

# LONG TERM LOCATIONAL PRICES AND INVESTMENT INCENTIVES IN THE TRANSMISSION OF ELECTRICITY

Yves SMEERS\*

4 April 2005

## **Abstract**

We present a model of generation and network investment in a competitive electricity system. The model focuses on the duality of long- and short-term locational signals introduced in the European Regulation 1228/2003 for enhancing cross border trade of electricity among Member States. The model assumes that the market consists of spot and transmission submarkets. Generators, consumers and a TSO operate on that market; none of these agents has market power. Lumpiness of investments is one of the problems that may render generation and network adequacy difficult to achieve. We take up this question and apply some formalism formerly developed by O'Neill and co-authors for the unit commitment problem in order to construct multipart tariffs that insure the adequate development of the resources both in generation and the grid. In the process, we recover the standard nodal pricing as part of that multipart tariff. We address the questions of cost reflectiveness and non discrimination imposed by Regulation 1228/2003.

---

\*Tractebel Professor of Energy Economics, Department of Mathematical Engineering and Center for Operations Research and Econometrics, Université catholique de Louvain, Louvain-la-Neuve, Belgium.

Support from the French "Commission de Régulation de l'Energie (CRE)" is gratefully acknowledged.

# 1 Introduction

The economic principles that rule capacity expansion under constant return to scale are well known (e.g. Crew et al. (1995)). One invests so as to equalize short run and long run marginal costs. If these convenient assumptions prevailed in the transmission of electricity, one would invest so as to equalize the marginal transmission capacity cost with marginal congestion cost. This is not possible. The transmission of electricity suffers from many undesirable economic properties that make the direct application of these principles impossible. It combines both economies of scale and lumpy investments which make the definition of long run marginal cost illusory. It is also plagued by considerable externalities. Because the external benefits of transmission investments are not appropriable by those who are at the origin of these investments (Bushnell and Stoft (1996)), there is no hope to fully internalize these externalities.

Our objective in this paper is to present a model of network investments in a competitive electricity system based on the idea of long-term locational signals introduced in the European Regulation 1228/2003 on Cross Border Trade of Electricity, hereafter referred to as “the Regulation” (European Parliament and Council (2003)). We conduct this analysis under the following drastic simplifying assumptions.

- The architecture of the market essentially consists of a spot and transmission markets. These are coordinated by the TSO. Such architectures have been extensively explored in the literature and are now commonly implemented. Current proposals of the European Transmission System Operators (ETSO) and Power Exchanges (EuroPEX) also go in that direction (see Section 2) even if Europeans are still far from their implementation.

- We neglect the distinction between day-ahead and real time markets and only assume one spot market that integrates both. This simplification is acceptable if the day ahead and real time markets are well arbitrated, a currently heroic assumption in the European context. We justify this assumption by our objective to concentrate on the mix of long-term and short-term locational signals covered by the Regulation. This requires to somewhat disregard the detail of short-term operations.
- We assume no market power and disregard contracts. The reader will thus find no reference to physical and financial contracts, or bilateral and centralized markets. This simplification is important because these contracts influence the incentive of agents in imperfect competitive markets. This influence disappears if we assume that there is no market power as we do in this paper. The contract issue should be taken up as soon as one departs from the assumption of price taking agents.

Except for taking network indivisibilities explicitly on board, we essentially make all the assumptions of the perfect competition model that we particularize, as described above, to accommodate some idiosyncracies of the electricity system. We then concentrate on network indivisibilities by adapting and slightly generalizing a paper of O’Neill et al. (2004) dealing with indivisibilities in the unit commitment problem.

The paper is organized as follows. Section 2 introduces the notion of locational prices in the context of the Regulation. Section 3 casts the paper in the economic literature on electrical networks. Section 4 presents a reinterpretation of O’Neill et al.’s (2004) in term of network pricing. Section 5 discusses the interpretation of these results in terms appearing in the Regulation, namely “cost reflectiveness” and “non discrimination”. Section 5 also presents a transposition to the network problem of comments made by Hogan and Ring (2003) on O’Neill et al. (2004). Section 6 extends the model to

address cost causality. This requires an extension of O'Neill's Theorem 2: it is presented in Appendix A, using the theory of conic duality. Section 7 takes up the question of non discriminatory prices. It indicates that contrarily to the common wisdom underlying the European debates preceding and following the introduction of the Regulation, non discriminatory prices may entail welfare losses in a decentralized market with indivisibilities. Section 8 briefly discusses institutional matters. Last, Section 9 lists remaining questions that should be addressed in future work. The conclusions terminate the paper. They can be summarized as follows: the notion of locational price is fraught with dangers that are largely unexplored. The introduction of this notion in the Regulation may have helped achieve a political compromise. The task remains to make it workable.

## **2 Investment, locational signals and regulation**

The concept of long-term locational prices as a signal for guiding investment and location of generation plants in restructured electricity markets appears in the Regulation on “Conditions for Access to the Network for Cross-border exchanges in Electricity” that came into force on July 2004. This law is essentially an outgrowth of the work of the Florence Regulatory Forum (see DG TREN web site) installed by the European Commission to find means to facilitate electricity exchanges between Member States in the so-called Internal Electricity Market. The Regulation contains five parts

- (i) Articles 1 and 2 set the scope of the law and introduce important definitions.
- (ii) Articles 3 and 4 discuss the conditions for accessing the network. Article 3 presents a system of cross border compensations between Transmission System Operators. Article 4 introduces the notion of locational prices which is the main focus of this paper.

- (iii) Articles 5 and 6 discuss congestion management at the interconnections. Article 5 makes general informational recommendations. Article 6 requires that congestion management be market based.
- (iv) Article 7 discusses the possible exemption of new interconnectors from the general rules.
- (v) Last, Articles 8 to 15 elaborate on procedural but crucial matters and in particular the role of Comitology (Article 13).

This paper concentrates on long-term locational prices that it considers together with congestion management. The Regulation requires that locational prices be efficient, cost reflective and non discriminatory (Article 4, paragraph 1). At the time of the adoption of the Regulation (June 2003), none of the documents elaborated by the Florence Regulatory Forum had proposed any method for computing locational prices that would satisfy these three criteria. The Forum and related studies (e.g. Pérez Arriaga et al. (2002)) certainly looked at long-term locational signals. But they did so in purely accounting terms, that is by allocating network costs among network users. Cost allocation methods only reflect costs in a very weak sense that does not imply any causality: agents are not charged for the cost that they cause; they only collectively pay for the total costs incurred. Cost allocation methods are also discriminatory in an economic sense because they do not rely on marginal costs. The fact that the marginal cost of the electricity network is difficult to define, makes the combination of these two criteria difficult if not impossible to implement in a non ambiguous way. Last, cost allocation methods are not economically efficient in the sense of inducing investments in the right location (see Currien (2003), Green (2003) and Lévêque (2003)). The lack of economic efficiency is particularly dangerous in the restructuring context: it distorts the incentive to invest and hence endangers the security of electricity supply. Lack of efficient locational signals and economic incentives to invest did not matter

much in the old regulatory days. The regulatory obligation to satisfy demand implied some so-called optimal mix of capacities and locational decisions (taking constraints such as site adequacy and environment into account). This is no longer true in a competitive system where long-term and short-term locational prices are the sole market instruments that can induce both the right mix of equipments and their proper location. It is thus of the essence that the prices be right and produce the good incentive to invest. Unfortunately, economic theory tells us that allocated costs offer no guarantee in that respect.

The absence of any precise reference to efficient long-term locational signals in the work of the Florence Regulatory Forum should not come as a surprise. We do not know at this stage how to construct efficient long-term locational signals, let alone efficient, cost reflective (in the strong sense of cost caused) and non discriminatory long-term price signals. The reason is simple: a generator locates on some site or it does not. Location and the choice of a plant type are discrete decisions, and we know very little on how to induce the right discrete decisions through prices. It is indeed a basic principle of economic theory that discrete decisions are difficult if not impossible to drive through price mechanism because of non convexity phenomena (see Scarf (1994)) for a clear statement of the problem and Pérez Arriaga and Smeers (2003) for a verbal discussion of the question in electricity networks). The problem becomes particularly acute if one notes that the impact of these decisions generally covers a time period of several years.

While long-term locational signals are a rather unexplored area, short-term locational signals for dealing with congestion are a well known but highly controversial subject in Europe. These are treated in Articles 5 and 6 of the Regulation. Congestion management involves continuous decision variables that one understands much better how to decentralize through a market process. The consequence is that different market based congestion methods exist

and others are proposed. It is certainly surprising that the Regulation recommends introducing long-term locational prices that one does not know how to construct and implement but avoids suggesting well understood short-term locational prices that have proved effective in several systems.

### 3 Related literature

Different approaches to grid investments exist in the literature. They are discussed in Rosellon (2003)'s survey. We take stock of that paper to put our work in perspective and refer the reader to that survey for further information and a guide to the literature. A global and in depth analysis of investment in transmission can be found in Woolf (2003).

Hogan (2002 and 2004) and several co-authors (Pope and Harvey (2002)) extended the theory of nodal prices and Financial Transmission Rights (FTR) to the problem of investments in the grid. Their objective is to provide a market driven mechanism for expanding the grid. This analysis underlies the notion of merchant plants found in FERC Standard Market Design (SMD) proposal, which eventually also found its way in the Regulation (Article 7). We discuss neither merchant plants nor their compatibility with the rest of the Regulation but instead focus on how our model relates to Hogan et al.'s theory. The basic idea of these authors is that it is possible to decentralise (at least some) investments in the grid to economic agents (Gencos, consumers, investors ...) provided these receive long-term FTRs that guarantee the payment of congestion rents over the life of the project. The System Operator grants the long-term rights in an auction. Because investments in new lines can destroy existing transmission rights, some restrictions on the allocation of the long-term rights are necessary to keep the set of granted rights physically feasible for the network. This latter process is rather complex (Hogan (2002)) but the bottom line is that congestion rents remunerate the investors who

invest in the network. Joskow and Tirole (2003) argue that the theory suffers from several shortcomings and Hogan replied to some of these comments (Hogan (2004)). We do not elaborate on these discussions that the reader can find in summarised form in Roselon's paper but concentrate on one of the points unveiled in Joskow and Tirole (2003) and clearly recognized in Hogan (2002 and 2004), namely investment lumpiness.

FTR were primarily introduced as hedges against random congestion charges (Hogan (1992)). FTR are forward contracts on spot transmission prices. They are special forward contracts though, in the sense that they need to be traded through the Transmission System Operator in order to guarantee that they are compatible with the physical possibilities of the network. FTR satisfy the standard property derived from finance theory that the forward price is the expectation, in some risk neutral probability, of the congestion price. In other terms, the economic signal embedded in long-term FTRs only contains congestion charges. In order to see this more clearly consider a deterministic world. There is no need to hedge and a FTR is then exactly equivalent to a payment of the congestion charge in real time. Resorting to long-term FTRs for inducing investments in the grid is then equivalent to using congestion charges as incentives to invest in the grid. Investment lumpiness limits this potential incentive. We elaborate in the following on the consequences of this point and argue that, notwithstanding investment lumpiness, capacity expansion of the grid can still be decentralised provided one introduces more prices than the sole congestion charges. We come back in Section 8 on the interpretation of that result in terms of Hogan's theory.

The Transco model is a second approach to investments in the grid. A Transco is a company that builds and operates the network for profit. In the terms of this paper and taking the concepts of the Regulation on board, the Transco is remunerated with both long-term and short-term locational charges.



It then develops the grid by solving a network capacity expansion problem with the view of maximising its profit. The network is a natural monopoly that gives market power to the Transco. If not regulated, the company will set long and short-term charges in a way that maximises its profit but degrades welfare. The question is thus to find charges that provide the incentive to be efficient and permit the firm to recover its costs. This has different aspects.

