we get

$$q_{17} = \frac{(ad_3 - bc_1 - d_3c_1 + bc_3) - (b + d_3)\tau_{17} + b\tau_{37}}{bd_1 + bd_3 + d_1d_3} \qquad \textbf{(7a)}$$

and

$$q_{37} = \frac{(ad_1 - bc_3 - d_1c_3 + bc_1) - (b + d_1)\tau_{37} + b\tau_{17}}{bd_1 + bd_3 + d_1d_3} \qquad \textbf{(7b)}$$

In the range of τ_{17} and τ_{37} where $q_{37} \leq 0$ we have

$$q_{17} = \frac{(a - c_1) - \tau_{17}}{b + d_1} \qquad \textbf{(7c)}$$

while in the range where $q_{17} \leq 0$ we have

$$q_{37} = \frac{(a - c_3) - \tau_{37}}{b + d_3} \qquad \textbf{(7d)}$$

Note that, except for the kink at the point where the demand for the other FTR becomes zero, these are very ordinary linear demand functions. The parallel shift due to substitutability is symmetric.

<u>Example 2:</u> Instead of the smoothly upward-sloping electricity supply function used above, electricity generation is often modeled with a stepwise rather than linear supply function. In order

to capture this feature we assume in a second example that node 1 has low-cost generation with a limited capacity $\overline{q}_1$. In contrast, node 3 has high-cost generation with unlimited capacity. Thus, the inverse supply curve for node 1 is $p_1 = c_1$ for $0 \le q_1 \le \overline{q}_1$, while for node 3 it is $p_3 = c_3$ for $0 \le q_3 \le \infty$. We further assume $c_1 < c_3$. These assumptions imply

$$W(q_{17}, q_{37}) = \int_0^{\tilde{q}_7} (a - bq_7) dq_7 - \int_0^{\tilde{q}_{17}} (c_1) dq_{17} - \int_0^{\tilde{q}_{37}} (c_3 \ ) dq_{37}$$

$$- \tau_{17} q_{17} - \tau_{37} q_{37} - \lambda (q_7 - q_{17} - q_{37}) - \mu_{17} q_{17} - \mu_{37} q_{37} - \gamma_{17} (q_{17} - \overline{q}_1) \tag{8}$$

After eliminating the term with λ by imposing the equality condition, the first order conditions are

$$\frac{\partial L}{\partial q_{17}} = a - c_1 - bq_{17} - bq_{37} - \mu_{17} - \tau_{17} - \gamma_{17} = 0 \tag{9a}$$

$$\frac{\partial L}{\partial q_{37}} = a - c_3 - bq_{37} - bq_{17} - \mu_{37} - \tau_{37} = 0 \tag{9b}$$

$$\frac{\partial L}{\partial \mu_{17}} = q_{17} \ge 0 \text{ and } \frac{\partial L}{\partial \mu_{37}} = q_{37} \ge 0 \text{ and } \frac{\partial L}{\partial \gamma_{17}} = q_{17} - \overline{q}_1 \le 0 \tag{9c}$$

This implies the following <u>inverse</u> demand functions of the two FTRs

$$\tau_{17} = a - c_1 - bq_{17} - bq_{37} \tag{10a}$$

$$\tau_{37} = a - c_3 - bq_{37} - bq_{17} \tag{10b}$$

for $0 \le q_{17} \le \overline{q}_1$ and $0 \le q_{37} \le \infty$.

Note that in this case the direct demand functions have some discontinuities. The principal condition is that q_{17} and q_{37} can both be positive only if $\tau_{17} + c_1 = \tau_{37} + c_3$. If the r.h.s. is smaller than the l.h.s. only q_{17} can be positive and if the l.h.s. is smaller than the r.h.s. only q_{37} can be positive. If the equality holds the size of both FTR's is indeterminate except that q_{17} is constrained and both have to add up to q_7, which itself depends on the sum of marginal generation cost and transmission charge.

