Chapter 2
Transmission Pricing

Ignacio J. Pérez-Arriaga, Luis Olmos, and Michel Rivier

2.1 Introduction

The transmission grid has a major impact on the operation and investment decisions in electric power systems. This impact is more noticeable when the electricity sector is organized around a wholesale market, where the transmission network becomes the meeting point of producers and consumers. The relevance of transmission is presently increasing with the growing penetration of intermittent renewable energy sources, frequently distant from the main load centres and significantly adding to the variability of flow patterns.

This chapter examines the economic impact of the transmission network on its users. This impact is twofold. On the one hand the network modifies the bulk prices of electrical energy, due to the presence of network losses and congestions. On the other hand, the costs of investment and operation of the transmission network have to be allocated to its users, according to some reasonable criterion. In principle both impacts should have a locational component. Injections or withdrawals of power in the grid affect losses and constraints differently depending on the node where they occur. Besides, the responsibility of network users in the reinforcements to the network generally depends on the location of these generators and loads. Thus, the allocation of the cost of the grid to its users should be guided by the location of the latter.

The chapter starts by discussing in Sect. 2.2 the effect of the transmission grid on system operation costs: how network constraints modify the economic dispatch of generation plants, and the costs of transmission losses. Section 2.3 presents the concept of nodal prices (locational marginal prices) and how to compute them.

I.J. Pérez-Arriaga (✉)
Comillas Pontifical University, Alberto aguilera 23, 28015 Madrid, Spain

Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
e-mail: ignacio@iit.upcomillas.es; ipa@mit.edu

L. Olmos • M. Rivier
Comillas Pontifical University, Alberto aguilera 23, 28015 Madrid, Spain

J. Rosellón and T. Kristiansen (eds.), *Financial Transmission Rights*,
Lecture Notes in Energy 7, DOI 10.1007/978-1-4471-4787-9_2,

The main properties of nodal prices are explained in Sect. 2.4. Section 2.5 describes how the impact of the transmission network on electricity energy prices is accounted for in practice in different power systems. Finally, Sect. 2.6 examines the allocation of the transmission network costs among the network users in the form of regulated network tariffs Sect. 2.7 concludes.

2.2 The Effect of Transmission on System Operation Costs

The unavoidable physical limitations of transmission networks when connecting producers to consumers have three main undesirable effects on the operation of the system. First, part of the energy transmitted over the lines and other grid facilities is transformed in heat and, therefore, never reaches the consumption centres. The difference between the amount of power injected at one end of a line and that withdrawn at the other end is called the loss of power in the line, or line **loss**. Second, the transmission grid imposes **constraints** due to a variety of technical reasons on any given set of power transactions that the network users want to make happen. Third, problems affecting the integrity and well functioning of the grid may result in the interruption of power supply to certain (or all) loads or the deterioration of the quality of electricity supplied. Thus, the **quality of the electricity supply service** may be affected by the grid.

Therefore the transmission grid may affect both the operation costs and the set of power injections and withdrawals that are allowed to take place. Conversely, the specific location of generators and loads in the grid is the driving factor behind the need for network expansion, which tries to improve the reliability of electricity supply and to reduce the operation costs derived from losses and network constraints. The locational differentiation of energy prices and network charges sends locational signals to prospective network users to be located so that these adverse effects are minimized.

Understanding the origin of network related operation costs, as well as the main drivers behind these costs and their impact on the system economic dispatch is of essence. Next, each of the three main effects of the grid on the system operation and its costs is discussed separately.

2.2.1 Network Losses

Most of the energy losses in electric power grids are due to the resistance of conductors to the circulation of electric current flows. These are known as **ohmic losses**. Other losses are due to the **corona effect** whereby electrical discharges take place in the air surrounding high voltage line conductors. Losses also occur within network devices like **transformers, reactors, capacitors**. Due to existing losses, consumers receive less energy than generators produce.

Transmission network losses result in additional system costs. More energy has to be produced than is consumed, because part is lost while being transported. These costs correspond to additional production costs, i.e. they are not network costs per se, though they are a consequence of the need to transport power over the transmission grid. The cost of losses is affected by transmission expansion and operation decisions. It is therefore advisable to set efficiency incentives encouraging the System Operator and network users to reduce these costs.

Ohmic losses in a line are nearly proportional to the square of the power flow over the line (more precisely, they are proportional to the square of the current in the wires). This means that the increase in losses per unit increase in the system load (marginal increase in losses) is approximately twice as large as the average amount of losses per unit of load (total amount of losses/total system load). Consequently, the marginal cost of transmission losses (transmission losses cost increase/increase in system load) exceeds their average cost (total cost of losses/total system load).

The increase in transmission losses in the system due to a marginal increase in the load at a certain node depends on the location of this node in the grid, since the resulting changes in line flows depend on the latter as well. Therefore, transmission losses create geographic differences in the marginal cost of supplying electric energy. This implies that the marginal cost of meeting a marginal increase in demand can only be correctly assessed if the exact node where demand is increased is specified. Other factors contributing to these differences are described in the next subsection.

Due to transmission losses, some power plants may take precedence in the merit order of the economic dispatch over other plants whose production costs are lower. The merit order of power plants in the dispatch must be affected by the loss factor corresponding to each plant according to its location in the grid.[1]

2.2.2 Network Constraints

Networks restrict in many ways the power transactions that can take place in the system. Most typically, transactions cannot result in a current intensity (roughly proportional to the power flow, for a given voltage level) over any line that exceeds the maximum one that can be handled by this line. The underlying reason to limit the current intensity over a transmission line may be thermal – and therefore dependent upon the physical characteristics of the facility – or related to the conditions of system operation as a whole, like the provisions to guarantee an appropriate system dynamic response to disturbances or to avoid stability related problems that usually increase with the length of lines. Another typical grid constraint is the need to keep **voltages** within certain limits at all nodes, which may call for having some generation unit connected near the node experiencing problems. The maximum allowable **short-circuit power** may also limit grid configuration. Generally speaking, the chief effect of grid constraints is to condition system operation, leading to deviations from the most

[1] The loss factor at a certain node represents the increase in transmission losses in the system resulting from a unit increase in the power injected at this node. Loss factors depend on the existing system operation conditions.

efficient one from an economic point of view. Most common constraints in distribution grids are related to voltage limits and maximum line capacities.

Just as in the case of network losses, the mere existence of the transmission network adds to system costs by requiring the dispatch of more costly generation units to surmount the physical limitations imposed by the grid. This does not imply that network design or development is flawed, since network investments required to ensure the total absence of constraints in the system would probably not be economically justified. Some network constraints may therefore be justified from an economic point of view (provided that they do not systematically prevent the coverage of demand).