Maybe more than any other regulatory question, investments in the grid have a strong flavour of asymmetry of information. The electric network is indeed a complex technology that maybe difficult to grasp from outside the industry. The problem of incentivising the Transco will then require the regulator to offer a menu of contracts. Models of this type usually assume an explicit demand function for the services of the regulated company (e.g., Vogelsang (2001)). For the sake of analytical convenience, these analyses also resort to simplifications that neglect the multi-product nature of transmission services. We depart from these considerations and see the Transco as a profit maximising firm among others. We also consider the demand for transmission services not as explicitly given by a demand function but as derived from the interactions between profit maximizing generators and consumers. Last, we also assume an electricity network that offers different point-to-point services. Our objective, in following that path, is not to overlook the important problem of the market power of the Transco, but to attempt to construct a partial equilibrium model with price taking agents that we intend to expand in future developments.

Our objective can be stated as follows: consider a world of perfect information with price taking agents (including the Transco). The question is to find the price signals that induce the Transco to invest in an efficient network and allows one to recover the cost of building this network. The need for adequate price signals for a Transco derives from the now well recognized fact that

a (price taking) transmission company that only receives congestion charges and an additional fixed revenue has an incentive to under-invest in order to increase congestion. It is also admitted that the sole recovery of congestion charges only allows the Transco to recover a relatively small fraction of the total costs of the grid (Pérez-Arriaga et al. (1995)). One can thus expect to see multi part tariffs emerge as necessary instruments for achieving both the efficiency and cost recovery objectives. Non linear tariffs are commonly encountered in the regulatory literature for dealing with asymmetry of information. They also appear in non convex economies (see Bjørndal and Jörsten (2004) for a recent treatment through dual price functions). This latter context is the one adopted in this paper.

In short, and possibly in contrast with the existent literature on Transco, we disregard the question of market power, but concentrate on the existence of price signals that induce a price taking Transco to manage congestion and develop the network in an efficient way even though investments are lumpy. We do this by constructing a model that fully accounts for the multiproduct nature of the transmission infrastructure and develop an equilibrium framework where the demand for transmission services results from other agents actions. This was the approach formerly adopted in studies of operations and market design. We hope to pave the way to similar developments in two directions; on the engineering side, we introduce the capability to accommodate realistic grid representations; on the economic side, we derive price taking equilibrium conditions susceptible of being extended to accommodate manifestations of market power.

## **4 A model of locational and congestion pricing**

The mix of discrete decisions (that one does not know how to decentralize through prices) and continuous decisions (that one knows very well how to

decentralize through prices) found in the electricity grid is not unique. It also appears in an other area of electricity restructuring namely unit commitment and plant dispatching. The mentioned papers of O’Neill et al. (2004) and Hogan and Ring (2003) address the question of finding price signals for inducing plants to start up and shut down. We transpose and slightly extend their reasoning to the questions of locational and congestion pricing.

#### 4.1 The grid equilibrium model

The following Primal Mixed-Integer Program (PIP) was considered by O’Neill et al. (2004) for studying price based decentralization of start up decisions in unit commitment problems

$$\max \nu_{\text{PIP}} = \sum_k c_k x_k + \sum_k d_k z_k \quad (1)$$

$$\text{s.t.} \quad \sum_k A_{k1} x_k + \sum_k A_{k2} z_k \leq b_0 \quad (2)$$

$$B_{k1} x_k + B_{k2} z_k \leq b_k \quad \forall k \quad (3)$$

$$x_k \geq 0, \quad z_k \in \{0, 1\}^{n(k)}. \quad (4)$$

This model is a standard mixed integer program (see any textbook on integer programming, e.g. Wolsey (1998)). It is easily interpreted in the context of optimal centralized investments in a regulated electricity system. We accordingly assume a system where investment decisions in the grid and in power generation are under the supervision of a benevolent, perfectly informed and perfectly rational (with unlimited computational possibilities) Regulator. Let  $j = 1, \dots, J$  denote the nodes of the electrical network. We cast O’Neill et al. model in the locational pricing context through the following interpretation.

- (i) Let  $k = 1, \dots, K$  designate an agent, consumer or generator, that decides to connect to the network.
- (ii)  $z_k$  is the vector of binary variables representing locational decisions of agent  $k$ . A component of  $z_k$ , let  $z_{kj}$ , is equal to 1 if  $k$  connects to the

network at location  $j$ ,  $j = 1, \dots, J$ . It is zero otherwise. All components of  $z_k$  are 0 if agent  $k$  does not connect to the network. The decision to connect to the network is a simplified view of a more complex decision: an agent, whether a consumer or generator, connecting to the network indeed builds a plant of a certain type, a feature that is not represented yet in the variable  $z$ . From a welfare point of view, the decision to connect therefore also implies the decision to build and hence a cost. We limit ourselves in this section to the sole locational decision. We shall later introduce different technologies,  $i = 1, \dots, I$ , and extend the definition of  $z_k$  by considering  $z_{kij} = 1$  when agent  $k$  connects a plant of technology  $i$  to location  $j$  of the network.

- (iii)  $x_k$  is the vector of power injection/withdrawal, of generator/consumer  $k$ . There are potentially as many components of  $x_k$  as there are nodes in the network. We replace  $x_k$  by  $\text{sign}_k x_k$  in O'Neill et al.'s expressions (1) and (2) in order to distinguish between injections and withdrawals. In this updated expression, all  $x_k$  are assumed positive and  $\text{sign}_k$  is a diagonal matrix whose components are  $-1$  for injections and  $+1$  for withdrawals.  $\text{sign}_k x_k$  is thus the vector of the net nodal withdrawals of agent  $k$ .
- (iv)  $d_k$  is the vector of fixed costs before any payment for locational charges incurred by agent  $k$  when it connects to the network.  $d_{kj}$  is thus the fixed cost of building a plant of (currently) unspecified technology in location  $j$ . Alternatively  $d_{kj}$  can be interpreted as a vector of fixed benefits or costs accruing to a consumer  $k$  when it connects to the network in location  $j$ .
- (v)  $c_k$  is the vector of the marginal costs of the generators; alternatively,  $c_k$  is the marginal utility of the consumer  $k$ . Both marginal costs and utilities are taken constant as in O'Neill et al.'s model. A slight complication of the notation makes it possible to approximate nonlinear concave utility and convex costs.

(vi) Given the interpretation of the  $x$  variables, relation (2) represents the load flow based constraints that limit the injections and withdrawals in the network.  $A_{k1}$  is thus identical for all generators and consumers  $k$  and can be noted  $A$ . The matrix  $A$  consists of two parts. One row of the matrix expresses the balance between injections and withdrawals in a lossless network. This is written

$$\underline{1} \cdot \sum_k \text{sign}_k x_k = 0 \quad (5)$$

where  $\underline{1}$  is a row vector of 1. The other rows of  $A$  and the associated constraint express the limitations on the injections and withdrawals resulting from thermal limits on the lines. These are expressed using the negatives of the Power Distribution Factors (PTDF) that give the flow in each line of the grid as a result of the injections and withdrawals  $\text{sign}_k x_k$ . We keep the inequality formulation of O'Neill et al., even though relation (5) is an equality. This facilitates the notation and can easily be justified.

(vii)  $A_{k2}$  is taken as 0 in this application and will no longer be used in this section.

(viii) The first component of  $b_0$  is the righthand side of (5) and is thus zero. The other components of  $b_0$  are the thermal limits on the lines of the network (that one can only writes in one direction in order to simplify notation). One thus sees that (2) represents the usual DC load flow approximation of network flow.

(ix)  $B_{k1}$  is an identity matrix and  $B_{k2}$  a diagonal matrix of the capacity (noted  $m$ ) installed by generator/consumer  $k$ ;  $b_k$  is identically 0.

Using these notations, PIP can be restated as program PIP<sub>LOC</sub>

$$\max \nu_{\text{PIP}_{\text{LOC}}} = \sum_k c_k \text{sign}_k x_k + \sum_k d_k \text{sign}_k z_k \quad (6)$$

$$\text{s.t.} \quad A \sum_k \text{sign}_k x_k \leq b_0 \quad (7)$$

$$x_k - m_k z_k \leq 0 \quad (8)$$

$$x_k \geq 0, \quad z_k \in \{0, 1\}^{n(k)}. \quad (9)$$

## 4.2 Extensions and restrictions

This program involves both investment ( $z$ ) and operation ( $x$ ) variables. One immediately notices that  $\text{PIP}_{\text{LOC}}$  boils down to the usual optimal dispatch/welfare optimization problem extensively studied in the congestion management literature when the  $z_k$  variables are fixed. This reduced problem is the basis of the study of short-term locational signals.  $\text{PIP}_{\text{LOC}}$  therefore imbeds the question of congestion management covered in Articles 5 and 6 of the Regulation. The aim of this paper is to free the  $z_k$  variables so as to also cover the long-term signals of Article 4.

The model can easily be extended to accommodate several time segments and contingencies. Let  $x_k^\tau$  be the power generation/withdrawal of generator/consumer  $k$  in time segment or contingency  $\tau$ ; the model can be rewritten as program  $\text{PIP}_{\text{LOC}}^{mp}$  (where  $mp$  stands for multiperiod)

$$\begin{aligned} \max \nu_{\text{PIP}_{\text{LOC}}^{mp}} &= \sum_k \sum_\tau c_k^\tau \text{sign}_k x_k^\tau + \sum_k d_k \text{sign}_k z_k \\ \text{s.t.} \quad A^\tau \sum_k \text{sign}_k x_k^\tau &\leq b_0^\tau && \forall \tau \\ x_k^\tau - m_k^\tau z_k &\leq 0 && \forall \tau \\ x_k^\tau \geq 0, \quad z_k &\in \{0, 1\}^{n(k)}. \end{aligned}$$

As mentioned above, the model can also be easily extended to allow  $z_k$  to represent the decision to locate at a certain node and to build a certain capacity. This extension only makes sense in a multiperiod (time segment) model. This model would thus consider variables  $z_{kij}^\tau$  and  $x_{kij}^\tau$  where  $i$  is the type of plant.

In order to simplify the discussion and to facilitate the establishment of the correspondence between this paper and O'Neill et al., we first limit ourselves to the simple formulation (6) to (9).

Program  $\text{PIP}_{\text{LOC}}$  should be interpreted as the problem to be solved by a regulator operating under the following ideal conditions:

- (i) Perfect information: the Regulator knows the marginal willingness to pay of the consumers and the marginal costs of the generators. The Regulator also knows the locational fixed costs (the fixed cost of building a plant (of undefined technology at this stage) at some location).
- (ii) Perfect competition: agents are price takers
- (iii) Perfect congestion management: congestion is managed by nodal prices.
- (iv) Simplified electrical assumption: the network is lossless and its structure is described by a PTDF matrix.

Assumptions (ii), (iii) and (iv) are common in studies of restructured electricity systems. Assumption (i) may look harsh, but is not more stringent than the implicit assumption that prevailed in the former regulatory regime where the Regulator was also assumed to decide or accept investments according to a perfect optimization cost minimization model. A main difference is that the Regulator of the former monopoly regime did not have to interpret the results of this optimization model in terms of locational prices. He could be wrong with less damaging consequences. It is on these prices that we concentrate, assuming all the above simplifying assumptions to facilitate their discovery.

### 4.3 Locational charges

O'Neill et al. suppose that one can solve Problem  $\text{PIP}_{\text{LOC}}$  and we follow suit. The assumption is indeed not that unrealistic given the astonishing numerical progress made in mathematical programming these last 20 years.