To the extent that the inequality $\tau_{17} + c_1 < \tau_{37} + c_3$ holds, we get $\frac{\partial q_{17}}{\partial \tau_{17}} = -\frac{1}{b} and \frac{\partial q_{37}}{\partial \tau_{37}} = 0$ and, vice versa, for $\tau_{17} + c_1 > \tau_{37} + c_3$ we get $\frac{\partial q_{37}}{\partial \tau_{37}} = -\frac{1}{b}$ and $\frac{\partial q_{17}}{\partial \tau_{17}} = 0$. Where the inequalities hold the two cross derivatives are zero, while at the equality they are undefined (infinite).

6 The regulatory model

The regulatory model uses FTRs as the definition of output. In principle, this simplifies the problem for the Transco, and provides a link to the merchant model by employing the critical output definition. The regulator then sets a price index constraint using some set of weights. In general the index would be adjusted to reflect productivity and inflation effects. Here we simplify by assuming stable costs and demand conditions, and consider the repeated application of the incentive mechanism with a myopic Transco optimizing in each period.

6.1 The sequencing of moves

In applying Vogelsang (2001) without FTRs the sequence of moves would be

(1) A pre-existing network and point-to-point transmission prices that the Transco has charged up to the present are in place.

(2) The regulator sets the regulatory pricing constraint.

(3) The Transco collects information about generation supply and electricity demand at all relevant geographical locations (or at each node).

(4) The Transco invests in grid capacity.

(5) The Transco sets point-to-point transmission prices.

(6) Generators and loads sign bilateral electricity contracts and buy point-to-point transmission services.

(7) There can be excess supply or excess demand for transmission services on a point-to-point basis. Excess supply could hurt the Transco but would not cause any feasibility problems. Excess demand could cause feasibility problems (although, together with excess supply for other point-to-point relationships the total sum might still be feasible). The Transco could then use non-price rationing and sell point-to-point transmission services to bilaterals on a first-come-first-serve basis. Regulators could impose penalties, giving the Transco incentives to price in such a way that excess demand does not occur (creating excess supply).

(8) The Transco calculates the fixed fee from the regulatory constraint and charges it to the loads.

Alternatively, moves (5)-(8) could be replaced by:

(5a) There is an ISO, who asks for (sequences of) bids from generators and loads at each node and then calculates nodal prices. Loads (ex post) pay the ISO and generators receive payment in such a way that markets always clear. The Transco receives as congestion payments the difference between what loads pay and what generators receive. Fixed fees are then calculated from the regulatory constraint and are paid by the loads. In this case the Transco does not set prices but only makes available capacities.

Under the FTR mechanism described above in the institutional setup of Section 2 the sequence would include some extra steps. Corresponding to (1)-(8) it would look:

(1) A pre-existing network and point-to-point transmission prices that the Transco has charged up to the present are in place.

(2) The regulator sets the regulatory pricing constraint.

(3) The Transco collects information about generation supply and electricity demand at all relevant geographical locations (or at each node).

(4) The Transco invests in grid capacity.

(5) The Transco auctions off point-to-point FTRs, based on the available grid capacity.

(6) There is an ISO, who asks for (sequences of) bids from generators and loads at each node and then calculates nodal prices. Loads (ex post) pay the ISO according to their last bids and generators receive payment of their last bids in such a way that markets always clear. The owners of FTRs receive as congestion payments the difference between what loads pay and what generators receive. Any excess congestion payments that cannot be allocated to an FTR (because less FTRs were sold than the point-to-point transmission available), go to the Transco. If FTRs are presented, for which there are no point-to-point "flows" to collect, congestion charges from the charges are paid by the Transco (could only happen if capacity is less than FTRs were sold, otherwise the congestion charge would be zero.).

(7) Fixed fees are then calculated from the regulatory constraint, based on congestion charges, and are paid by the loads. In this case the Transco does not set prices but only makes available capacities.[13]

6.2 A More General Framework

Assume a transmission network with several nodes, and an institutional structure of a Transco (combined with an ISO) that sells LTFTRs in order to carry out transmission expansion projects.[14] These projects possibly involve large and lumpy meshed networks. LTFTRs are assumed to be point-to-point balanced financial transmission right obligations. There are various established

[13] We assume (as in equation 12 below) that the fixed fees go to the Transco. That means the auction only yields revenues based on the expected congestion charges. We also assume that the grid investment is common knowledge, and that the participants in the auction are well informed about electricity demands and generation supplies. Likewise, the auction prices are the prices in the Transco's optimization problem in equation 12 below.