The cost of grid constraints, like that of losses, corresponds to additional generation costs that are associated with the characteristics of the network. Therefore, these costs are not part of the cost of the network itself. Operation and expansion decisions may affect the cost of grid constraints, which advises sending economic signals encouraging parties in the system to reduce this cost.

Both losses and grid constraints result in changes in the economic dispatch. The merit order of generation units depends not only on their production costs but also on their location in the grid and their impact on losses and grid constraints. The marginal cost of supplying load depends on the location in the grid of the former and therefore, may vary from one node to another. Additional costs associated with losses and constraints must be assigned to network users.

As explained below, nodal prices applied to the electric energy sold or purchased are economic signals that efficiently internalize all the short-term effects of the network on electricity supply costs. Due to their relevance, next Sects. 2.3 and 2.4 are devoted to discussing nodal prices and their properties.

2.2.3 Quality of Service

Transmission networks have also an impact on the **quality of the electricity supply** service. In countries where the electricity system is well developed, generation outages or lack of total generation capacity are hardly ever responsible for electricity supply interruptions. In a small percentage of cases, the origin of interruptions lies in joint generation and transmission security failures (although the consequences of such events are usually very severe, since they affect large areas in the system). Supply disruptions are in fact practically always due to local distribution grid failures. Distribution business regulation should strike a balance between the cost of developing the grid and the resulting enhancement of end consumer quality of service. The effect of the transmission grid on the quality of service is not so notorious and will not be discussed further in this section.

2.3 Nodal Prices: Definition and Computation

Losses and grid constraints result in differences in the local marginal value of energy among transmission nodes.[2] Locational energy prices affect the short and long term efficiency of the functioning of the system by driving market agent decisions on how much power to produce or consume at each time, as well as where to site the new generation or load they plan to install, which may in turn affect the development of the transmission network.

Short-term locational energy prices also vary over time. Separate prices are computed for each hour in day-ahead markets and in some power systems they are also computed as close as several minutes ahead of real time. Signals sent through these prices are needed to achieve maximum system efficiency. They aim to ensure that the generators with the lowest variable costs are the ones dispatched and demand can respond to the actual costs of supplying energy at each location. Besides, these signals also drive the expansion of the system, since expectations about future values of energy prices at the different locations affect market agents' long-term decisions on the siting of new generation and demand facilities.

2.3.1 Concept of Nodal Prices

Nodal pricing represents the most sophisticated and efficient expression of locational energy prices. The marginal cost of electricity in a system corresponds to the extra cost incurred to serve a differential increase in the system load. It can be demonstrated that pricing the electricity produced or consumed in each node at the local marginal cost leads generators and loads in the system to make efficient operation decisions.

As a result of the existence of the grid, the marginal cost of electricity varies from one node to another. The nodal electricity price, also called locational marginal electricity price, in each node k is the short term cost of supplying most economically a marginal increase in demand in this node while complying with grid constraints. Nodal energy prices can be computed both for active and reactive power, as discussed in (Schweppe et al. 1988). However, nodal prices of reactive power have not been used in any real life system.[3]

When taking into account the actual features of electricity systems, which obviously must include the transmission network, any computed marginal system costs must be node specific. The uniform marginal system cost considered in several electricity markets results from disregarding the effect of the transmission network on the generation economic dispatch. Both short and long term marginal costs can be

[2] Nodal prices are also called locational marginal prices. In the pioneering work on this subject, see (Schweppe et al. 1988), the most general term "spot prices" is used.

[3] In some systems, like UK, energy and capacity payments associated to the production of reactive power have been paid to agents located in specific areas of the system where voltage problems may occur. However, no systematic nodal or zonal reactive power pricing scheme has been applied.

computed at system level and for each node. Long term marginal costs consider the option to marginally increase transmission or generation capacity to meet an increment of the load at a certain node.

2.3.2 Computation of Nodal Prices

Nodal prices can be readily obtained as by-products of the models widely available to compute the economic dispatch in the short-term taking into account the transmission grid. Models used may be as complex as needed. Using a very simple model, we aim to illustrate the process of computation of nodal energy prices within a centralized economic dispatch where network effects are considered.

In model (2.1) of the system economic dispatch, we make use of linear equations representing the flow of power over the grid according to Kirchhoff laws[4] (DC model). For the sake of simplicity, ohmic losses in each line have been represented as a function of the flow over this line and assigned to the extreme nodes of the line, thus being equivalent to an extra demand in each of the two nodes (half of the losses would be assigned to each node). For other representations of line losses, see, for instance, (Rivier et al. 1990). Besides, in order to make the formulation simpler, the only grid constraints considered are maximum line capacities.

$$
\begin{aligned}
&\max \sum\nolimits_{\mathrm{i}} \{B_{\mathrm{i}}(d_{\mathrm{i}}) - C_{\mathrm{i}}(g_{\mathrm{i}})\} \\
&s.t. \\
&d_i - g_i + \sum\nolimits_m \left\{\tau_{im} \cdot \phi_{im} - L_{i,m}(\phi_{im},\ R_{im})\right\} = 0 && \forall i && \pi_i \\
&\tau_{im} \frac{\theta_i - \theta_m}{x_{im}} = \phi_{im} && \forall i, m && \xi_{im} \\
&\phi_{im} \le \overline{\phi_{im}} && \forall i, m && \mu_{im} \\
&\theta_{ref} = 0 \\
&g_i \le \overline{g_i} && \forall i && \beta_i \\
&d_i \le \overline{d_i} && \forall i && \alpha_i
\end{aligned}
\tag{2.1}
$$

[4] Kirchhoff laws are two. First one states that at each node, power injections must equal power withdrawals. Second one states that, when flowing among two nodes, power is split among the different parallel paths between these nodes in inverse proportion to the electrical distances along these paths.

In the formulation in (2.1), i is an index representing the set of nodes; m is an alias of i; $B_{\mathrm{i}}(d_{\mathrm{i}})$ is either the benefit obtained by agents at node i from the power d_{i} they consume or the offer by agents at node for the power d_{i} they consume. Then, $B_{\mathrm{i}}(d_{\mathrm{i}})$ is equal to the cost of electricity for consumers plus the consumer surplus; $C_{\mathrm{i}}(g_{\mathrm{i}})$ is the either the cost incurred by agents at node i when producing g_{i} units of power or the bid by agents at node i to produce g_{i} units of power; $\tau_{\mathrm{i}m}$ is a binary variable whose value is 1 when nodes i and m are connected through a line and is 0 otherwise; $\phi_{\mathrm{i}m}$ is the flow over the line between nodes i and m in the direction from i to m; $L_{\mathrm{i},m}(\phi_{\mathrm{i}m}, R_{\mathrm{i},m})$ is the fraction (half) of transmission losses in the line between nodes *i* and *m* that has been assigned to node i and has therefore been represented as an extra load in this node: losses over a line depend on the flow and resistance of the line; θ_{i} is the phase angle at node i; $x_{\mathrm{i}m}$ is the reactance of the line between nodes i and m; $\overline{\phi_{\mathrm{i}m}}$ is the maximum flow allowed over the line from i to m in the relevant direction; θ_{ref} is the reference phase angle; $\overline{g_{\mathrm{I}}}$ is the maximum power production in node i and $\overline{d_{\mathrm{I}}}$ is the maximum demand for power in the same node. Apart from that, π_{i}, $\xi_{\mathrm{i}m}$, $\mu_{\mathrm{i}m}$, β_{i} and α_{i} are the dual variables of the corresponding constraints, which are obtained, together with the primal variables, when solving the optimization problem.