Let  $\bar{z}_k \forall k$  be the optimal location of the generators/consumers, we state a Primal Linear Program PLIP( $\bar{z}$ )

$$\begin{aligned} \max \sum_k c_k \text{sign}_k x_k + \sum_k d_k \text{sign}_k z_k \\ A \sum_k \text{sign}_k x_k \leq b_0 \quad (u_0) \end{aligned} \quad (10)$$

$$x_k - m_k z_k \leq 0 \quad (u_k) \quad \forall k \quad (11)$$

$$z_k = \bar{z}_k \quad (w_k) \quad \forall k \quad (12)$$

$$x_k \geq 0$$

where  $u_0, u_k$  and  $w_k$  are the dual variables of the constraints (10), (11) and (12) respectively.

The dual DLIP<sub>LOC</sub>( $\bar{z}$ ) of that problem is written

$$\min u_0 b_0 + \sum_k w_k \bar{z}_k \quad (13)$$

$$u_0 A \text{sign}_k + u_k \geq c_k \text{sign}_k \quad \forall k \quad (14)$$

$$-u_k m_k + w_k \geq d_k \text{sign}_k \quad \forall k \quad (15)$$

$$u_0 \geq 0, \quad u_k \geq 0, \quad w_k \text{ unconstrained} \quad \forall k. \quad (16)$$

Note from the above discussion that  $\text{sign}_k$  in relation (14) multiplies the nodal components of  $u_0 A$  by  $-1$  or  $+1$  depending on whether agent  $k$  is a generator or a consumer at that node. The dual variables of the problem can then be interpreted as follows.

- (a) The first component of  $u_0$  is the price of electricity at the hub node. The other components of  $u_0$  are the values of the capacities of the lines (flowgates). These components of  $u_0$  are indeed the dual variables of the thermal capacities of the lines and should be interpreted in the same sense as in usual discussions of congestion management. One can easily verify that this implies that  $u_0 A$  is the vector of the nodal prices of electricity.



(b)  $u_k$  is the vector of scarcity premium on tight generation and consumption capacities. The constraints (14) express that the nodal price at some active generation node (the corresponding nodal component of  $u_0 A$ ) is equal to the variable cost (the corresponding nodal component of  $c_k$ ) of a generator at that node which operates at a positive level, but below capacity. This price is equal to the variable cost plus a scarcity premium if the generator operates at full capacity at that node. (Note that  $\text{sign}_k = -1$  implies  $(u_0 A)_k - u_k \leq c_k$ ). A similar interpretation holds for consumers.

(c)  $w_k$  are locational prices in the sense that a generator/consumer  $k$  pays  $w_{kj}$  to locate at  $j$ . Note that  $d_k$  are costs for generators. A constraint (15) expresses that the locational price is greater or equal to the capacity margin minus the fixed cost of the plant. The equality holds for a plant that is effectively located at that node. In that case the locational price and the capacity margin pay for the cost of locating the plant at that node. Note that  $w_k$  is unconstrained: its components can be positive or negative. If negative a component of  $w_k$  should be interpreted as a payment to be given to the agent to locate.

#### 4.4 Efficiency properties of locational charges

We now transpose O'Neill et al.'s result in the interpretation of locational and congestion pricing.

**Definition 1** *Efficient short-term and long-term locational prices are prices  $\bar{u}^0$  (congestion) and  $\bar{w}_k$  (locations) such that agents (generators and consumers) when charged  $(\bar{u}^0, \bar{w})$  behave efficiently in the following sense. Suppose that these agents maximize their profits under the sole capacity constraints*

$$x_k - m_k z_k \leq 0.$$

In other words, agent  $k$  solves the following profit maximization problem

$$\max(c_k - \bar{u}^0 A) \text{sign}_k x_k + (d_k \text{sign}_k - \bar{w}_k) z_k \quad (17)$$

$$\text{s.t.} \quad x_k - m_k z_k \leq 0 \quad (18)$$

$$x_k \geq 0, \quad z_k \in \{0, 1\}^{n(k)} \quad (19)$$

Then the solutions  $\bar{x}'_k, \bar{z}'_k$  that they obtain are those desired by the Regulator who solves  $PIP_{LOC}$ . Moreover the injections and withdrawals  $\bar{x}'$  balance and comply with the network constraints

$$A \sum_k \text{sign}_k \bar{x}'_k \leq b_0. \quad (20)$$

Note that this definition only refers to economic efficiency. It corresponds to the notion of market clearing price in O'Neill et al. but does not embed the notions of cost reflectiveness or non discrimination also foreseen by Regulation. We examine these notions in subsequent sections of this paper.

The transposition of O'Neill et al.'s results to our context can then be expressed in the following proposition.

**Proposition 1** *Suppose that the Regulator successively solves problems  $PIP_{LOC}$  and  $PLIP_{LOC}(\bar{z})$  where  $\bar{z}$  is the solution to  $PIP_{LOC}$ . Let  $\bar{u}^0$  and  $\bar{w}_k$  be part of the optimal dual solution of  $PLIP_{LOC}(\bar{z})$ . Then  $\bar{u}^0$  and  $\bar{w}_k$  are respectively efficient short run and long run locational prices. Let  $\bar{x}'_k$  and  $\bar{z}'_k$  be the decision of agent  $k$ . One has  $(c_k - \bar{u}^0 A) \text{sign}_k \bar{x}'_k + (d_k \text{sign}_k - \bar{w}_k) \bar{z}'_k = 0$ .*

**Proof.** The proposition is a direct application of the Theorem 2 in O'Neill et al., which is recalled in generalized form in Appendix A.

In this system, the Regulator or the System Operator receives the congestion charges  $\bar{u}^0 A \sum_k \text{sign}_k \bar{x}'_k$  and the location charges  $\sum_k \bar{w}_k \bar{z}'_k$ . Because the statement of the problem does not contain any information on the cost of the network, Proposition 1 contains no result on the balance of the TSO's budget.

Similarly, there is no information on how the cost of the network depends on the locational decision of the agents; hence the result cannot express any cost causality.

It is remarkable that generators that decide to locate and produce make a zero profit. Similarly, consumers that decide to locate and consume have a zero net welfare gain. This is the standard result of perfectly competitive equilibrium market which is recovered here by the application of Theorem 2 of O'Neill's et al. All other agents would incur a loss if they were to locate on the network and generate/consume electricity.

## 5 Discussion

The above congestion and locational prices satisfy one objective of the Regulation, namely economic efficiency. They leave the decision to locate, generate and consume to the agents who pay these prices. The prices  $\bar{w}_k$  and  $\bar{u}^0$  are also quite distinct. The payment  $\bar{w}_k$  results from the decision to locate on the network and can be interpreted as the long-term signals of Article 4 of the Regulation. These payments must be made irrespective of the consumption or generation levels provided the decision is made to connect to the network. The nodal prices  $\bar{u}^0 A$  entail congestion charges that are directly proportional to consumption and generation. They comply with the obligation to use market based congestion management methods of Article 6 of the Regulation. Besides offering the two price signals required by the Regulation,  $\bar{u}^0 A$  and  $\bar{w}$  also form true two-part tariffs in the sense commonly understood by economists.

As in O'Neill et al., the approach extends the commodity space beyond the sole energy space. Both energy and the locational rights are priced. One should note that it is of the essence not to convert the fixed part  $\bar{w}_k$  of the tariff into a proportional charge that would come on top of  $\bar{u}^0 A$ . Doing this would indeed convert the two-part tariff into two single part linear prices,

which cannot sustain the equilibrium in the presence of indivisibilities. This important remark does appear neither in the Regulation nor in the allocation mechanisms discussed by the TSOs and Regulators in the Florence Forum. While Proposition 1 of Section 4 states that these nonlinear tariffs satisfy the efficiency objective of Article 4 of the Regulation, none of the other desired properties (cost reflectiveness and non discrimination) required by the Regulation is achieved. This is not specific to this construction as two-part tariffs are often discriminatory in two senses. By construction they lead to average prices that are decreasing with quantity. This second degree (e.g. Tirole (1998)) price discrimination is allowed by Courts. Two-part tariffs are also first degree (Tirole (1998)) discriminatory prices. They differ by agent in the sense that, except for their proportional part which is equal to the marginal cost, the fixed part does not reflect the cost incurred by the generator but the willingness to pay of the agent. This is unlikely to be accepted by competition authorities without further justification. We therefore elaborate on this discrimination.

- (a) The  $\bar{w}_k$  are first degree discriminatory in the sense that two generators that locate at the same site may not necessarily pay the same locational price. The only possible justification of that discrimination is that, in contrast with the common wisdom of the Florence process, price discrimination is sometimes necessary in order to achieve economic efficiency. This is what happens here: the signals are much more than simply locational; they are individual signals, targeted at each candidate in each location.
- (b) The  $\bar{w}_k$  are not cost reflective. They bear no relation to the cost of the network which is completely absent from the statement of the problem.

These prices also lack other important desirable properties. First there is no guarantee that the sum of the congestion payments and the long run

locational price recovers the cost of the network, let alone induces an optimal development of the network. Second the locational prices do not reflect the incremental cost that the TSO would incur because of the decision of an agent to locate at some node of the network. This is not surprising as the model  $\text{PIP}_{\text{LOC}}$  does not contain any representation of the cost of the network.

In short efficiency is achieved at the cost of sacrificing non discrimination, cost reflectiveness and other desirable properties. One may thus search for a more involved tariff structure that exhibits more components, some of them related to the cost of the network. Before turning to this point, it is useful to review Hogan and Ring (2003) comments on O'Neill et al.'s model and to see whether they can be transposed to the problem of locational and congestion prices.

Hogan and Ring (2003) note that there may be many  $w_k$  capable of supporting the dispatch equilibrium in a unit commitment problem. We want to show that the same holds for the locational equilibrium. To see this, recall first that  $\bar{u}_k$  is the scarcity rent captured by the optimally located plants when operating in a perfectly competitive system. There may be several  $\bar{u}_k$  but take one of them. Transposing Hogan and Ring's reasoning to locational pricing, we first consider the minimal payment  $w_{kj}(\bar{u})$  that should be paid to agent  $k$  in location  $j$  in order to induce a decision compatible with the  $\bar{z}$  desired by the Regulator. Using Hogan and Ring's notation and considering the case of a generator as an example, one defines

$$\bar{\Pi}_{kj} = -d_{kj} \bar{z}_{kj} + \bar{u}_{kj} m_{kj} \quad (21)$$

$$\bar{\Pi}_{kj}^+ = \max(-d_{kj} + \bar{u}_{kj} m_{kj}) z_{kj}, \quad z_{kj} \in \{0, 1\}. \quad (22)$$

$\bar{\Pi}_{kj}$  is the profit collected by agent  $k$  at location  $j$  if it behaves according to the Regulator's plan. This profit consists of a capacity scarcity margin  $\bar{u}_{kj} m_{kj}$  minus the incurred fixed cost  $d_{kj} \bar{z}_{kj}$ .  $\bar{\Pi}_{kj}^+$  is the profit that the same agent would collect if it made its own decision to invest in location  $j$  on the basis

of the expected capacity scarcity margin  $\bar{u}_{kj}$  and its investment cost  $d_{jk}$ . The profit  $\Pi_{kj}^+$  is what would be driving the generator locational decision in a pure nodal system. Define

$$w_{kj}^-(\bar{u}) = \min(0; \bar{\Pi}_{kj} - \Pi_{kj}^+). \quad (23)$$

Suppose  $w_{kj}^-(\bar{u})$  is strictly negative, then one can verify that a payment of  $-w_{kj}^-(\bar{u})$  to generator  $k$  at  $j$ ) compensates this agent for

- the opportunity cost  $\Pi_{kj}^+$  of not following its own optimal strategy;
- the cost of following the strategy of the Regulator.

Any higher payment would over compensate the agent.