[14] We abstract from the other markets such as the market for bilateral contracts, and the market for capacity reserves.

agents (generators, Gridcos, marketers, etc.) interested in the transmission grid expansion that do not have market power in their respective markets or in "adjacent" markets.[15]

There is a sequence of auctions at each period t where participants buy and sell LTFTRs[16] (and therefore possibly reconfigure the existing FTR allocations), culminating in a real time auction at which time all FTRs are cashed out. We assume that no valuable FTRs remain unallocated at each t.[17] The Transco seeks to maximize expected profits at each auction subject to simultaneous feasibility constraints, and a two-part tariff cap constraint.[18]. We use regulation of the Transco's price structure to promote adequate signals for efficient expansion. The Transco carries out a long-run intertemporal maximization that also considers recovery of fixed costs. The transmission outputs are the incremental LTFTRs between consecutive periods.

Suppose we define the problem using the least cost solution for the network configuration that meets a given demand. Over the domain, where $\iota^t q = 0$, let

$$c^*\left(q, K^{t-1}, H^{t-1}\right) = \underset{K^t \in \mathrm{K}, H^t \in \mathrm{H}}{Min} \left\{ c\left(K^t, K^{t-1}, H^t, H^{t-1}\right) \middle| H^t q \leq K^t \right\}. \quad \textbf{(11)}$$

where:

---

[15] Joskow and Tirole (2000) analyze the welfare implications of market power in the markets of financial and physical transmission rights.
[16] LTFTRs are sold in each auction for the total length of the periods.
[17] This assumption eliminates the need for proxy awards. Note however that there are always unallocated FTRs (e.g. due to change in flows) but not all of them are "valuable". Hogan (2002a) and Kristiansen and Rosellón (2006) assume unallocated capacity and FTRs that permit the ISO to handle negative externalities due to an expansion project through proxy awards.
[18] For a joint energy and transmission rights auction, and assuming a fixed transfer matrix, O´Neill et al (2002) show that revenue adequacy is met whenever capacity does not decrease over periods. The simultaneous feasibility constraints guarantee that all payments required under the LTFTR obligations are met whenever capacity in period $t+1$ (K^{t+1}) is greater than capacity at period t (K^t).

q^t = the net injections in period t (FTRs are derived from : $\sum_j \tau_j^t = q^t \, ; \tau_j^t = \begin{bmatrix} -x \\ 0 \\ 0 \\ . \\ . \\ +x \\ 0 \end{bmatrix}$)[19]

K^t = available transmission capacity in period t

H^t = transfer admittance matrix at period t

ι^t = a vector of ones

Here, $c\left(K^t, K^{t-1}, H^t, H^{t-1}\right)$ is the cost of going from one configuration to the next. The cost level can also be affected by a change in flows only, even when there is no line capacity expansion (note that this is not likely to depend only on the difference in the configuration; e.g., there are economies of scale and scope). At each time t, when expansion takes place matrix H will be affected due to possible changes in the geometry of the network including changes in impedances.[20] Here the output is still defined in terms of incremental FTRs.[21]

For a load-flow model for real power (DC load approximation), the Transco's profit maximization problem is given by:

$$\underset{t^t, F^t}{Max} \pi^t = \tau^t \left(q(\tau^t) - q^{t-1}\right) + F^t N^t - c^*\left(q(\tau^t), K^{t-1}, H^{t-1}\right) \tag{12}$$

subject to

$$\tau^t Q^w + F^t N^t \le \tau^{t-1} Q^w + F^{t-1} N^t \tag{13}$$

---

[19] 'q' refers to net injections of the form q^i, while the FTRs are of the form q^{ij}. The FTRs form a matrix $Q = [q^{ij}]$ so that the vector of net injections is $q = Qe$, where e is a unit vector. Since we are assuming that FTRs are point-to/point obligations, we can indistinctively use net injections or FTRs as output (see Hogan, 2002b).