The nodal price at node k, ρ_{k}, is, in this simple case, the dual variable of the k$^{\mathrm{th}}$ nodal balance equation, π_{k}.[5]

$$\rho_{\mathrm{k}} = \pi_{\mathrm{k}} \tag{2.2}$$

The system economic dispatch can be modeled using an alternative formulation, see (2.3). In model (2.3), ohmic losses in the transmission grid are easily represented.

$$\max \sum\nolimits_{\mathrm{i}} \{B_{\mathrm{i}}(d_{\mathrm{i}}) - C_{\mathrm{i}}(g_{\mathrm{i}})\}$$

$$s.t.$$

$$\begin{array}{lll}
\sum\nolimits_{\mathrm{i}} (d_{\mathrm{i}} - g_{\mathrm{i}}) + L(d,\ g) = 0 & & \gamma \\
\sum\nolimits_{\mathrm{i}} \left\{PTDF_{\mathrm{i},l} \cdot (g_{\mathrm{i}} - d_{\mathrm{i}})\right\} = \phi_l & \forall l & \xi_l \\
\phi_l \le \overline{\phi_l} & \forall l & \mu_l \\
g_{\mathrm{i}} \le \overline{g_{\mathrm{I}}} & \forall i & \beta_{\mathrm{i}} \\
d_{\mathrm{i}} \le \overline{d_{\mathrm{I}}} & \forall i & \alpha_{\mathrm{i}}
\end{array} \tag{2.3}$$

[5] Strictly speaking, the nodal price expression will be $\rho_{\mathrm{k}} = \pi_{\mathrm{k}} + \alpha_{\mathrm{k}}$, although α_{k} will be non-zero only at those nodes where all the demand is fully unserved.

Most symbols in model (2.3) have already been used in this section. New ones are described next. $L(d, g)$ represents transmission losses in the system expressed as a function of power injections and withdrawals; $PTDF_{i,l}$, is the Power Transfer Distribution Factor of the flow over line l with respect to the power injection at node i, i.e. it is the sensitivity of the flow over this line with respect to the power injected at this node; ϕ_{l} is the flow over line l and $\overline{\phi_{\mathrm{l}}}$ represents the maximum amount of power allowed over line l in the direction in which the flow actually goes in the scenario considered. Finally, when used as an index, l refers to the set of lines in the system.

Nodal energy prices can be computed from the solution of the economic dispatch in (2.3) as a linear combination of several of the dual variables of constraints in this problem.

$$\rho_k = \gamma + \eta_k = \gamma + \gamma \cdot \frac{\delta L}{\delta d_k} - \sum\nolimits_l (\mu_l \cdot PTDF_{k,l}) \tag{2.4}$$

where $\frac{\delta L}{\delta d_k}$ is the loss factor corresponding to node k; and η_k is a variable representing the difference between the energy system price and the nodal price at each node. Then, generally speaking, this variable should be different from zero for all nodes but that taken as a reference for computing the system price. It reflects the impact of the grid on the value of energy at each node and depends on the reference node chosen.

The formulation of the optimization problem in (2.3) should depend on the identity of the node chosen as a reference nodes. The derivative of the systems losses with respect to demand at node k and PTDFs depend on the choice of reference node. The amount of power being dispatched at each node; the overall value of accepted bids; and nodal prices are not affected by the choice of the reference node but dual variable γ is. Therefore, (2.4) should instead read as in (2.5).

$$\begin{aligned}\rho_k = \gamma_s + \eta_{k,s} = \gamma_s + \gamma_s \cdot LF_{k,s} - \sum\nolimits_l (\mu_l \cdot PTDF_{k,l,s}) = \rho_s \cdot (1 + LF_{k,s}) \\ - \sum\nolimits_l (\mu_l \cdot PTDF_{k,l,s})\end{aligned} \tag{2.5}$$

2.4 Main Properties of Nodal prices

Nodal electricity prices consider the impact of the transmission network on the short term marginal value of energy both from a technical and an economic point of view. The level of these prices depends, at any time and node of the network, on system operation conditions including the following: set of available generation and transmission facilities and their technical features (capacities, line impedances); load level at each node; and variable costs of generators. Amongst nodal electricity prices main properties, there are their ability to send efficient short term signals; the efficient allocation among parties of the cost of losses and network constraints; their ability to recover only part of the cost of the grid; and the option to decompose them into a system (or energy) a loss and a congestion part. The remainder of this subsection discusses these properties in detail.

2.4.1 Property 1: Efficient Short-Term Energy Prices

It can be easily demonstrated that nodal prices are optimal short term economic signals that internalise all the grid effects in a single value – the price to buy or sell energy in €/$/£ per kWh – separately computed for each node. In other words, when the energy produced or consumed at a certain node i is priced at the corresponding nodal price π_i, market agents located at this node are encouraged to behave most efficiently in order to maximize the social benefit of the system. The proof of this statement can be found in (Schweppe et al. 1988).

Consumers decisions will only be optimal if they exhibit some elasticity to the price of energy (the larger the amount of power purchased, the smaller value they place on, or the price they are willing to pay for, an extra unit of power). However, most consumers do not decide how much power to purchase at any given moment in time based on the price they will have to pay for it. Hence, the amount of power retrieved by consumers at the majority of nodes can be considered to be an input to the dispatch problem where prices are determined. Unless the dispatcher has access to the true utility function of consumers (B_i for consumer at node i), nodal prices will not maximize social welfare in the short term.

Besides, achieving an optimal operation of the system requires bids from generators corresponding to their true production cost function (we have worked under the hypothesis that the cost function $C_i(g_i)$ used in the dispatch and the agent problems is the same). However, generators may bid strategically deviating from their true cost function since, in reality, some degree of market power always exists.