Conversely, one may also be interested in computing the maximal positive charge that can be levied on the generator that decides to invest before it modifies its decision. Define

$$w_{kj}^+(\bar{u}) = \Pi_{kj}^+. \quad (24)$$

One can verify that a levy  $w_{kj} \leq w_{kj}^+$  will not induce a generator  $k$  that spontaneously decide to invest in location  $j$  to change its decision. One can thus imagine that a Regulator could try to finance (part of) its payments  $w_{kj}^-(\bar{u})$  from charges  $w_{kj}^+(\bar{u})$ . More generally, let  $KJ^+$  and  $KJ^-$  be respectively the pairs  $(k, j)$  for which  $w_{kj}^+$  and  $w_{kj}^-$  have been defined. A  $w'$  supports the locational equilibrium if

$$w'_{kj} \leq w_{kj}^- \quad (k, j) \in KJ^- \quad (25)$$

$$w'_{kj} \leq w_{kj}^+ \quad (k, j) \in KJ^+. \quad (26)$$

This implies that

$$- \sum_{(k,j) \in KJ^+} w_{kj}^+(\bar{u}) - \sum_{(k,j) \in KJ^-} w_{kj}^-(\bar{u}) \quad (27)$$

is the minimal net payment that the Regulator must be prepared to make in order to induce the different agents to abide to its locational strategy  $\bar{z}$ . Nothing guarantees that this minimal payment is negative, that is, that the regulator will break even. Any deficit should be covered otherwise.

A similar reasoning applies to the consumers with a like conclusion. The overall conclusion in terms of covering revenue requirements is more positive as soon as one considers generators and consumers together. Supposing that the total surplus of the market is positive (a very mild assumption if one starts from nothing), then it is always possible by playing on the  $w$  according to relations (25) and (26) to satisfy the revenue requirement.

The need for the Regulator to break even may lead one to increase the discrimination between agents with respect to the original  $w$ . This suggests introducing constraints on the  $w$  that are directly inspired by commonly found regulatory objectives. One may for instance wish to achieve

(i) a budgetary balance. This would imply

$$\sum_k \sum_m w'_{km} \bar{z}_{km} = 0 \quad (28)$$

or

$$\sum_k \sum_m w'_{km} \bar{z}_{km} = \text{Fixed Network Cost.} \quad (29)$$

(ii) Non discrimination constraints. These impose some equalities between the negative  $w'_{kj}$  and/or other equalities between the positive  $w'_{kj}$ .

Needless to say there is no guarantee that imposing these regulatory constraints on top of (25) and (26) will lead to feasible  $w'_k$ . However it is always possible to try to get close to regulatory objectives by minimizing the violation of the constraints. Note that one cannot hope to collect the deficit accumulated by locational payments to the generators by charges levied on customers such that  $w'_{kj} > w^+_{kj}$  as these would leave a negative surplus to

these agents. We shall not elaborate any further on these variations in this section and return to the problem of cost causality.

## 6 Cost causality and nonlinear locational prices

Section 5 concludes that the  $\bar{w}_k$  derived from problem  $\text{PLIP}_{\text{LOC}}$  are discriminatory and not cost reflective. This is not surprising. Non discriminatory prices are well known not to be able to support an efficient competitive market when there are indivisibilities and economies of scale. The absence of cost reflectiveness could also be expected as the model  $\text{PIP}_{\text{LOC}}$  used to compute the  $\bar{w}_k$  does not link agents' locational decisions to the grid structure and hence to the grid costs. Also, one shall note that the above tariff only prices energy and location, while tariffs that include a capacity charge and hence a third component in the tariff are common in practice. The following develops a richer set up with the view of introducing cost causality in the model.

### 6.1 Representation of the network

Consider an extension of problem  $\text{PIP}_{\text{LOC}}$ , constructed as follows.

- (a) There exists a set of possible network configurations  $n = 1, \dots, N$  each of cost  $d_0^n$  and PTDF matrix  $\text{PTDF}^n$ . The System Operator selects one of these configurations. Let  $z_0$  be an  $N$ -dimensional vector with  $z_0^n = 1$  if the TSO selects the  $n^{\text{th}}$  network configuration.
- (b) The locational price takes the form of a two-part tariff, namely a fixed charge for accessing the network and a proportional charge for reserving capacity (a capacity or demand charge). This reserved capacity can be a single variable valid for the whole year, or a vector if one implements seasonal reservations e.g. summer and winter (see Perrez Arriaga and Smeers (2003)). We first work with a single annual capacity reservation



in order to facilitate the discussion. With this interpretation  $x_k$  plays both the role of energy generated or consumed and reserved capacity.

- (c) One introduces some causality between the decision to locate, the reserved injection and withdrawal capacity and the network structure. This causality is complex in the real world. Strictly speaking, it derives from a network expansion planning problem. Because our methodology uses optimization type techniques to explore economic problems, we adopt an approach theoretically justified in optimization terms but practically still to be elaborated and represent causality through two types of constraints.

- (i) The “skeleton network constraints”

$$F_0 z_0 - \sum_k F_{k \neq 0} z_k \leq 0 \quad (30)$$

imply that the decisions of the agents to locate (the  $z_k, \forall k \neq 0$ ) impose a certain minimal structure on the network. This results in a set of allowed network structures  $z_0$  (there may be several  $z_0$  for a given set of location decisions  $z_k$ ).

- (ii) The “incremental network constraints”

$$- \sum_{k \neq 0} G_{1k} x_k + G_{20} z_0 \leq 0 \quad (31)$$

imply that the capacities reserved by the different agents (the  $x_k, \forall k \neq 0$ ) further restrict the set of allowable  $z_0$  of the network.

- (iii) The selection constraint

$$\sum_{n=1}^N z_0^n \leq 1 \quad (32)$$

implies that the TSO can only select a single network configuration.

Together these constraints restrict the set of possible configurations of the network compatible with the locational and capacity reservation decisions of the agents. This might be sufficient to fully determine the

configuration of the network, but not need to be so. We therefore also leave the possibility that the TSO optimally (that is, through cost minimization) selects the configuration. One could obviously argue that the constraints (31) can be merged with (30) and hence that one only needs to write one set of constraints. A reduction of the number of constraints serves no useful purpose whether computationally or economically. From a computational point of view, the formulation that explicitly retains the two types of constraints is likely to be tighter in an integer programming sense (see Wolsey (1998)) and hence more efficient. From an economic point of view, the two constraints express different cost causality and hence should be made explicit. We recognize that the construction of the constraints (30) and (31) may be difficult in practice, but leave it to further research to investigate how they can be inferred from both engineering practice and mathematical programming models. Before leaving this point, we note that the absence of any agent connecting to the grid ( $z_k = 0, x_k = 0, \forall k$ ) eliminates the need for a network and hence allows the solution  $z_0 = 0$ . This justifies setting the right hand side of these constraints to zero. This also means that these constraints define a cone when the integer restrictions are relaxed.

- (d) Constraints (30) and (31) delineate the set of acceptable network configurations. To each of them, one associates a matrix of PTDF coefficients,  $\text{PTDF}^n$ , and a vector of thermal limit constraints  $b^n$ . We define the set of acceptable injections and withdrawals in network configuration  $n$  as

$$K^n = \{x_0^n \mid \text{PTDF}^n x_0^n \leq b^n, \underline{\mathbf{1}} x_0^k = 0\} \quad (33)$$

where  $\underline{\mathbf{1}}$  is a row vector of 1 of appropriate dimension. Note that we neglect the constraint  $-b^n \leq \text{PTDF}^n x_0^n$  in order to simplify the presentation.  $K^n$  is thus the set of balanced injections and withdrawals that result in flows that do not exceed the thermal limits of the lines in

configuration  $n$ . We also introduce a cone  $C^n$  associated to the network configuration  $n$  as

$$C^n \equiv \left\{ (x_0^n, z_0^n) \mid \frac{x_0^n}{z_0^n} \in K^n, \underline{1}x_0^n = 0 \text{ if } z_0^n > 0, x_0^n = 0 \text{ if } z_0^n = 0 \right\} \quad (34)$$

(recall that  $z_0^n$ , as defined above is equal to 1 if one selects network configuration  $n$  and is zero otherwise)

Define

$$x_0 = (x_0^n, n = 1, \dots, N), \quad z_0 = (z_0^n, n = 1, \dots, N) \quad (35)$$

and the cone

$$C_0 \equiv \{(x_0, z_0) \mid (x_0^n, z_0^n) \in C_0^n, n = 1, \dots, N\}. \quad (36)$$

It is obvious that (34) defines a convex cone. We also know that the cartesian product of convex cones is a convex cone.  $(x_0, z_0) \in C_0$  is thus a convex constraint that we shall impose on the network company.

The Transmission System Operator selects a single  $z_0^n$  that minimizes its cost  $d_0^n$  and satisfies (30) – (31). Given this selection, the TSO can offer injection and withdrawal services  $x_0$  such that  $(x_0, z_0) \in C_0$ . The agents  $k$  request injection and withdrawal services  $x_k$  such that

$$\sum_k \text{sign}_k x_k = \sum_n x_0^n \quad (37)$$

$$(x_0, z_0) \in C_0 \quad (38)$$

$$\sum_{n=1}^N z_0^n = 1, z_0 \in \{0, 1\}^N. \quad (39)$$

Relation (37) equalizes the production and consumption of the network services (injection/withdrawal). It also transforms the different  $x_k$  into a single nonzero vector  $x_0$  of injection/withdrawal services for the selected network configuration. Relation (38) expresses that the TSO can only offer network services compatible with the selected configuration of the grid. This is the convex part of the production set of the transmission company. In contrast,

(39) describes its non convex part, namely that the TSO can only select one configuration of the network. In order to simplify the notation, we rewrite relation (37) as  $E_0 x_0 - \sum_{k \neq 0} E_k x_k = 0$ . In order to fully use conic duality, we also define the cone  $C_k$  for each generator or consumer  $k$ .

$$C_k = \{(x_k, z_k) \mid x_k - m_k z_k \leq 0, x_k \geq 0, z_k \geq 0\}.$$

## 6.2 The Regulator's global problem

Introducing a diagonal matrix  $\text{sign}_0$  with only  $-1$  components, the extension of the  $\text{PIP}_{\text{LOC}}$  problem can be stated as

### Problem $\text{EPIP}_{\text{LOC}}$

$$\max \sum_{k \neq 0} c_k \text{sign}_k x_k \qquad + \sum_k d_k \text{sign}_k z_k \qquad (40)$$

$$E_0 x_0 - \sum_{k \neq 0} E_k x_k \qquad = 0 \qquad (41)$$

$$F_0 z_0 - \sum_{k \neq 0} F_k z_k \leq 0 \qquad (42)$$

$$- \sum_{k \neq 0} G_{1k} x_k \qquad + G_{20} z_0 \qquad \leq 0 \qquad (43)$$

$$\underline{1} \cdot z_0 \qquad = 1 \qquad (44)$$

$$(x_k, z_k) \in C_k, \forall k \qquad (45)$$

$$z_k \in \{0, 1\}^{n(k)}. \qquad (46)$$

In this model, (41) expresses the equality between produced and consumed network services; (42) and (43) relate the selection of the network configuration and the locational decisions of the agents; (44) states that only a single network configuration is allowed; (45) describes the production sets of all agents, TSO, generators and consumers. Specifically  $C_0$  imposes the flows to be feasible for the selected network configuration and  $C_k$  requires that an agent first needs to locate and reserve some injection or withdrawal capacity before generating or consuming.

Note that in contrast with problem  $\text{PIP}_{\text{LOC}}$  which is a linear mixed integer program, the conic constraint (45) makes problem  $\text{EPIP}_{\text{LOC}}$  a nonlinear mixed integer program. This non linearity should not be a real concern though as  $\text{EPIP}_{\text{LOC}}$  can easily be restated as a linear mixed integer program of the type precedingly developped by power engineers for network capacity expansion (see Latorre et al. (2003) for a survey of these models). The nonlinear version is used here because of its tractability for extending O'Neill et al.'s Theorem 2. We therefore again assume that this problem can be solved (for instance by solving the linear mixed integer version of it) and note  $(\bar{x}_0, \bar{z}_0)$  and  $(\bar{x}_k, \bar{z}_k), \forall k$  an optimal solution.