[20] As in O´Neill et al (2002), the constraint on H could alternatively be defined under the assumption of a fixed network topology (e.g., $H(q^t, Z^t) \equiv H(z^t)q^t \le K^t, \forall t$). We chose now the more general approach of allowing changes in the geometry of the network due to an expansion project.

[21] See Gribik et al. (2004) for a discussion of separate rights associated with impedances and prices for impedances as in O'Neill et al. (2005).

where:

τ^t = vector of transmission prices between locations in period t

F^t = fixed fee in period t

N^t = number of consumers in period t

$Q^w = (q^t - q^{t-1})^w$

w = type of weight.

Equation (12) provides the profit function of a Transco that auctions LTFTRs at each period t. Incomes of the Transco are variable ($\tau^t(q(\tau^t) - q^{t-1})$: income from LTFTRs) and fixed (F^tN^t: income from fixed charges to consumers). The cost c of going from one configuration to the next in (1) depends on FTRs, capacity K, and the transfer admittance matrix H.

As in Vogelsang (2001), the proposed price cap index (13) is defined on two-part tariffs: a variable fee τ^t and a fixed fee F. However, the output is now incremental LTFTRs. The weighted number of consumers N^t is assumed to be determined exogenously. Note that period (t-1) basically provides prices, quantities and costs that are needed for regulation in the next period.

If the demand and optimized cost functions are differentiable,[22] the first order optimality conditions are:

$$\nabla q(\tau - \nabla_q c^*) = Q^w - (q(\tau) - q^{t-1}) \qquad \textbf{(14)}$$

The analysis of Vogelsang (2001) points to this key relationship in analyzing the incentive properties of the regulatory price cap constraints. In the case of transmission expansion, the

[22] The demand function is probably well behaved. The properties of $c^*\left(q\left(\tau^t\right), K^{t-1}, H^{t-1}\right)$ remain an open question. Note, that by construction we always have $t^t q\left(\tau\right) = 0$. It is reasonable to assume $c^*\left(q\left(\tau^t\right), K^{t-1}, H^{t-1}\right) < \infty$, so there is always a solution. It should be also true that $c^*\left(q\left(\tau^t\right), K^{t-1}, H^{t-1}\right)$ is piecewise partially differentiable in q almost everywhere. In general, it is not likely that $Hq = K$. There are many contingency constraints in security constrained dispatch. In addition, the feasible sets of topologies, impedances and capacities is not convex, and the cost function itself is not likely to be convex. Hence the optimized cost function $c^*\left(q\left(\tau^t\right), K^{t-1}, H^{t-1}\right)$ could be very complicated.

demand and cost functions are only piecewise differentiable, and in general they are not separable. However the local properties might in many circumstances be those of well-behaved functions, and this can give some insight about the direction of the incentive effects

6.2.1 Example

Consider again the example in Hogan (2000, pp. 7-17), we reformulate (12) through (13) as:

$$\underset{\tau_{17}^t,\tau_{37}^t,F^t}{Max}\ \tau_{17}^t(q_{17}^t(\tau^t)-q_{17}^{t-1})+\tau_{37}^t(q_{37}^t(\tau^t)-q_{37}^{t-1})+F^tN^t-c^*(q(\tau^t),K^{t-1},H^{t-1}) \quad \textbf{(12a)}$$

subject to

$$\tau_{17}^tQ_{17}^w+\tau_{37}^tQ_{37}^w+F^tN^t\le\tau_{17}^{t-1}Q_{17}^w+\tau_{37}^{t-1}Q_{37}^w+F^{t-1}N^t \quad \textbf{(13a)}$$

Assuming that the constraint is binding, equation (13a) can be substituted into term F^tN^t of the objective function. The maximization problem then turns into:

$$\underset{\tau_{17}^t,\tau_{37}^t}{Max}\ \tau_{17}^t(q_{17}^t(\tau^t)-q_{17}^{t-1})+\tau_{37}^t(q_{37}^t(\tau^t)-q_{37}^{t-1})+(\tau_{17}^{t-1}-\tau_{17}^t)Q_{17}^w+(\tau_{37}^{t-1}-\tau_{37}^t)Q_{37}^w-c^*(q(\tau^t),K^{t-1},H^{t-1}) \quad \textbf{(15)}$$