2.4.2 Property 2: Efficient Allocation of Network Losses Associated Costs and Redispatch Costs due to Network Constraints

Nodal prices result in those network users located in areas where the power they produce or consume cause significant losses or network congestion facing less favorable prices (higher for consumers, lower for producers) than those network users that, due to their location in the grid, contribute to reducing network losses or alleviate congestion in the grid. Therefore, besides producing optimal short-term signals, nodal prices are locational signals encouraging agents to install new load or generation in places where the resulting ohmic losses and network congestion are as small as possible.

Note however that, while nodal prices send economic signals in the direction of reducing losses and congestion costs, they are not assigning to agents the social cost of losses and network congestion. This is remarkably clear for losses, since, due to the fact that losses increase with the square of power flows, nodal price differences due to losses result in larger net revenues for the system than the cost of system losses.

2.4.3 Property 3: Contribution to the Recovery of Network Investment and Maintenance Costs

The application of nodal prices to the power injected and withdrawn in each node gives rise to a net revenue NR_t at each time t, whose expression is provided in (2.6). The overall net revenue for the whole system over a certain period of time, normally a year, is widely known as the Variable Transmission Revenues (*VTR*) of the system, whose mathematical expression is provided in (2.7):

$$NR_t = \sum\nolimits_n \left(\pi_{n,t} \cdot d_{n,t} - \pi_{n,t} \cdot g_{n,t}\right) \tag{2.6}$$

$$VTR = \sum\nolimits_t NR_t \tag{2.7}$$

where n represents the set of nodes in the system and t the time.

As shown in (Olmos 2006), *VTR* can also be computed line by line according to (2.8). Each line l between nodes *in* and *out*, where power flows from node *in* to node *out*, can be considered an arbitrageur buying energy $P_{l,in,t}$ injected in the line at node *in* and time t and selling energy $P_{l,out,t}$ retrieved from the line at node *out* and the same time. Given that the amounts of power injected into and withdrawn from line l differ by the amount of ohmic losses in this line, and nodal prices at time t at both line nodes $\pi_{in,t}$ and $\pi_{out,t}$ also differ, the commercial exploitation of line l will result in a net revenue at time t represented in (2.8) by $NR_{l,t}$.[6]

$$VTR = \sum\nolimits_{l,t} NR_{l,t} = \sum\nolimits_t \sum\nolimits_l \left(\pi_{out,t} \cdot P_{l,out,t} - \pi_{in,t} \cdot P_{l,in,t}\right) \tag{2.8}$$

Variable Transmission Revenues computed according to (2.7) and (2.8) are the same. These revenues are associated with differences among nodal prices and powers injected into and withdrawn from the grid due to transmission losses and congestion. Network revenues associated with congestion are also known as congestion rents.

[6] Exceptionally, "**network revenues**" may be negative when line losses are very large due to corona discharge. Note that network revenue is the profit that the transmission network would earn if energy were purchased from generators at their **nodal price** and sold to consumers at theirs. However, the transmission network should not be allowed to conduct free market transactions, but must rather be treated like a regulated monopoly with pre-established remuneration. Exceptions, namely merchant lines, may be justified for individual lines under special circumstances.

VTR critically depend on the level of development of the grid. Overdeveloped grids will result in small losses and congestion, thus leading to small differences among nodal prices. These, in turn, will result in small *VTR*. On the other hand, underdeveloped grids will result in large differences among nodal prices probably leading to large *VTR* (although losses will probably be large as well).

Pérez-Arriaga et al. (1995) demonstrate that, under ideal conditions affecting the planning of the grid, *VTR* in an optimally developed network would amount to exactly 100 % of the network investment costs. Ideal conditions affecting the development of the network to be met for network variable revenues to amount to 100 % of the network costs are investigated in Rubio and Pérez-Arriaga (2000), and mainly include the following:

- Static and dynamic network expansion plans are the same and planning errors do not occur.
- Investments in transmission are continuous.
- Economies of scale do not exist in the transmission activity.
- Reliability constraints considered in system development planning are also considered in system operation.

In real life systems, *VTR* fall short of total transmission costs. The former only manage to recover about 20 % of the costs of the grid, according to estimates in Pérez-Arriaga et al. 1995. Main reasons for revenues from the application of nodal prices being so low are briefly discussed next.

First, economies of scale and the discrete nature of network investments result in an overdevelopment of networks in practice (see Dismukes et al. 1998). In effect, building lines with a large capacity is generally preferable over building a larger number of small lines even when the former are not going to be fully used during the first years of their economic life. As we have just explained, overinvesting in the development of the grid results in small nodal price differences and a small *VTR*.

Second, certain reliability constraints and a wide range of scenarios shall be considered when planning the expansion of the grid, This is due to the fact that there is a high level of uncertainty about the operation conditions that may occur in the system throughout the economic life of investments being decided. However, some of these restrictions and all these scenarios but one will not be considered when computing the operation of the system. Due to the fact that the set of constraints considered when computing the expansion of the system tends to be larger than that considered for operation planning, long term nodal prices computed assuming grid investments are continuous would also differ substantially from short term nodal prices. Specifically, differences among long term nodal prices, and therefore also revenues from their application (which should amount to the exact cost of the grid assuming continuous investments), would be much larger than those computed for short term prices.

Therefore, even if (short term) nodal prices are applied, revenues from their application will not suffice to recover the cost of the grid. Additional transmission charges will need to be levied on network users to complete the recovery of this cost. This is discussed in Sect. 2.6.

Given that Financial Transmission Rights entitle owners to receive the difference between the energy prices at the nodes that these rights refer to, the aggregate value for market agents of all the simultaneously feasible transmission rights (defined as obligations) that can be issued in the system would equal the expected overall net revenues from the application of energy prices. Due to the fact that, as just mentioned, these revenues tend to be much smaller than the total cost of an optimally developed transmission grid, it is highly unlikely that the financing of investments in the transmission grid through the issuance of FTRs would result in an appropriate development of the grid. Most of the required reinforcements could not be financed through this scheme.

Authorities must bear in mind that revenues of transmission companies or the System Operator should generally not depend on revenues resulting from the application of nodal prices. Otherwise, they will have a perverse incentive not to invest in the further development and maintenance of the grid so as to increase nodal price differences and therefore their revenues. Revenues of transmission service providers should generally be regulated (not dependent on nodal price revenues), though *VTR* should probably be devoted to finance part of the payments to these companies.