### 6.3 The continous version of the Regulator's problem

Let thus  $\bar{z}$  be the locational vector selected by the Regulator after solving  $\text{EPIP}_{\text{LOC}}$ . Define

**Problem  $\text{EPLIP}_{\text{LOC}}(\bar{z})$**

$$\max \sum_{k \neq 0} c_k \text{sign}_k x_k + \sum_k d_k \text{sign}_k z_k \quad (47)$$

$$\text{s.t.} \quad E_0 x_0 - \sum_{k \neq 0} E_k x_k = 0 \quad (u_0) \quad (48)$$

$$F_0 z_0 - \sum_{k \neq 0} F_k z_k \leq 0 \quad (\lambda_f) \quad (49)$$

$$- \sum_{k \neq 0} G_{1k} x_k + G_{20} z_0 \leq 0 \quad (\lambda_g) \quad (50)$$

$$\mathbf{1} \cdot z_0 = 1 \quad (\lambda_0) \quad (51)$$

$$z_k = \bar{z}_k \quad \forall k \quad (w_k) \quad (52)$$

$$(x_k, z_k) \in C_k, \forall k \quad (x_k^*, z_k^*) \quad (53)$$

Note that  $c_0 = 0, G_{10} = 0$  and  $G_{2k} = 0$  for  $k \neq 0$ . We slightly generalize O'Neill et al.'s Theorem 2 by relying on the extension of standard LP duality to (convex) conic programming duality (Ben Tal and Nemirowski (2001)) presented in Appendix A. Program  $\text{EPIP}_{\text{LOC}}(\bar{z})$  is a conic program which

comprises the conic constraint (53). As usual we write the dual variables of the constraints at the right of each of them. The dual variables  $x_k^*, z_k^*$  are associated to  $x_k$  and  $z_k$  appearing in the conic constraint (53). The conic dual of  $\text{EPLIP}_{\text{LOC}}(\bar{z})$  can be written as

**Problem EDLIP<sub>LOC</sub>( $\bar{z}$ )**

$$\min \lambda_0 + \sum_k w_k \bar{z}_k \quad (54)$$

$$\text{s.t.} \quad u_0 E_0 = x_0^* \quad (55)$$

$$-u_0 E_k - \lambda_g G_{1k} - c_k \text{sign}_k = x_k^* \quad (56)$$

$$\lambda_f F_0 + \lambda_g G_{20} + \lambda_0 + w_0 + d_0 = z_0^* \quad (57)$$

$$-\lambda_f F_k + \lambda_0 + w_k - d_k \text{sign}_k = z_k^* \quad (58)$$

$$(x_k^*, z_k^*) \in C_k^*, \quad (59)$$

$$\lambda_f, \lambda_g \geq 0 \quad u_0, \lambda_0, w_k \quad \forall k \text{ unconstrained} \quad (60)$$

Our interest in this extended model is to find prices that induce the TSO to invest into the appropriate network configuration and the generators and consumers to invest in the right location and operate efficiently.

## 6.4 Short-term locational charges

Before getting into this development, we first interpret the different constraints of the dual of  $\text{EPLIP}_{\text{LOC}}(\bar{z})$  and particularly those involving dual cones. Consider first the dual of the cone  $C_0$  associated to the TSO. Recalling the relations (34) till (36), one can write

$$C_0 \equiv \prod_n \{(x_0^n, z_0^n) \mid \text{PTDF}^n x_0^n \leq b^n z_0^n, \mathbf{1} \cdot x_0^n = 0 \text{ if } z_0^n > 0, \\ x_0^n = 0 \text{ if } z_0^n = 0\}.$$

The following lemma characterizes the dual cone  $C_0^*$ .

**Lemma 1**

$$C_0^* \equiv \prod_n C_0^{*n} \equiv \prod_n \{(x_0^{*n}, z_0^{*n}) \mid x_0^{*n} = \mu^n \underline{1} - \lambda^n \text{PTDF}^n, \\ z_0^{*n} \geq \lambda^n b^n, \text{ for some } \mu^n \text{ and some } \lambda^n \geq 0\}.$$

**Proof.** Consider the following linear program

$$\min x_0^{*n} x_0^n + z_0^{*n} z_0^n \quad (61)$$

$$\text{s.t. } -\text{PTDF}^n x_0^n + b^n z_0^n \geq 0 \quad \lambda^n \quad (62)$$

$$\underline{1} \cdot x_0^n = 0 \quad \mu^n \quad (63)$$

$$x_0^n \text{ unconstrained, } z_0^n \geq 0 \quad (64)$$

where the dual variables of the constraints are  $\lambda^n$  and  $\mu^n$  respectively. The dual of this problem can be restated as

$$\max 0$$

$$\text{s.t. } -\lambda^n \text{PTDF}^n + \mu^n \underline{1} = x_0^{*n} \quad x_0^n \quad (65)$$

$$\lambda^n b^n \leq z_0^{*n} \quad z_0^n \quad (66)$$

$$\lambda^n \geq 0, \mu^n \text{ unconstrained}$$

$(x_0^{*n}, z_0^{*n})$  belongs to  $C_0^{*n}$  iff the minimal value of the primal problem is attained and is zero. This happens if and only if there exists a dual solution  $\lambda^n \geq 0$ ,  $\mu^n$  unconstrained, such that  $x_0^{*n} = \mu^n \underline{1} - \lambda^n \text{PTDF}^n$  and  $z_0^{*n} \geq \lambda^n b^n$ . This completes the proof.

The proof of the lemma immediately suggests an interpretation of the constraint  $(x_0^{*n}, z_0^{*n}) \in C_0^{*n}$ . The condition (65) indeed implies that  $x_0^{*n}$  is a set of nodal prices where  $\mu^n$  is the electricity price at some hub and  $\lambda^n$  are the prices of the flowgate capacities in network configuration  $n$ . (66) implies that  $z_0^{*n}$  is bounded below by the merchandizing surplus. Because the value of the primal and the dual are equal, one also has

$$x_0^n x_0^{*n} + z_0^n z_0^{*n} = 0$$

and by duality  $z_0^{*n} = \lambda^n b$  when  $z_0^n > 0$ ;  $z_0^{*n}$  is then exactly equal to the merchandizing surplus. The above can be summarized by saying that  $(x_0^{*n}, z_0^{*n}) \in C_0^{*n}$  iff  $x_0^{*n}$  is a vector of nodal prices associated to flow  $x_0^n$  in configuration  $n$  and  $z_0^{*n}$  is the corresponding revenue of the TSO accruing from congestion charges.

## 6.5 The TSO's problem

We are now equipped to interpret the role of the TSO in this set up. Denote with an upper  $\bar{\cdot}$  the dual variables that are solutions of the dual problem  $\text{EDLIP}_{\text{LOC}}(\bar{z})$ . Consider the TSO's problem  $\text{CIP}_0$

$$\max \nu_{\text{CIP}_0} = -\bar{u}_0 E_0 x_0 - (d_0 + \bar{\lambda}_f F_0 + \bar{\lambda}_g G_{20} + \bar{w}_0) z_0 \quad (67)$$

$$(x_0, z_0) \in C_0 \quad (68)$$

$$\underline{1} \cdot z_0 = 1 \quad z_0 \in \{0, 1\}^N \quad (69)$$

Note by an upper  $\bar{\cdot}$  and a  $'$  the optimal solution of  $\text{CIP}_0$  (e.g.  $\bar{x}'_0$ ). We briefly discuss the interpretation of the different expressions appearing in this problem.

Note first that  $c_0$  is identically zero because the TSO does not have variable operating costs. By (55),  $\bar{u}_0 E_0$  is equal to  $x_0^*$  and hence, as shown in Lemma 1 is a vector of nodal prices in network configuration  $n$ . As a result,  $-\bar{u}_0 E_0 \bar{x}'_0 = -\bar{x}_0^{*n} \bar{x}'_0{}^n = \bar{z}_0^{*n} \bar{z}'_0{}^n = \lambda^n b^n$  is the merchandizing surplus in network configuration  $n$ , and hence the revenue of the TSO in that configuration.  $d_0$  is the vector of investment costs of the different network configurations.  $\bar{w}_0$  is a vector of payments/charges imposed by the Regulator to induce the TSO to select the adequate network configuration. Last  $-\bar{\lambda}_f F_0$  and  $-\bar{\lambda}_g G_{20}$  are locational payments respectively resulting from the generators/consumers decision to locate and reserve certain injection and withdrawal capacities. These two last terms do not carry much intuitive interpretation. This is not really surprising: one could not expect to find an easy interpretation of the cost



causality when the causal relationships that drive the development of the grid are themselves murky. We shall come back to these terms later in the discussion. In conclusion, relations (67) to (69) represent the problem of a TSO that maximizes the profit accruing from short and long-term payments when investing in the grid.

## 6.6 The generator and consumer's problem

We now turn to the sets  $C_k$ . Recall that  $C_k$  is defined as

$$C_k \equiv \{(x_k, z_k) \mid m_k z_k - x_k \geq 0, x_k \geq 0, z_k \geq 0\}.$$

We have the following lemma

### Lemma 2

$$C_k^* \equiv \{(x_k^*, z_k^*) \mid z_k^* \geq 0, \frac{z_k^*}{m_k} + x_k^* \geq 0\}.$$

**Proof.** Consider the following linear program

$$\min z_k^* z_k + x_k^* x_k \tag{70}$$

$$\text{s.t.} \quad m_k z_k - x_k \geq 0 \quad \lambda_k \tag{71}$$

$$z_k \geq 0, x_k \geq 0 \tag{72}$$

The corresponding dual can be stated as

$$\max 0$$

$$\text{s.t.} \quad \lambda_k m_k \leq z_k^*$$

$$-\lambda_k \leq x_k^*$$

$$\lambda_k \geq 0$$

$(x_k^*, z_k^*) \in C_k^*$  iff the primal problem has a solution and the objective function value is equal to zero. This requires that the dual problem also has a solution and hence that

$$x_k^* + \frac{z_k^*}{m_k} \geq 0 \text{ and } z_k^* \geq 0.$$

This completes the proof of the lemma.

Note that  $z_k > 0$ ,  $x_k > 0$  implies by the complementarity conditions that  $\lambda_k m_k = z_k^*$  and  $-\lambda_k = x_k^*$  and hence that  $x_k^* + \frac{z_k^*}{m_k} = 0$ .

We are now equipped to interpret the actions of agent  $k$ . This latter solves the problem

$$\begin{aligned} \max \nu_{\text{CIP}_k} &= (c_k \text{sign}_k + \bar{\lambda}_g G_{1k} + \bar{u}_0 E_k) x_k + (d_k \text{sign}_k + \bar{\lambda}_f F_k - w_k^*) z_k \\ \text{s.t. } &(x_k, z_k) \in C_k. \end{aligned}$$

Denote again with an upper  $\bar{\cdot}$  and a  $'$  the optimal solution of problem  $\text{CIP}_k$  (e.g.  $\bar{x}'_k$ ). The different expressions appearing in this problem can be interpreted as follows.  $c_k \text{sign}_k \bar{x}'_k$  is the generation cost or the willingness to pay of agent  $k$ .  $\bar{u}_0 E_k \bar{x}'_k$  is the revenue collected/charge paid by agent  $k$  for injection/withdrawal of electricity in a nodal price system.  $\bar{\lambda}_g G_{1k} \bar{x}'_k$  is the demand charge component due to the long-term locational prices. It can be positive or negative depending on the matrix  $G_{1k}$ .  $d_k \text{sign}_k \bar{z}'_k$  is the cost incurred by agent  $k$  in order to locate on the network and to build the equipment.  $\bar{\lambda}_f F_k \bar{z}'_k$  is the charge accruing from the locational decision. This charge can also be positive or negative depending on the matrix  $F_k$ .  $w_k^* \bar{z}'_k$  is a charge/payment levied by the Regulator in order to induce agent  $k$  to select the right location. Again  $\text{CIP}_k$  represents the problem of an agent that maximizes the profit accruing from short and long-term payments when deciding to locate and to operate on the grid.