The two first-order conditions of (15) lead to:

$$(\tau_{17}^t-\frac{\partial c^*}{\partial q_{17}^t})\frac{\partial q_{17}^t}{\partial\tau_{17}^t}+(\tau_{37}^t-\frac{\partial c^*}{\partial q_{37}^t})\frac{\partial q_{37}^t}{\partial\tau_{17}^t}=Q_{17}^w-q_{17}^t+{q^{t-1}}_{17} \quad \textbf{(16a)}$$

and

$$(\tau_{37}^t-\frac{\partial c^*}{\partial q_{37}^t})\frac{\partial q_{37}^t}{\partial\tau_{37}^t}+(\tau_{17}^t-\frac{\partial c^*}{\partial q_{17}^t})\frac{\partial q_{17}^t}{\partial\tau_{37}^t}=Q_{37}^w-q_{37}^t+{q^{t-1}}_{37} \quad \textbf{(16b)}$$

Note that we need to allow for expansion and contraction in order to keep the FTRs on the cost function. All costs would then be instantaneous costs. We will initially consider two types of weights: chained Laspeyres weights (*w* = *t-1*) and idealized weights (*w* = *). As discussed above, Laspeyres weights are easily calculated and have shown nice economic properties under stable cost and demand conditions (see Ramírez and Rosellón, 2002). Idealized weights correspond to perfectly predicted quantities and posses strong efficiency properties (Laffont and Tirole, 1996).

6.2.1.1 Case one: Idealized weights

If we use idealized weights with $Q_{17}^{w} = q_{17}^{*} - q_{17}^{t-1}$ and $Q_{37}^{w} = q_{37}^{*} - q_{37}^{t-1}$ then we get

$$(\tau_{17}^{t} - \frac{\partial c^{*}}{\partial q_{17}^{t}})\frac{\partial q_{17}^{t}}{\partial \tau_{17}^{t}} + (\tau_{37}^{t} - \frac{\partial c^{*}}{\partial q_{37}^{t}})\frac{\partial q_{37}^{t}}{\partial \tau_{17}^{t}} = q_{17}^{*} - q_{17}^{t} \tag{17a}$$

and

$$(\tau_{37}^{t} - \frac{\partial c^{*}}{\partial q_{37}^{t}})\frac{\partial q_{37}^{t}}{\partial \tau_{37}^{t}} + (\tau_{17}^{t} - \frac{\partial c^{*}}{\partial q_{17}^{t}})\frac{\partial q_{17}^{t}}{\partial \tau_{37}^{t}} = q_{37}^{*} - q_{37}^{t} \tag{17b}$$

These two necessary are satisfied for prices to equal marginal costs. Under the regulatory constraint and a unique solution, idealized weights are sufficient for transmission nodal prices to equal marginal costs whenever

$$(\frac{\partial q_{17}^{t}}{\partial \tau_{17}^{t}}) \cdot (\frac{\partial q_{37}^{t}}{\partial \tau_{37}^{t}}) \neq (\frac{\partial q_{37}^{t}}{\partial \tau_{17}^{t}}) \cdot (\frac{\partial q_{17}^{t}}{\partial \tau_{37}^{t}}) \tag{17}$$

This does not appear to be a very restrictive condition.[23]

6.2.1.2 Case two: Last period's quantities as weights

If we use last period's weights with $Q_{17}^{w} = q_{17}^{t-1} - q_{17}^{t-2}$ and $Q_{37}^{w} = q_{37}^{t-1} - q_{37}^{t-2}$, the first order conditions imply

$$(\tau_{17}^{t} - \frac{\partial c^{*}}{\partial q_{17}^{t}})\frac{\partial q_{17}^{t}}{\partial \tau_{17}^{t}} + (\tau_{37}^{t} - \frac{\partial c^{*}}{\partial q_{37}^{t}})\frac{\partial q_{37}^{t}}{\partial \tau_{17}^{t}} = (q_{17}^{t-1} - q_{17}^{t-2}) - (q_{17}^{t} - q_{17}^{t-1}) \tag{18a}$$

and

$$(\tau_{37}^{t} - \frac{\partial c^{*}}{\partial q_{37}^{t}})\frac{\partial q_{37}^{t}}{\partial \tau_{37}^{t}} + (\tau_{17}^{t} - \frac{\partial c^{*}}{\partial q_{17}^{t}})\frac{\partial q_{17}^{t}}{\partial \tau_{37}^{t}} = (q_{37}^{t-1} - q_{37}^{t-2}) - (q_{37}^{t} - q_{37}^{t-1}) \tag{18b}$$