2.4.4 Property 4: Decomposition of Nodal Prices in Their Energy, Losses and Congestion Components

As already pointed out when discussing the computation of nodal prices in Sect. 2.3.2, the nodal energy price in each node can be decomposed into three components: one associated with the marginal cost of producing electricity in the system; another one associated with the effect that increasing the demand in this node has on ohmic losses and the marginal cost of electricity; and a third one related to the effect of marginally increasing the demand in the node on transmission constraints and the cost of these constraints. The decomposition of nodal electricity prices is investigated by Rivier and Pérez-Arriaga (1993), where the mathematical expression of the nodal price in node *k* provided in (2.9) is derived.

$$\rho_k = \gamma_s + \eta_{k,s} = \gamma_s + \gamma_s \cdot LF_{k,s} + \sum\nolimits_j \mu_j \cdot NC_{j,ks} \tag{2.9}$$

where, γ_s can be deemed the cost of producing electricity in the system, which is common to all nodes whose prices are to be computed; and $\eta_{k,s}$ is the part that can be deemed specific to each node k, which comprises the cost of losses caused by an increase in the node demand, $\gamma_s \cdot LF_{k,s}$, and the cost of restrictions affected by this demand increase, $\sum_j \mu_j \cdot NC_{j,ks}$. $LF_{k,s}$ is the loss factor of node k; μ_j is the cost of each restriction j; and $NC_{j,ks}$ is the impact of an increase in demand in node k on the system variable constrained in restriction j. $LF_{k,s}$ and $NC_{j,ks}$ are therefore sensitivity factors measuring changes of losses and any constrained parameter of the system, respectively, for an increase in demand at node k.

However, defining a one-to-one relationship between each nodal price ρ and its energy, losses and constraint components is not possible. As highlighted in (2.9), components of the price at a node k must be defined taking as a reference the nodal price at another node s, which we shall call reference node from now on. Thus, the energy component of price ρ_k, γ_s, corresponds to the nodal price at node s; the losses component is defined in terms of $LF_{k,s}$, which is the loss factor at node k taking as a reference node s, meaning the increase in the ohmic losses in the system resulting from an increase in power injected in node s to supply a marginal increase in electricity demand in node k; finally, the constraint component of price ρ_k is defined in terms of the impact $NC_{j,ks}$ on the system variable constrained by grid constrain j of an increase in power injected in node s to supply a marginal increase in electricity demand in node k. Changing node s taken as a reference for the computation of nodal price ρ_k would result in a change of the value of its energy, losses and constraint components, while the nodal price itself would not change.

The reference node s may be chosen to be the one(-s) where the marginal generator(-s) in the system economic dispatch is(are) connected. Then, the energy component of the nodal price at node k would refer to the cost of producing electricity with the most efficient generation unit(-s) available, while the losses and constraint components would correspond to the cost for the system of transporting electricity produced by the marginal generator(-s) in the dispatch to node k. This, in any case, must be deemed an arbitrary decomposition of nodal price ρ_k, since the system marginal generator may change depending on the set of active constraints and existing losses, and therefore the production cost of this generator cannot be deemed independent of constraints and losses in the system. Therefore, decomposing nodal prices into its energy, losses and constraint components may have practical applications but one should be aware of the limitations of such a composition.

An interesting corollary of the decomposition of nodal prices just discussed is the existing relationship between the prices in any two nodes k1 and k2 in the system, which is provided in (2.10).

$$\rho_{k1} = \rho_{k2} \cdot (1 + LF_{k1,k2}) + \sum_j \mu_j \cdot NC_{j,k1,k2} \tag{2.10}$$

Equation (2.10) results from deriving the expression of nodal price ρ_{k1} according to (2.9) when taking node k2 as the reference one. Rivier and Pérez-Arriaga (1993), discuss other less-relevant properties of nodal prices. Other algorithms have been proposed more recently to overcome the dependence of the decomposition of prices on the reference bus chosen (see Cheng and Overbye 2006). This and other research works try to get around this challenge by imposing constraints on the decomposition problem that determine the identity of the reference bus.

2.4.5 Dependence of the Sensitivity of Line Flows with Respect to Nodal Power Injections on the Choice of Reference Node

Factor $PTDF_{k,l}$ refers to the sensitivity with respect to the power injection in node k of a specific type of constrained variable: the flow over line l.[7] Sensitivity factors of line flows are commonly used in regulatory approaches normally related to the allocation of the costs of transmission lines. Factor $PTDF_{k,l,s}$ is commonly claimed to represent the marginal use of line l by agents located in node k.

As already mentioned, the value of PTDFs depends on the reference node considered when computing them. Then, the sensitivity of the flow in line l with respect to the power injection in node k must be denoted $PTDF_{k,l,s}$, thus refering to the specific node s where an increase in the power withdrawn balances the aforementioned increase in the power injected in node k (neglecting losses, the extra power withdrawn in node s must be the same as that injected in node k[8]).

Given the role that PTDFs have in the allocation of the cost of transmission lines according to some of the methods proposed for this (namely the so-called Marginal Participations method), discussing the effect of the selection of the reference node on the value of these factors is relevant. If losses are neglected and line flows are assumed to be a linear function of power injections and withdrawals, applying the superposition principle it can be easily proved that the PTDFs of line l with respect to the power injected at node k computed using reference nodes $s1$ and $s2$ are related by the expression in (2.11).

$$PTDF_{k,l,s2} = PTDF_{k,l,s1} + PTDF_{s1,l,s2} \tag{2.11}$$

Note that $PTDF_{s1,l,s2}$ does not depend on the reference node chosen. This involves that changing the slack node results in a uniform increase (either positive or negative) of the sensitivities of the flow in each line with respect to the power injected in all nodes of the system. Therefore, absolute differences among the sensitivities of the flow in a line with respect to power injections in different nodes of the system do not depend on the reference node used to compute these sensitivities.

[7] Power Transfer Distribution Factors are normally defined as the sensitivities of flows with respect to power injections, while sensitivity factors of constrained variables in general, NC, are normally defined with respect to power withdrawals. Therefore, changing the sign of factors NC corresponding to line flows is necessary to compute PTDFs. Besides, it must be noted that PTDFs are defined by some authors as the sensitivity of line flows with respect to point to point transactions rather than power injections. Thus, for example, authors in Galiana et al. (2003) compute the sensitivity of line flows with respect to equivalent bilateral power exchanges (whereby each demand is assigned a fraction of each generation and each generator is assigned a fraction of each demand in a uniform manner) to allocate the cost of these lines to their users.

[8] If losses are considered, the amount of power withdrawn in the reference node should not be 1 MW (a unit increase) but an amount slightly larger or smaller depending on the effect on transmission losses in the system of the considered power transaction between node k and reference node s.

Then, if part of the cost of transmission lines is allocated to agents according to the sensitivities of the flows in the former with respect to power injections by the latter, differences among the transmission charges to be paid by different agents would not depend on the reference node chosen to compute line flow sensitivities. However, this does not mean that charges computed using any reference node make engineering and economic sense. As explained in Olmos and Pérez-Arriaga (2009), only those cost allocation methods whose underlying principles are sound can be deemed to produce sound transmission charges.