## 6.7 Compatibility between the TSO and other agents decisions

One needs to show that the decisions of the generators and consumers, when solving their respective subproblems, combine to give energy flows that satisfy the constraints of the configuration selected by the TSO. One also need to explore whether financial flows balance. Last, one also wishes that the obtained behaviours do not depart from those found in the usual theory of congestion

management. These properties are obtained from the solutions of the primal and dual problems  $EPIP_{LOC}(\bar{z})$  and  $EDLIP_{LOC}(\bar{z})$  as we shall now show. We begin with the less usual result coming from conic duality.

The following lemmas cast our model in the standard theory of nodal pricing and hence directly relate to congestion management.

**Lemma 3** *Let  $(x'_0, z'_0)$  and  $(x'^*_0, z'^*_0)$  be respectively part of the primal and dual optimal solutions of  $EPLIP_{LOC}(\bar{z})$  and  $EDLIP_{LOC}(\bar{z})$ . One has*

$$x'^n_0 = 0 \text{ if } z'^n_0 = 0. \quad (73)$$

*If  $z'^{\bar{n}} > 0$ , then  $x'^{\bar{n}}_0$  solves*

$$\max x_0^{*\bar{n}} x_0^{\bar{n}} \quad (74)$$

$$\text{s.t.} \quad PTDF^{\bar{n}} x_0^{\bar{n}} \leq b^{\bar{n}} \quad (75)$$

$$\underline{1} x_0^{\bar{n}} = 0 \quad (76)$$

*where  $x_0^{*\bar{n}}$  is a vector of nodal prices in network configuration  $\bar{n}$ .*

**Proof.** Let  $\bar{n}$  be the network configuration selected by the TSO ( $z_0^{\bar{n}} = 1$  in (52)). (73) follows immediately from the definition of  $C_0$ . The result for  $n = \bar{n}$  follows by noting that the problem (74), (75) and (76) is identical to the problem (61) to (64) of Lemma 1 after replacing  $z_0^{\bar{n}}$  by 1.  $x_0^{*\bar{n}}$  can be interpreted as a vector of nodal prices as shown in Lemma 1.

Lemma 3 recovers a fundamental result of congestion management in nodal pricing (Hogan (1992)). Given a structure  $\bar{n}$  of the network,  $\bar{x}_0^{*\bar{n}}$  is the vector of nodal prices and the TSO selects an offer  $x_0^{\bar{n}}$  of injection and withdrawal services that maximizes the value of the network. (73) imposes that the energy flows can only be different from zero in the selected configuration  $\bar{n}$ . (75) imposes that the thermal limits of the lines in the selected configuration are not violated. Altogether problem (74)–(75) imposes that the system operator maximizes the value of the network capacities, using the nodal prices  $\bar{x}_0^{*\bar{n}}$ .

**Lemma 4**  $\bar{u}_0 E_k \bar{x}_k$  is the congestion charge paid by agent  $k$ .

**Proof.** As argued in the discussion following Lemma 1, the constraint (55) at optimum boils down to

$$u_0 = x_0^{*\bar{v}}.$$

The result then follows by noting that  $E_k = \text{sign}_k$ .

Combining Lemmas 3 and 4, one sees that the congestion management part of the problem is identical to the one described in the usual theory of nodal pricing when the network configuration is given.

In contrast, the remaining expressions  $\bar{\lambda}_f^T F_0$ ,  $\bar{\lambda}_f^T F_k$ ,  $\bar{\lambda}_g G_{20}$  and  $\bar{\lambda}_g^T G_{1k}$  do not have any standard interpretation. There are the “cost causal” parts of the tariffs that “price” the impact of the decision to locate and to reserve capacity on the network cost incurred by the TSO. The lack of interpretation comes from the fact that constraints (30) and (31) are of a combinatorial nature and do not have any particular economic interpretation. But we shall see that they induce the right behaviour on the TSO and the agents  $k$ . In no way do these “locational prices” satisfy the criterion of transparency demanded by the Regulation. But that should not be expected, as the network expansion process is itself a complex problem that is not transparent. Before exploring the global compatibility of the behaviour of the different agents, we note that complementarity slackness implies at least some accounting relations between the charges imposed on the agents and the revenue of the TSO.

**Lemma 5** (i)  $\bar{\lambda}_f F_0 z'_0 = \sum_{k \neq 0} \bar{\lambda}_f F_k z'_k$ . The sum of the locational charges  $\bar{\lambda}_f^T F_k z'_k$  paid by the generators and consumers is equal to the locational revenue  $\bar{\lambda}_f^T F_0 z'_0$  received by the TSO.

(ii)  $\bar{\lambda}_g G_{20} z'_0 = \sum_k \bar{\lambda}_g G_{1k} x'_k$ . The sum of the locational charges  $\bar{\lambda}_g^T G_{1k} x'_k$  paid by the generators and consumers is equal to the locational revenue  $\bar{\lambda}_g^T G_{20} z'_0$  received by the TSO.

**Proof.** Obvious from (49), (50) and complementarity conditions.

## 6.8 Decentralization of the decisions

It remains to show that the decentralized solutions of the TSO problem and of the different agent  $k$  problems solve the Regulator's problem, that is, that they induce the TSO to select the network configuration desired by the Regulator and similarly that they incentivize generators and consumers to locate and develop as desired by the Regulator. Last we want to be sure that the electricity and transmission markets clear.

The following proposition is a direct consequence of Theorem 1 stated in Appendix A.

**Proposition 2** *Suppose all agents  $k$ , when solving problem  $CIP_k$ , are charged/pay the price  $\bar{u}_0, \bar{\lambda}_f, \bar{\lambda}_g, \bar{w}_k$  found in the solution of  $EPLIP_{LOC}(\bar{z})$ . Then the solutions found by these agents are identical to the solution of problem  $EPLIP_{LOC}$ . Moreover energy markets and the transmission service market clear.*

**Proof.** The result immediately derives from Theorem 2 in Appendix A.

This implies the following corollary.

**Corollary 1** *Lemmas 3, 4 and 5 hold if the solutions of problem  $EPLIP_{LOC}(\bar{z})$  are replaced by those of the  $CIP_k$ .*

This implies that prices  $\bar{u}_0, \bar{\lambda}_f, \bar{\lambda}_g, \bar{w}_0$  and  $\bar{w}_k$  allow one to fully decentralize decisions among agents. Congestion management boils down to nodal pricing and retains its usual properties. TSO maximizes the value of the network and collects the merchandizing surplus. Also the demand charges components of the location prices paid by the users of the network and received by the TSO balance.

## 7 Non discriminatory prices

The Regulation and the work of the Florence Regulatory Forum request non discriminatory prices in order to induce competition. The above tariffs are only partially non discriminatory. The line congestion charges in the vector  $\bar{u}_0$  are the same for all agents  $k$  and hence are non discriminatory. This is the usual result of congestion management by nodal prices. The same is true for the demand charges of the long-term locational signals, at least if one accepts the description of causality given by relations (30) and (31). In contrast the  $w_k$  are truly discriminatory. This discrimination is justified by the objective of economic efficiency. In contrast with the common wisdom underlying both the Regulation and the work of the Forum, there are indeed situations where discrimination is necessary for achieving efficiency. Ramsey pricing is the best known illustration of the usefulness of discriminatory prices in the network industries: its justification is that it makes everyone better off. Still it remains a discrimination and hence may be unlawful. One may wonder whether one can obtain non discriminatory price, needless to say, at the cost of losing some economic efficiency. This is achieved by constructing a model where one explicitly imposes that the  $w_k$  are equal for all generators and consumers. We briefly turn to that question.

Recall that the model of Section 6 leads to three parts locational prices that are efficient and cost reflective. The congestion and demand charges are non discriminatory; the locational price is cost reflective and non discriminatory but the locational price (the  $w$ ) remains both non cost reflective and discriminatory. The question we address here is the removal of the discrimination. We first take up the question on the simple model of Section 4 and then extend the discussion to the more involved model of Section 6.

Consider the problem  $\text{PIP}_{\text{LOC}}$  and the dual problem  $\text{DLIP}_{\text{LOC}}(\bar{z})$  of problem  $\text{LIP}_{\text{LOC}}(\bar{z})$ . We modify  $\text{DLIP}_{\text{LOC}}(\bar{z})$  into  $\text{CDLIP}_{\text{LOC}}(\bar{z})$  (Con-

strained Dual LIP problem) by imposing the non discrimination constraint that all locational charges are identical. This leads to

$$\min u_0 b_0 + w \sum_k \bar{z}_k \quad (77)$$

$$\text{s.t.} \quad u_0 A \text{ sign}_k + u_k \geq c_k \text{ sign}_k \quad \forall k \quad (78)$$

$$-u_k m_k + w \geq d_k \text{ sign}_k \quad \forall k \quad (79)$$

$$u_0 \geq 0, \quad u_k \geq 0, \quad w \text{ unconstrained.} \quad (80)$$

The only difference between  $\text{CDLIP}_{\text{LOC}}(\bar{z})$  and  $\text{DLIP}_{\text{LOC}}(\bar{z})$  is the implicitly introduction of the non discrimination constraint

$$w_k - w = 0 \quad \forall k$$

that expresses that locational charges are identical for all agents.

The corresponding primal problem is written  $\text{CPLIP}_{\text{LOC}}(\bar{z})$

$$\max \sum_k c_k \text{ sign}_k x_k + \sum_k d_k \text{ sign}_k z_k \quad (81)$$

$$\text{s.t.} \quad A \sum_k \text{ sign}_k x_k \leq b_0 \quad (82)$$

$$x_k - m_k z_k \leq 0 \quad \forall k \quad (83)$$

$$\sum_k z_k = \sum_k \bar{z}_k \quad (84)$$

$$x_k \geq 0. \quad (85)$$

Note that the constraint  $x_k - m_k z_k \leq 0$  guarantees that  $z_k \geq 0$ . It is obvious that problem  $\text{CPLIP}(\bar{z})$  is a relaxation of  $\text{PIP}_{\text{LOC}}(\bar{z})$  because it aggregates the objectives  $z_k = \bar{z}_k \quad \forall k$  of the Regulator into the single objective (84). Strictly speaking, this does not solve the Regulator's problem but only approximates it. But it may be a very good approximation as we discuss now. The solution of  $\text{PIP}_{\text{LOC}}(\bar{z})$  aims at selecting the optimal choice of the capacities for each agent in different locations. It is unlikely that the Regulator has enough information to differentiate between the cost structure of the different agents. Suppose indeed that the Regulator cannot differentiate between the different agents.

He is then only able to select the best mix of technologies and capacities at the different locations. The vector  $\sum_k \bar{z}_k$  describes the total capacity of each technology built at each location. The price signal  $\bar{w}$  therefore depends on both a location and a technology but not on the agent. As a locational signal, it may still be considered as discriminatory to the extent that it differentiates the locational price as a function of the technology. This may be forbidden by law except if one can argue that this differentiation is justified on the basis of cost causality. For instance gas and coal plants are not expected to operate in the same way and hence could imply different investments in the network. But at least the approach does not differentiate the locational signal by agent, something that would certainly be forbidden.

The problem of this formulation is that the prices found by solving problem (77) to (80) do not guarantee that the constraint (84) will be met in the decentralization process. Agents may still build too much or too little capacity. This is the welfare loss implied by the non discrimination constraint.