Note that the right hand side of equations 18a and 18b represents the negative of the growth rates of the incremental FTRs. Given the assumption of stationary cost and demand functions positive growth in period *t* is associated with a reduction in price, while negative growth can also be associated with a reduction in price as long as $q^{t} - q^{t-1} > 0$. We first consider the conditions for a

[23] With invertible demand, $\det(\nabla q) \neq 0$,

positive growth rate. Assume that the cross-derivatives have the same sign. In fact, we would normally assume that they are equal to each other so that the integrability conditions hold. This would be true in the absence of income effects. Now, if they are complements and if prices are above marginal costs the current transmission quantities will exceed last period's quantities, which means that prices have been lowered. If, as in the examples above, the two services are substitutes we are only sure to get this effect if the cross effects are smaller than the direct effects. If prices are below marginal costs we get the opposite results. So, we get a closer approximation of prices to marginal costs unless cross effects are too large.

Now consider the case of a negative growth rate but still a positive incremental FTR in period t. This would hold if

$$(\tau_{17}^t - \frac{\partial c^*}{\partial q_{17}^t})\frac{\partial q_{17}^t}{\partial \tau_{17}^t} + (\tau_{37}^t - \frac{\partial c^*}{\partial q_{37}^t})\frac{\partial q_{37}^t}{\partial \tau_{17}^t} - (q_{17}^{t-1} - q_{17}^{t-2}) = -(q_{17}^t - q_{17}^{t-1}) < 0 \qquad \textbf{(19a)}$$

and

$$(\tau_{37}^t - \frac{\partial c^*}{\partial q_{37}^t})\frac{\partial q_{37}^t}{\partial \tau_{37}^t} + (\tau_{17}^t - \frac{\partial c^*}{\partial q_{17}^t})\frac{\partial q_{17}^t}{\partial \tau_{37}^t} - (q_{37}^{t-1} - q_{37}^{t-2}) = -(q_{37}^t - q_{37}^{t-1}) < 0 \qquad \textbf{(19b)}$$

Since last period's incremental FTRs are nonnegative, the conditions for prices to fall are actually weaker than those expressed in the previous paragraph; and they are weaker the larger last period's incremental FTRs are.

In summary, the two cases show that the process is likely to do well with idealized weights under fairly general conditions while this is not assured under Laspeyres weights

7 Merchant Transmission

The use of FTRs as the output definition links the regulatory model to the basics of merchant transmission investment. In effect merchant investment is like the regulated investment without the guarantee of the fixed charge or price index constraint.

7.1 Open Entry

A conceptual step further away from Vogelsang (2001) would be a merchant transmission approach based on the same basic regulatory constraint. It would start with an existing grid that could be owned by a Transco and add a free-entry feature for grid extensions (including deepening investments). Any additional grid capacity added would be entitled to revenues from FTRs on additional point-to-point transmissions made possible by this capacity expansion.[24] One problem is to calculate the additional outputs attributable to new generations of capacity. This is, however, a standard problem of the merchant transmission approach. In contrast, the current approach is compatible with economies of scale and scope that would be typical for transmission. There remains the issue of loop flow if that leads to diseconomies of scope. Under this mechanism, a merchant would have an incentive to invest if the congestion revenue to be expected from the additional capacity exceeds its cost or if average congestion revenue exceeds average cost of the new capacity. If there are no economies of scale for the new capacity one would get to the optimal capacity in a single step, whereas the pure Transco approach would only get there half way.

Joskow and Tirole (2005) critique proposed merchant investment schemes for transmission grids to a set of criticisms. An issue of particular importance is that of lumpiness or economies of scale. Joskow and Tirole show that merchant investment based on nodal pricing as the investment reward could not hope to recover optimal lumpy investment because the nodal prices would fall, as a result of the investment and therefore marginal prices could not cover average cost of such investment. By using a two-part tariff scheme the above merchant investment approach could, in principle, do better than simple linear nodal pricing.