If transmission losses are taken into account, the choice of the reference node has a small, albeit nonzero, influence on differences among the sensitivities of a line flow with respect to power injections in different nodes, as shown in (2.12) and (2.13), which have been derived from the discussion on the subject in Rivier and Pérez-Arriaga (1993):

$$(1 + LF_{k,s2}) = (1 + LF_{k,s1}) \cdot (1 + LF_{s1,s2}) \tag{2.12}$$

$$PTDF_{k,l,s2} = PTDF_{k,l,s1} + PTDF_{s1,l,s2} \cdot (1 + LF_{k,s1}) \tag{2.13}$$

Differences among line flow sensitivities with respect to different injection nodes are dependent on the choice of the reference node because the change in the sensitivity factor for a certain injection node resulting from a change of the reference node is a function of the loss factor of this injection node. However, differences among loss factors computed for different injection nodes are likely to be very small. Hence, generally speaking, differences among line flow sensitivity factors can be deemed slack node independent.

2.5 Main Locational Energy Pricing Schemes: Alternatives to Nodal Pricing

The management and pricing of the effect that the transmission network has on the energy dispatch is one of the areas where the power system academic community has been more prolific recently (see Chao and Peck 1996; Stoft 1998; Ruff 1999; Chao and Peck 2000; Tabors and Caramanis 2000; Boucher and Smeers 2001; ETSO 2001; Henney 2002; Hogan 2002; O'Neill et al. 2002; ETSO 2004; ETSO/EuroPEX 2004 as a sample of relevant works on the subject). The choice of the transmission pricing scheme to be applied should condition the definition of Financial Transmission Rights, as we shall explain below for each of the main types of schemes. Any transmission pricing scheme to be implemented must comply with sound engineering and economic principles but it must also be politically acceptable. This section describes and critically analyses the most relevant options for the pricing of the effects of transmission on power system dispatch. We discuss only market based methods, i.e. those which aim to maximize the economic value of energy and transmission capacity bids accepted.

Pricing schemes can be classified according to different criteria. Probably the two most relevant ones are (1) the type of interface involved in these schemes between energy and transmission pricing and (2) the level of locational differentiation (granularity) in final energy prices that result from them. According to the first criterion, pricing schemes can be classified into implicit schemes, where energy prices computed include the effect of the transmission grid on the economic value of energy, and explicit ones, where the effect of the network on the value of energy at each location is priced separately from energy itself. According to the second criterion, one may distinguish among nodal pricing, where a separate energy price is computed for each transmission node; zonal pricing, whereby the system is divided into pricing areas and a separate price is computed for each of them; and single pricing, where a single energy price is applied at all nodes in the system.

We shall here review main pricing alternatives according to the location differentiation in final energy prices they create. Within each main option corresponding to a level of disaggregation of prices, a distinction may be made between implicit and explicit schemes if appropriate.

2.5.1 Nodal Differentiation of Energy Prices

By far, the most relevant scheme within this category is nodal energy pricing (also called Locational Marginal Pricing), which produces a separate price for the energy consumed and generated at each transmission grid node. Energy prices computed through nodal pricing implicitly include the effect of grid losses and transmission congestions, internalising both effects in a single value (€/$/£ per kWh) that varies at each system node. Therefore, nodal pricing is an implicit transmission pricing scheme that produces perfectly efficient signals for decisions concerning the (short-term) economic operation of generation and demand, since nodal prices correctly convey the economic impact of losses and constraints at the different producer and consumer locations.

When focusing on the effect of grid congestion on the dispatch, nodal pricing may be seen as the outcome of a joint competitive auction of energy and physical rights to use the transmission capacity. O'Neill et al. (2002), provide an example of implementation of a contingency constrained auction for both energy and transmission rights where the authors consider both options and obligations. Auctions proposed in O'Neill et al. (2002), are different from other designs of implicit auctions in the sense that authors propose using them both in the short and the long term.

The academic community has come up with several designs to run implicit auctions in a decentralized manner. Thus, Aguado et al. (2004), decomposes the original problem into several simpler ones. The optimal outcome at regional level is found through an iterative process. The concept, properties and way to compute nodal energy prices have already been extensively discussed in the preceding sections within this chapter.

Instead of integrating the effect of transmission on the energy dispatch, one may think of separately pricing the effects that network congestion or losses should have on the final price of energy. However, if we are not able to define areas of uniform energy prices, which result from the application of a zonal, instead of a nodal, pricing scheme, separating the allocation of energy and capacity is not possible (or feasible from a practical point of view). When zonal prices cannot be defined, any power transaction significantly affects the flow through the congested lines and has to participate in the transmission capacity allocation process. Then, the unconstrained energy dispatch taking place after the allocation of transmission capacity (where limits to power flows imposed by the network are not considered) has to replicate exactly the outcome of the capacity allocation process (either the capacity auction or the outcome of the bilateral trading process taking place among agents to buy and sell transmission capacity rights).

However, the effect of transmission losses on efficient energy prices can effectively be computed separately from the energy system price (the so called lambda in nodal pricing nomenclature) through the application of loss factors. Therefore, there is no need to forgo the short-term loss signals that contribute to the economically efficient system operation. The losses attributable to each player, either computed as a marginal or average value, can be applied in the form of corrective factors to determine the prices to be paid or earned by this player or, rather preferably, the net amount of energy produced or consumed by the former. This should lead players to internalise the losses they are responsible for in their offers.

When energy prices differ by node, Financial Transmission Rights can be used to hedge against possible financial losses from the volatility in the differences among prices at two or more nodes (ETSO 2006). FTRs hedging a certain power transaction may be issued by any party. However, leaving their issuance in the hands of the TSO responsible for transmission among the nodes in the transaction would ensure revenue adequacy (Hogan 1992). According to this criterion, the issuing party should in this case be the corresponding national or State TSO for local transactions and the regional TSO for cross-border transactions.

Examples of **nodal pricing** can be found in electricity markets in Chile, Argentina, New Zealand and several Regional Transmission Organizations (RTOs) in the USA, such as the PJM system (Pennsylvania, New Jersey, Maryland), the Electric Reliability Council of Texas (ERCOT) system, or the California system. Loss factors are used for instance in the Irish Single Electricity Market.

Revenues from the application of nodal prices correspond to the economic value produced by the transmission grid by transporting power from nodes where it has a lower value (price) to those where its value is higher. Then, these revenues should be devoted to pay the regulated revenues to be earned by grid owner(-s).