This discussion can be extended to the model of Section 6. We consider a slightly different (but equivalent) formulation of the non discrimination problem based on that model.

$$\begin{aligned}
\inf_w \quad & \max \sum_k c_k \text{sign}_k x_k - (d_0 + w_0)z_0 + \sum_{k \neq 0} (d_k \text{sign}_k + w)z_k \\
& E_0 x_0 - \sum_{k \neq 0} E_k x_k & = 0 \\
& F_0 z_0 - \sum_{k \neq 0} F_k z_k & \leq 0 \\
& - \sum_{k \neq 0} G_{1k} x_k & + G_{20} z_0 & \leq 0 \\
& & 1 \cdot z_0 & = 1 \\
& (x_k, z_k) \in C^k \\
& & z_k \in \{0, 1\}^{n(k)}
\end{aligned}$$

This model is obtained from  $\text{PIP}_{\text{LOC}}(\bar{z})$  by two operations

- (i) one dualizes the constraint (52)  $z_k = \bar{z}_k$



(ii) one imposes that the prices associated to these constraints are the same for all generators and consumers.

One immediately sees that (ii) is the non discrimination constraint that imposes that all  $w$  are identical in the objective function of the problem.

This formulation will guarantee non discriminatory locational prices but it will not ensure that the decentralization process will lead to locations and investments that meet the objective  $\sum_k \bar{z}_k$  through  $z_k = \bar{z}_k$  of the Regulator. This is the same phenomenon as mentioned for the preceding model. Generators and consumers may over or under invest, therefore entailing a loss of economic efficiency.

It is impossible to a priori foresee the extent of the loss of efficiency that would result from this non discrimination constraint on the sole basis of the model formulation. But it is certainly quite doable to investigate the question numerically.

Note that an obvious improvement of both models would be to differentiate the locational charge depending on whether agent  $k$  injects or withdraws at some node. This implies defining  $w^+$  as the locational charge for all generators and  $w^-$  for all consumers. This extension is straightforward.

## 8 Institutional discussions

The above discussion only offers an abstract setting that it is useful to try to relate to current practical proposals. We again rely on the discussion provided by Rosellon (2003) that was already invoked in Section 3.

It should be clear that our formulation of the TSO problem best fits the Transco model. In this model a single company is in charge of both the investments and the operations of the transmission system. This proposal, which is

favored in Joskow and Tirole (2002) is implemented in the United Kingdom. Combined with an appropriate incentive regulation, this experience is credited to be very successful. The current European institutional framework makes no recommendation in favour of that system and some of the language of the Regulation (Article 6, paragraph 6) even seems not to recommend it. What is sure is that it is impossible to impose a Transco like solution throughout Europe, let alone to adopt a single Transco for the whole of Europe. Still it remains useful to retain this interpretation both because it is implemented in practice and that implementation is successful.

Our model obviously differs from the exact implementation of the Transco in the UK. In this paper, the TSO receives instructions on how to set the long-term signals; it also receives some money transfers via the  $w_0$  that induces it to select the right network configuration. The model also supposes a well informed, quite knowledgeable and very intrusive Regulator. Still, the Regulator does not need to solve problem  $PIP_{LOC}$  by himself. The Transco, or an independent consultant can do so on the basis of commonly agreed data and assumptions. But the Regulator is assumed to be comfortable with the whole process and to agree that the  $\bar{z}_k$  (or  $\sum_k \bar{z}_k$  in case of non discriminations) are desirable objectives.

The merchant line is the other approach to grid investments recalled in Section 3. It is analytically grounded in the theory of nodal prices and their extension into long-term financial transmission rights as hedging instruments. The ISO model, where operations and ownership of the grid are separated, provides the institutional background of the approach. Hogan et al.'s theory of merchant lines essentially extends the role of financial transmission rights from operations to both operations and investments. Joskow and Tirole (2002) argue that this can only be done under drastic assumptions that are violated in practice. We concentrate in this paper on one of these assumptions, namely

the lumpiness of investments, and explore its possible consequences. Our conclusion is that it is still possible to decentralise lumpy investment in the grid provided one invokes a more complex set of prices that covers not only congestion but also access charges. In other words, there exists a combination of access and congestion charges that provide the necessary incentives for investments. It is remarkable that this is exactly what the Regulation foresees, but without indicating how this can be done. The reality is that the derivation of these charges is a very demanding task. It indeed requires first to extend the role of the ISO from the sole operation of the existing system, to include both the operation and investments in the grid. This extension raises several questions that we do not discuss here; still it retains a key property of the ISO model in the sense that this latter does not need to own the grid. Our abstract model is thus also fully compatible with the notion of a large RTO. The creation of RTOs is still a long way off in the European context, but it is definitely more doable than a single Transco. The task of the RTO becomes formidable however. Instead of only auctioning physically feasible financial transmission rights that only covers congestion charges, this global RTO should now auction two types of long-term contracts, corresponding respectively to the access and congestion charges. In the same way as congestion charges need to be carefully tuned in order to "get the prices right", the access charges also need to be well tuned in order to induce the right investments in the grid. Needless to say ownership of access rights on top of congestion rights by generators as a result of the auction would exacerbates generator's market power in the sense treated in Joskow and Tirole (2000) and Gilbert et al. (2004). The answer may lie in Hogan's suggestion (Hogan (2002)) that transmission companies would probably be the main owners of these rights.

## 9 Discussion and further questions

The locational prices obtained in Section 6 satisfy two objectives of the Regulation. They are economically efficient and cost reflective. They also allow for a separation between short-term and long-term locational prices as foreseen by the Regulation. Also all the theory of nodal price/flowgate congestion management remains unchanged. Still, these locational prices leave several open problems. We briefly introduce some of them.

*Revenue adequacy* is a key element of the theory of congestion management. It states that the TSO will normally receive a non negative profit from congestion management. This non negative profit occurs in a short run equilibrium problem, that is for given network capacities, provided certain conditions related to ex ante and ex post contingencies are met. We retain the standard result of revenue adequacy in congestion management operated under nodal pricing in our more general set up. Moreover, by Theorem 1 all agents make zero profit at equilibrium. This would suggest that revenue adequacy can be extended to encompass both long-term and short-term signals. But there is a major difference between this extended revenue adequacy and the standard congestion management result. Congestion payments between TSOs and agents  $k$  are in balance as can be seen by multiplying relation (48) by  $\bar{u}_0$ . Revenues and expenses from demand charges also balance as can be seen from the complementary condition of constraints (42) and (43). In contrast, there is no guarantee that  $\sum_k \bar{w}_k z_k$  is equal to zero and it is indeed unlikely that it will be so. We discussed in Section 5 how Hogan and Ring expanded the scope of prices capable of sustaining the equilibrium, therefore possibly allowing to reduce or eliminate the revenue imbalance  $\sum_k \bar{w}_k z_k$  faced by the Regulator. Their reasoning can be extended to the more general model of Section 6. All this should be studied further, probably numerically.

The *realism* of the models can be improved. The model of Section 6 considers that the system is operated during a single time period, therefore assimilating the demand charge to an additional energy charge. This restriction can easily be removed. It has been a tradition among power engineers to plan the network for the peak period. This is unlikely to be correct because transmission flows are not highest during the peak. This suggests to introduce a set of reference periods (seasons) that are critical for designing the network and hence that can be used for determining the demand charge. This can easily be done by extending the model of Section 6. This would allow for a richer description of the cost causality and hence would lead to a richer set of nonlinear tariffs. The same can be said about the inclusion of the choice of a technology. The  $z_k$  variables have been defined to only refer to locational decisions. As argued in Section 4, it is easy to expand their interpretation to encompass a technological choice. This would also lead to a richer set of nonlinear tariffs.

Last, one shall remember that the models discussed in this paper fail to achieve fully non discriminatory tariffs. Bjørndal and Jörnsten (2004) were able to construct non discriminatory dual price functions for the unit commitment problem treated by O'Neill et al. (2004). The possibility of extending their analysis to the grid problem constitutes an other area of interest.

The following issues may be more challenging. The current model supposes that the cost causality can be expressed through relations of the type (42) and (43). This is certainly true in principle but does not say anything on how to derive these relations. This problem raises both power engineering and optimization questions. Consider first the electrical engineering issue. A main feature of model EPIP<sub>LOC</sub> is to imbed a description of the *production set of the TSO*. This description involves both causal relations ((30), (31)) as well as a description of the transportation possibilities of the different network

configurations contemplated by the TSO ((33) and (34)). The representations of the production set of the TSO in economic models is unusual but not totally absent from the literature (Vogelsang (2001)). Our relations ((30), (31) and 33)) are stylized from the following considerations. We assume that the TSO selects its investments using a capacity expansion model. There are several examples of these models in the electrical engineering literature (see e.g. Latorre (2003) for a survey of these models). Many of these models are of the mixed integer programming type. We suppose that one did select one of these models. Using the language of combinatorial optimization, relations (30) and (31) are valid inequalities, derived from this mixed integer program, that describe the polyhedron of feasible solution of  $EPIP_{LOC}$ . The construction of these inequalities is currently an area of intense research in combinatorial optimization. It is not clear however that these inequalities have so far been derived for the type of model involved in network capacity planning. Also current valid inequalities are limited to local description of the feasible polyhedron. In short, the introduction of relations (30) and (31) makes a lot of sense from the point of view of logic, but their construction remains to be explored.

Consider now the following economic issue: The current model is established under assumptions of perfect knowledge and absence of market power. Relaxing each of these assumptions creates a whole set of additional complexities. Consider the assumption of *perfect knowledge* first. Following a traditional assumption of the “old” theory of regulation, the Regulator is supposed to know the cost of the generators, the willingness to pay of the consumers and the set of possible network configurations of the TSO. The limitations of this type of assumptions have been extensively explored in the new theory of regulation, using the notion of asymmetry of information. Embedding this machinery in extensions of Theorem 2 of the appendix would add a whole new dimension to the work.

*The absence of market power* is another limitation of the model. An extension of the model discussed here to one that would accommodate market power is straightforward to formulate: the problem  $\text{EPIP}_{\text{LOC}}$  should be replaced by a mathematical program subject to equilibrium constraint (MPEC). This model would request that the Regulator selects the long and short-term locational prices in order to induce agents operating in an oligopolistic market to behave in such a way that they maximize welfare. But it is not clear that this problem has a solution. Again, the real difficulty is to extend the decentralization result of Theorem 2 to that more general set up.

The *static nature* of the model is certainly one of the major shortcomings of the above discussion. Static models are common when discussing equilibrium problems but the simplification is a real drawback when it comes to implementation. Bushnell and Stoft (1996) have initialized the discussion of the dynamics of investments in the electrical grids and their consequences on the validity of existing Financial Transmission Rights. This has since been extensively elaborated in various papers, among them the discussion of merchant lines. It is totally absent from this set up.

## 10 Conclusion

The idea of separating long-term and short-term locational signals in the Regulation on cross border exchanges of electricity is probably a good one. It offers the opportunity to depart from simple linear tariffs that are not sufficient to induce the right investments and locational decisions in a system plagued by discrete decisions. But this separation leaves it open how to construct the long-term locational signals. It is easy but misleading to pretend to solve the problem by cost allocation rules. Except proven otherwise by numerical experiment, there is no reason to believe that these will generate the right signals. This paper considers the goal of finding an alternative approach or at least to

identify ideal abstract conditions that would allow for such an alternative. It retains three criteria that are mentioned more or less explicitly in the Regulation, namely economic efficiency, cost reflectiveness and non discrimination. It finds that the three criteria cannot be simultaneously achieved, even under ideal conditions. But one can at least trade non discrimination for the other criteria. It also finds that transparency of the long-term signals is probably hopeless. But this is not surprising as the network expansion process is itself a murky one. Last and probably more important it provides a framework where both the long-term and short-term signals can both be cast, and the usual theory of short-term signals remains completely unaffected.