7.2 Irreversible investments

If we assume irreversible investments in the sense of pure sunk costs of line capacities (with infinite lives), capacities would only be expanded, not contracted. Thus, some transmission demand

[24] This requires duplicate calculations of all nodal prices, both with and without that additional capacity. Alternatively, flowgate rights could be calculated with the same information.

would have to grow in order to induce a change in capacities. This could relate to changes in generation supply (e.g., retirement of a generation plant or reduced generation costs at all plants) or increased electricity demands. If we assume a system that is currently optimized and there is an exogenous shock we would likely require some line expansion, while the resulting system would not be optimal compared to one that was originally designed for the new transmission demand. The cost of expansion would definitely be positive. We could also look at fluctuating transmission demands and capacities that can only be adjusted upwards and only over longer periods. The question then would be what the optimal capacity is for coping with peak problems, while optimizing capacity utilization the rest of the time.

8 Conclusion

In this paper we address regulatory approaches to electricity transmission expansion in a manner that is compatible with merchant investment in the context of price-taking generators and loads. The regulatory model is an extension of Vogelsang (2001) for meshed projects. Transmission output is defined in terms of incremental LTFTRs for lumpy and large transmission projects. With idealized weights, as well as under Laspeyres weights, we are able to identify the conditions for marginal cost pricing. This is a step in a research agenda. Principle questions include how to practically characterize the piecewise cost functions, incorporate changes in topology and how to address the global rather than local optimality properties of the incentives.

9 References

Brunekreeft, G., K. Neuhof, and D. Newbery, (2005). "Electricity Transmission: An Overview of the Current Debate," Utilities Policy, Elsevier, Vol. 13, No. 2, pp. 73-93, June.

Bushnell, J. B., and S. E. Stoft (1997) “Improving Private Incentives for Electric Grid Investment,” *Resource and Energy Economics* 19, 85-108.

California ISO and London Economics International (2003),”A Proposed Methodology for Evaluating the Economic Benefits of Transmission Expansions in a Restructured Wholesale Electricty Market,” mimeo, (http://www.caiso.com/docs/2003/03/25/2003032514285219307).

Gans, J. S. and S. P. King (1999), “Options for Electricity Transmission Regulation in Australia,” University of Melbourne, September 10.

Grande, O. S., and I. Wangensteen (2000) “Alternative Models for Congestion Management and Pricing Impact on Network planning and Physical Operation,” *CIGRE*, Paris, aug/sept.

Gribik, P. R., D. Shirmohammadi, J.S. Graves, J.G. Kritikson (2004), “Transmission Rights and Transmission Expansion,” IEEE Transactions on Power Systems, Vol. 20, No. 4 November 2004, pp. 1728-1737.

Harvey, S. M. (2002) “TCC Expansion Awards for Controllable Devices: Initial Discussion,” Mimeo.

Hogan, W. (2002a) “Financial Transmission Right Incentives: Applications Beyond Hedging.” Presentation to HEPG Twenty-Eight Plenary Sessions, May 31, http://www.ksg.harvard.edu/people/whogan.

Hogan, W. (2002b) “Financial Transmission Right Formulations,” Mimeo, JFK School of Government, Harvard Electricity Policy Group Harvard University, http://www.ksg.harvard.edu/people/whogan.

Hogan (2000) “Flowgate Rights and Wrongs”, Mimeo, JFK School of Government, Harvard Electricity Policy Group Harvard University, http://www.ksg.harvard.edu/people/whogan.

Hogan, W. (1992) “Contract Networks for Electric Power Transmission,” *Journal of Regulatory Economics*, 4, 211-242.

Joskow, P. and J. Tirole (2000) “Transmission Rights and Market Power on Electric Power Networks,” *RAND Journal of Economics*, vol. 31, no. 3, Autumn, pp. 450-487.

Joskow, P. and J. Tirole (2005) "Merchant Transmission Investment," *The Journal of Industrial Economics,* volume 53, issue 2, Page 233, June.