2.5.2 Zonal Differentiation of Prices

Zonal price differentiation schemes involve applying the same final energy price within each of a set of areas while allowing price differences to take place among these areas. Normally, under zonal price schemes, a single market price is applied to all agents in the system unless significant network congestion occurs restricting the energy flows among pre-defined areas. In the latter case, prices differ among areas but the same price is applied to all nodes within any of these areas. Therefore, zonal price differences are normally caused by grid congestion, though a system of zonal loss factors is applied in some systems.

Energy price differences among electrical zones can result from the application of both implicit and explicit schemes. Zonal type implicit pricing schemes are normally referred to as zonal pricing or market splitting. Explicit mechanisms normally take the form of a coordinated auction of the capacity of the corridors linking price zones.

Zonal pricing normally involves the computation of a single, centralized, energy dispatch in the whole national or regional system where network effects within each uniform price area are neglected. It is therefore a simplification of nodal pricing. Market splitting, which can be considered a particular case of zonal pricing, involves the consideration of only one offer curve and one demand curve for the whole system in a first step. If the resulting pattern of flows causes significant congestion on the corridors linking the predefined areas, separate offer and demand curves are considered for each price area and, according to these curves, power is transacted among areas so that existing congestion is solved. This implementation of market splitting agrees with that of many others in the academic literature and the industry (see ETSO 1999; Newbery et al. 2003). Market splitting is applied within the Nordel region and in Italy. Zonal pricing has been also used in California.

Alternatively, the network capacity of likely-to-be-congested corridors linking uniform price areas may be explicitly allocated prior to running an only-energy market within each area. Market agents must acquire the right to use the inter-area transmission capacity they need to carry out the commercial transactions they want to get involved in, i.e. physical transmission rights over this capacity. Agents may buy this capacity (the right to use it) in a centralized explicit auction where the right to use the transmission network is allocated to those agents who value it most. Alternatively, agents may negotiate bilaterally the acquisition of those rights previously issued by the corresponding TSO.

Chao and Peck were the first ones to propose the utilization of rights over the capacity of likely-to-be-congested flow-gates (corridors) (see Chao and Peck 1996), where authors demonstrate that, under ideal conditions, this system would converge towards efficient energy prices. Similarly, Oren and Ross 2002, propose in an auction for flow-gate rights prior to the energy dispatch. Authors propose a system whereby SOs responsible for the energy dispatch in the different control areas would coordinate to manage the flow on the congested lines that is the responsibility of transactions taking place within different areas. There are other works on the use of flow-gate rights in combination with unconstrained energy markets (see Tabors and Caramanis 2000, for an example).

Once transmission capacity rights have been assigned in one way or the other, the energy auction takes place. Only those transactions that have acquired capacity rights to access the congested transmission they use can participate in the energy market. Auctioning transmission capacity at regional level requires some centralized coordination (see ETSO 2001). If flow patterns due to the different transactions were not considered jointly they might result in unexpected violations of network constraints unless significant security margins were applied. But employing security margins would most likely result in an underutilization of the transmission grid.

In those systems where explicit auctions are used, local authorities are in charge of the dispatch of energy within their corresponding areas. Thus, areas or countries enjoy a high level of independence. For this reason, capacity auctions have been widely applied in real life power systems. Up till recently, this was the method used to manage congestion on the borders between Austria and the Czech Republic, Belgium and the Netherlands, Denmark and Germany or France and the United Kingdom, among others (see Consentec/Frontier 2004).

The implementation of both zonal pricing schemes and mechanisms for the explicit allocation of transmission capacity on congested corridors implies the definition of internally well-meshed areas which can be considered as super nodes for congestion management purposes. Nodal energy prices computed within any of these predefined areas should be very similar if losses are ignored and serious congestion is limited to the interconnections between areas. Then, these areas can be regarded as “Single Price Areas” (SPAs) as far as congestion management is concerned (Christie and Wangensteen 1998; Stoft 1998; Chao and Peck 2000).

Balanced transactions within a SPA should not significantly affect the flow over inter-area links. In other words, any bilateral transaction within a SPA should not create loop flows outside this area which may significantly contribute to congestions inter Single Price Areas. The definition of Single Price Areas, whenever applicable, is not a trivial matter without practical consequences, see (Boucher and Smeers 2001). In zonal pricing schemes it will affect the validity of the energy dispatch and energy prices computed. What is more, as explained when discussing nodal pricing schemes, if we were not able to define SPAs, separating the allocation of energy and capacity, and therefore applying explicit capacity pricing mechanisms, would not be possible.

Borders among Single Price Areas may probably not coincide with political ones. Thus, assuming SPAs that are the same as existing control areas or countries may result in an inefficient dispatch or, even worse, in one that is far from being feasible. Thus, the plans to implement an implicit auction among Power Exchanges in Europe should be reconsidered carefully (see ETSO/EuroPEX 2004).

Financial Transmission Rights to be defined in this case should refer to two or more of the pricing zones whose definition has just been discussed, as price differences to hedge within each of these zones would be zero.

Revenues from the application of pricing schemes with zonal differentiation should be devoted to the coverage of network allowed regulated revenues, as with nodal prices, since they are just a simplified version of the nodal pricing scheme.

2.5.3 Single Pricing

In densely meshed transmission grids with no systematic or structural congestions, applying pricing mechanisms introducing nodal or zonal energy price differentiation is often regarded to be an unnecessary sophistication. Then, a single energy price is computed for the whole system. Once the outcome of the only energy market is known, one can check whether the pattern of commercial transactions violates any network constraint. Only when a constraint is violated does the System Operator need to re-dispatch some generation. Several implementations of re-dispatch are possible. According to some of them, the cost of the re-dispatch carried out to solve any violation of the network constraints should be minimum (see Rau 2000; Tao and Gross 2002). In these cases, market-based mechanisms must be used to modify the pattern of generation in the system. In other words, changes to the dispatch must be based on the bids sent by market agents indicating how much they ask for in order to change their market positions. Other re-dispatch algorithms aim to minimize the number and size of the adjustments to the original dispatch (see Galiana and Ilic 1998; Alomoush and Shahidehpour 2000). Fang and David 1999, describe other possible schemes.

Alomoush and Shahidehpour (2000) and Biskas and Bakirtzis (2002), are aimed at re-dispatching generation and load in the context of regional markets. These algorithms must achieve coordination among the different zones. Thus, Biskas and Bakirtzis (2002), decomposes the original problem using Lagrangian relaxation techniques. The coordination variables are the prices of the power exchanges between zones.

Counter-trading is a specific implementation of the method of re-dispatch. In counter-trading, the System Operator nominates pairs of generators that modify their outputs to create a power flow that goes in the opposite direction to the one causing network congestion in the unconstrained energy dispatch. Obviously, one could generalize and say that re-dispatch is nothing but counter-trade, since any increase in the output of a generator has to be matched by a corresponding and identical (except for losses) reduction in the output of another generator.