## References

- [1] Ben-Tal, A. and A. Nemirovski (2001). *Lectures on Modern Convex Optimization*. Philadelphia, MPS-Series on Optimization.
- [2] Bjørndal, M. and K. Jörnsten (2004). Equilibrium supported by dual price functions in markets with non-convexities. Department of Finance and Management Science, Norwegian School of Economic School and Business Administration. Bergen, Norway.
- [3] Bushnell, J.B. and S.E. Stoft (1996). Electric grid investment under a contract network regime. *Journal of Regulatory Economics*, 10, 61–79.
- [4] Crema, A. (1995). Average shadow price in a mixed integer linear programming problem. *European Journal of Operational Research*, 85, 625–635.
- [5] Crew, M.A., Fernando, C.S. and P.R. Kleindorfer (1995). The theory of peak-load pricing: a survey. *Journal of Regulatory Economics*, 8, 215–248.



- [6] Curien, N. Cost allocation methods (2003). In F. Lévêque (ed.), *Transport Pricing of Electricity Networks*. London, Kluwer Academic Publishers, 73–101.
- [7] European Parliament and Council (2003). Regulation (EC) No 1228/2003 of the European Parliament and of the Council of 26 June 2003 on Conditions for Access to the Network for Cross-border exchanges in Electricity 26.6.2003, Official Journal L 176, 15/07/2003, pp. 0001-0010.
- [8] Hogan, W.W. (1992). Contract networks for electric power transmission. *Journal of Regulatory Economics*, 4, 211–242.
- [9] Hogan, W.W. (1998). Transmission investments and competitive electricity markets. Kennedy School of Government, Harvard University. Downloadable from website <http://www.whogan.com>
- [10] Hogan, W.W. (2002). Financial transmission right incentives: applications beyond hedging. Presentation at Harvard Electricity Policy Group. Downloadable from website <http://www.whogan.com>
- [11] Hogan, W.W. (2004). Transmission market design. Presentation at the conference “Electricity deregulation: Where to go from there”. Downloadable from website <http://www.hogan.com>
- [12] Hogan, W.W. and B.J. Ring (2003). On minimum-uplift pricing for electricity markets. Electricity Policy Group.
- [13] Gilbert, R., Neuhoff, K. and D. Newbery (2004). Allocating transmission to mitigate market power in electricity markets. *Rand Journal of Economics*, 35(4), 691-709.
- [14] Green, R. (2003). Cost recovery and the efficient development of the grid. In F. Lévêque (ed.), *Transport Pricing of Electricity Networks*. London, Kluwer Academic Publishers, 137–153.

- [15] Joskow, P. and J. Tirole (2000). Transmission rights and market power on electric power networks. *Rand Journal of Economics*, 31(3), 450–487.
- [16] Joskow, P. and J. Tirole (2003). Merchant transmission investment. CMI EP 24, available from CMI Electricity Power web site.
- [17] Latorre, G., Cruz, R.D., Areiza, J.M. and A. Villegas (2003). Classification of publication and models on transmission expansion planning. *IEEE Transactions on Power Systems*, 18(2), 938–946.
- [18] Lévêque, F. (2003). Legal constraints and economic principles. In F. Lévêque (ed.), *Transport Pricing of Electricity Networks*. London, Kluwer Academic Publishers, 3–33.
- [19] Mas-Colell, A., Whinston, M.D. and J.R. Green (1995). *Microeconomic Theory*, Oxford University Press.
- [20] O’Neill, R.P., Sotkiewicz, P.M., Hobbs, B.F., Rothkopf, M.H. and W.R. Steward Jr. (2004). Efficient market-clearing prices in markets with non convexities. *European Journal of Operations Research*.
- [21] Pérez-Arriaga, I., Olmos Camacho, L. and F.J. Rubio Odériz (2002). Report on cost components of cross border exchanges of electricity.
- [22] Pérez-Arriaga, J.I., Rubio, F.J., J.F. Puerta Gutiérrez, J. Arceluz and J. Marin (1995). Marginal pricing of transmission services: an analysis of cost recovery. *IEEE Transactions on Power Systems*, 10(1).
- [23] Pérez-Arriaga, I. and Y. Smeers (2003). In F. Lévêque (ed.), *Transport Pricing of Electricity Networks*. London, Kluwer Academic Publishers, 175–203.
- [24] Pope, S.L. and S.M. Harvey (2002). TCC awards for transmission expansions. Downloadable from website <http://www.pjm.com/services/trans/rtp-meetingnote3.html>

- [25] Rosellon, J. (2003) Different approaches towards electricity transmission expansion. *Review of Network Economics*, 2(3), 238–269.
- [26] Scarf, H.E. (1994). The allocation of resources in the presence of indivisibilities. *Journal of Economic Perspectives*, 8(4), 111-128.
- [27] Tirole, J. (1998). *The Theory of Industrial Organization*. Cambridge (Mass.), The MIT Press.
- [28] Vogelsang, I. (2001). Price regulation for independent transmission companies. *Journal of Regulatory Economics*, 20, 141–165.
- [29] Wolsey, L.A. (1998). *Integer Programming*. New York, John Wiley & Sons.
- [30] Woolf, F. (2003). *Global Transmission Expansion: Recipes for Success*. Penn Well, Tulsa.

## Appendix A: Methodological background

The following provides a small extension of the main theoretical result of O'Neill et al. The extension is appropriate for introducing cost causality in the network problem.

Consider the following mixed integer *conic* problem.

$$\max \nu_{CIP} = \sum_k c_k x_k + \sum_k d_k z_k \quad (\text{A.1})$$

$$\text{s.t.} \quad \sum_k A_{k1} x_k + \sum_k A_{k2} z_k \leq b_0 \quad (\text{A.2})$$

$$B_{k1} x_k + B_{k2} z_k \leq b_k \quad \forall k \quad (\text{A.3})$$

$$(x_k, z_k) \in C_k \quad z_k \in \{0, 1\}^{n(k)} \quad (\text{A.4})$$

where  $C_k$  is a convex cone. The only difference between this CIP problem and O'Neill et al.'s PIP is the replacement of the constraint  $x_k \geq 0$  by  $(x_k, z_k) \in C_k$ . Because  $x_k \geq 0$  also defines a convex cone CIP generalizes PIP. Let  $\tilde{C}_k = \{\tilde{x}_k, \tilde{z}_k \mid \tilde{x}_k x_k + \tilde{z}_k z_k \geq 0 \forall (x_k, z_k) \in C_k\}$  be the dual cone of  $C_k$ , one knows that conic programming satisfies the same duality properties as linear programming. We then extend the formalization of O'Neill et al. as follows. Let  $z_k^*$  be the value of the  $z_k$  dual variables in an optimal solution to CIP, define the following primal conic program

$$\begin{aligned} \text{PCIP}(z^*) \quad \max \nu_{\text{PCIP}} &= \sum_k c_k x_k + \sum_k d_k z_k \\ \text{s.t.} \quad &\sum_k A_{k1} x_k + \sum_k A_{k2} z_k \leq b_0 \quad \forall k \\ &B_{k1} x_k + B_{k2} z_k \leq b_k \quad \forall k \\ &z_k = z_k^* \quad \forall k \\ &(x_k, z_k) \in C_k \end{aligned}$$

Using duality theory in conic programming, one states the following dual problem

$$\begin{aligned}
\text{DCIP}(z^*) \quad \min \nu_{\text{DCIP}} = & y_0 b_0 + \sum_k y_k b_k + \sum_k w_k z_k^* \\
\text{s.t.} \quad & y_0 A_{k1} + y_k B_{k1} - c_k = x_k^* \\
& y_0 A_{k2} + y_k B_{k2} - d_k + w_k = z_k^* \\
& y_0 \geq 0 \quad y_k \geq 0 \\
& w_k \text{ unconstrained} \\
& (\tilde{x}_k, \tilde{z}_k) \in \tilde{C}_k
\end{aligned}$$

Theorem 1 of O'Neill et al. is readily extended into

**Theorem 1.**

$$\nu_{\text{CIP}}^* = \nu_{\text{PCIP}}^* = \nu_{\text{DCIP}}^*$$

where  $*$  indicates the optimal solution value for the respective problems.

The proof is identical to the one of O'Neill et al.

Consider now a set of price vectors  $(P_0, P_k^z)$ . One adapts the agent problem of O'Neill et al. into

$$\begin{aligned}
\text{CIP}_k \quad \max \nu_{\text{CIP}_k} = & (c_k x_k + d_k z_k) - P_0(A_{k1} x_k + A_{k2} z_k) - P_k^z z_k \\
& B_{k1} x_k + B_{k2} z_k \leq b_1 \\
& (x_k, z_k) \in C_k \quad z_k \in Z^k
\end{aligned}$$

We then define a competitive equilibrium as a set of prices  $\{P_0^*, P_k^{z*}\}$  for all  $k$  and allocations  $\{x_k^*, z_k^*\}$  for all  $k$  such that

1. At the prices  $\{P_0^*, P_k^{z*}\}$ , the allocations  $\{x_k^*, z_k^*\}$  solve  $\text{CIP}_k$  for all  $k$ , and
2. The market clears:  $\sum_k A_{k1} x_k + \sum_k A_{k2} z_k \leq b_0$ .

O'Neill's Theorem 2 can then be generalized as

**Theorem 2.** Let  $\{x_k^*, z_k^{*'}\}$  be the solution to  $\text{CIP}_k(z^*)$  and  $\text{PCIP}(z^*)$  and let  $\{y_0^*, y_k^*, w_k^*\}$  be the solution to  $\text{CDIP}(z^*)$ . If  $y_0^* = P_0$  and  $w_k^* = P_k^z$  then the prices  $\{y_0^*, w_k^*\}$  and allocations  $\{x_k^*, z_k^{*'}\}$  for all  $k$  is a competitive equilibrium.

**Proof.** The proof is almost identical to the one reported by O'Neill et al. The only difference is the replacement of the two complementarity conditions

$$\begin{aligned} 0 &\leq (y_0^* A_{k1} + y_k^* B_{k1} - c_k) \perp x_k \geq 0 && \forall k \\ 0 &\leq (y_0^* A_{k2} + y_k^* B_{k2} + w_k^* - d_k) \perp z_k \geq 0 && \forall k \end{aligned}$$

by the generalized complementarity condition

$$\tilde{C}_k \ni \begin{pmatrix} \tilde{x}_k \\ \tilde{z}_k \end{pmatrix} \perp \begin{pmatrix} x_k \\ z_k \end{pmatrix} \in C_k$$

which states that

$$\begin{aligned} &\begin{pmatrix} x_k \\ z_k \end{pmatrix} \in C_k \\ &\begin{pmatrix} y_0^* A_{k1} + y_k^* B_{k1} - c_k \\ y_0^* A_{k2} + y_k^* B_{k2} + w_k^* - d_k \end{pmatrix} = \begin{pmatrix} \tilde{x}_k \\ \tilde{z}_k \end{pmatrix} \in \tilde{C}_k \\ &(y_0^* A_{k1} + y_k^* B_{k1} - c_k)x_k + (y_0^* A_{k2} + y_k^* B_{k2} + w_k^* - d_k)z_k = 0 \end{aligned}$$

These generalized complementarity conditions are used to prove that  $\nu_{\text{CIP}_k}^{**} \leq \nu_{\text{CIP}_k}^*$  by noting that

$$(c_k - y_0^* A_{k1} - y_k^* B_{k1})x_k^{**} + (d_k - y_0^* A_{k2} - y_k^* B_{k2} - w_k^*)z_k^{**} \leq 0$$

because

$$\begin{aligned} &\begin{pmatrix} y_0^* A_{k1} + y_k^* B_{k1} - c_k \\ y_0^* A_{k2} + y_k^* B_{k2} + w_k^* - d_k \end{pmatrix} \in \tilde{C}_k \\ &\begin{pmatrix} x_k^{**} \\ z_k^{**} \end{pmatrix} \in C_k \end{aligned}$$

and the definition of  $\tilde{C}_k$ .