Kristiansen, T., Rosellón, J., 2006. A Merchant Mechanism for Electricity Transmission Expansion. *Journal of Regulatory Economics,* 29, no. 2 (March), pp. 167-193

Laffont, J.J., and J. Tirole (1996) "Creating Competition through Interconnections: Theory and Practice," *Journal of Regulatory Economics*, 10: 227-256.

Léautier, T.-O. (2000) "Regulation of an Electric Power Transmission Company," *The Energy Journal*, vol. 21, no. 4, pp. 61-92.

Littlechild, S. (2003) "Transmission Regulation, Merchant Investment, and the Experience of SNI and Murraylink in the Australian National Electricity Market", mimeo, (http://www.ksg.harvard.edu/hepg/Papers/Littlechild.Transmission.Regulation.Australia.pdf) .

Littlechild, S. and C. Skerk. (2004) "Regulation of Transmission Expansion in Argentina: Part I – State Ownership, Reform and the Fourth Line." CMI Working Paper 61. 15 November 2004. (http://www.ksg.harvard.edu/hepg/Papers/Littlechild.Regulation.Transmission.Expansion.Argentina.PartI.pdf)

Pérez-Arriaga, J. I., F. J. Rubio and J. F. Puerta Gutiérrez et al. (1995) "Marginal Pricing of Transmission Services: An Analysis of Cost Recovery," *IEEE Transactions on Power Systems*, vol. 10, no. 1, February.

O'Neill, R., E.S. Barthlomew, B.F. Hobbs, R. Baldick (2005), "A Complete Real-time Electricity Market Design," IEEE (mimeo), December 2005.

O´Neill R. P., U. Helman, B. F. Hobbs, W. R. Stewart Jr., and M. H. Rothkopf (2002) " A Joint Energy and Transmission Rights Auction: Proposal and Properties," Mimeo.

Ordover, J. and J. Panzar (1982), "On the Nonlinear Pricing of Inputs," *International Economic Review* 23, pp. 659-675.

Pope, S. (2002) "TCC Awards for Transmission Expansions," Mimeo.

Ramírez, J. C. and J. Rosellón (2002) "Pricing Natural Gas Distribution in Mexico," *Energy Economics*, vol. 24, no. 3, pp. 231-248.

Rosellón, J. (2007), "An Incentive Mechanism for Electricity Transmission Expansion in Mexico," *Energy Policy*, Article in Press, (available on line 12 December 2006).

Rubio-Odériz, J., and I. J. Pérez-Arriaga (2000) "Marginal Pricing of Transmission Services: A Comparative Analysis of Network Cost Allocation Methods," *IEEE Transactions on Power Systems*, vol. 15, no. 1, February.

Sappington, David and David Sibley (1988), "Regulating Without Cost Information: The Incremental Surplus Subsidy Scheme," *International Economic Review* 29, pp. 297-306.

Sheffrin, Anjali, and F. A. Wolak (2001) "Methodology for Analyzing Transmission Upgrades: Two Alternative Proposals," Mimeo.

Stoft, S, (2006), "Transmission Investment in a Deregulated Power Market," Chapter 4 in in F. Lévêque, Competitive Electricity Markets and Sustainability, Edward Elgar Publishing Limited.

Stoft, S. and F. Graves, (2000) "PBR Designs for Transcos: Towards a Comparative Framework, The Electricity Journal, Vol.13, No. 7, August 2000.

Transpower (2002). Confirmed Pricing Methodology. Final Design Principles. Mimeo.

Vogelsang, I. (2006). "Electricity Transmission Pricing and Performance-based Regulation," *Energy Journal,* 27(4): 97(30).

Vogelsang, I. (2001), "Price Regulation for Independent Transmission Companies," *Journal of Regulatory Economics*, vol. 20, no. 2, September.

Wilson, R. (2002), "Architecture of Power Markets," *Econometrica*, Vol. 70, No. 4, pp. 1299-1340.

Wolak, F. A. (2000) “An Empirical Model of the Impact of Hedge Contract on Bidding Behavior in a Competitive Electricity Market,” *International Economic Journal*, Summer, 1-40.