Typically, the extra cost of re-dispatch or counter-trade is socialized to all consumers thus leading to uniform energy prices in the whole system (single pricing). In this case, any economic signals resulting from the management of congestion, which could have been used to emulate nodal or zonal pricing, are lost. Conceptually speaking, assigning the cost of re-dispatch to those market agents that “create” the network constraint is possible. Economic signals would thus not be completely lost. This is a technically complex task, nevertheless Rivier and Pérez-Arriaga (1993), and others. Tao and Gross (2002), allocate the cost of re-dispatch taking into account the participations of agents (injections and withdrawals considered separately) in the flow over the congested lines. In order to do this, they express the flow over the congested lines as a function of power injections and withdrawals. Similarly, Baran et al. (2000), determines the participation of each transaction in the flow over congested lines. Afterwards, the total cost of re-dispatch is allocated among congested lines taking into account both the marginal cost of the restriction on the flow through each congested line and the incremental cost of the re-dispatch necessary to avoid violating this restriction.

Experience with counter-trade in California shows that those schemes based on re-dispatch may be subject to gaming by market agents who artificially create congestion in the grid in order to be paid afterwards to remove it. In any case, nodal pricing or implicit auctions seem to be superior to congestion management mechanisms based on re-dispatch. Singh et al. (1998), compare nodal pricing to a mechanism based on decentralized bilateral trade among market agents, followed by the minimum cost re-dispatch necessary to solve infeasibilities. They conclude that price signals resulting from nodal pricing are more efficient, unless the cost of re-dispatch is efficiently allocated to the agents responsible for congestion in the grid. However, as we have explained before, efficiently allocating the cost of re-dispatch is not an easy task nor is there an indisputable way to do it.

However, nodal and zonal pricing schemes may also result in extra incentives to exercise market power when, due to the reduction in the size of the relevant market under these schemes in the presence of congestion, market agents gain power to unilaterally affect the energy price in any of the pricing zones the system is divided into. Auctioning Financial transmission Rights may aggravate this problem when market agents enjoying market power in an importing area are allowed to buy transmission rights into this area, see Olmos and Neuhoff 2006.

Applying a single energy pricing scheme does not result in net revenues (congestion rents) to be devoted to partially covering the cost of regulated transmission grid lines. On the other hand, as already mentioned, if redispatching generation and or load is necessary, a net cost will be incurred. Many national power systems apply single energy pricing schemes internally (within their borders). These include almost all European countries and Colombia.

Obviously, implementing single pricing within a system would render FTRs useless at local level, since there would not be energy price differences to hedge. Market agents would only need to be hedged against potential differences among single energy prices applied in different local (national, State) systems. For the most part, this is the case within the Internal Electricity Market of the European Union.

2.6 Completing the Recovery of the Network Cost: Computation of the Complementary Transmission Charges

2.6.1 Fundamentals

Electric power transmission is nearly regarded a natural monopoly. Therefore, transmission should be a regulated business. Both under traditional cost of service regulation and under incentive regulation, the allowed annual revenues of the regulated transmission company, which are set by the regulator, must be paid by transmission network users. We discuss here how network related economic signals should be designed to achieve the recovery of the allowed transmission revenues while promoting efficiency in the short-term (encouraging agents to make optimal operation decisions) and in the long-term (driving agents' decisions on the location of new generators and loads).

As already shown, energy prices applied may have different levels of spatial differentiation due to the existence of losses and constraints in the grid. Energy price differences among nodes give raise to location-related economic signals to network users and result in some partial recovery of the total allowed revenues of the regulated transmission company. As already explained, revenues from the application of nodal prices comprise those obtained well ahead of real time through the sale of Financial (or Physical) Transmission Rights over the capacity of likely to be congested corridors, or hedging differences in prices among different nodes, and those obtained in the day-ahead and real time markets through the application of these prices to power injections and withdrawals. However, as Rubio and Pérez-Arriaga (2000), show, the net revenue resulting from the application of nodal prices amounts only to a small fraction of the total regulated cost of the grid. Revenues resulting from the application of alternative energy pricing schemes are expected to be lower. The fraction of regulated transmission revenues recovered from the application of energy prices is normally referred to as Variable Transmission Revenues (VNR).

Therefore, completing the recovery of the cost of the grid requires applying additional charges, normally called complementary charges, that relate to the fraction of the grid cost not recovered through energy prices. Complementary charges should also send economic signals to agents encouraging them to reduce the cost of expansion of the grid. Therefore, these charges should encourage agents to install new generation or load in those locations where reinforcements needed for the grid to cope with the resulting incremental flows are least costly.

Additionally, complementary charges should be compatible with the application of efficient short-term economic signals. Complementary charges refer to all transmission business costs associated with network infrastructure including investment costs (asset depreciation as well as a return on net fixed assets), operation and maintenance costs, and other administrative and corporate costs. On the other hand, line losses and generation costs due to grid constraints, System Operator costs and those costs related to the provision of Ancillary Services should be levied on system users through other charges. Then, complementary charges are related to the allocation of long term costs not to be affected by short-term decisions by agents (the cost of lines already existing is not conditioned by how much power each generator or load is transacting at each time). As a consequence of this, complementary charges should interfere as little as possible with short-term economic signals, so as not to compromise the efficiency of system operation.

Transmission charges can be divided into Connection charges and Use of the System (UoS) charges. Connection charges are employed to allocate the cost of transmission facilities directly connecting a network user, or group of users, to the rest of the grid. UoS charges are related to the costs of the rest of transmission facilities. Economic principles advocate allocating the cost of each transmission line according to the responsibility of grid users on the construction of that line. Applying this principle is easy when it is about allocating the cost of connection facilities: those responsible for their construction are the users connecting through them to the rest of the system. On the other hand, determining the responsibility of generators and loads in the construction of the bulk of the transmission grid is much more difficult, especially when the grid is meshed. The remainder of this section is devoted to the discussion of the design of UoS charges. Both the allocation method employed to determine which fraction of the grid should be paid by each agent and the design of transmission charges are discussed next.

2.6.2 Allocation of the Cost of the Main Grid to Its Users

Determining those generators and loads that were responsible for the construction of some lines has proven to be a very difficult task. Then, it is most sensible to use some proxy of cost causality, such as the level of network utilization of each line by each agent, as the basic criterion for the allocation of the cost of this line. This involves assuming that the responsibility of each agent in the construction of a line is proportional to the amount of use of the line by the agent.

However, the cost of those expensive lines that only benefit a subset of network users, in non-well-meshed networks, should be allocated according to the responsibility of network users in the construction of the former. The fraction of the cost of each line that each network user is responsible for can be computed based on the a priori estimation of the benefits produced by this line for this user.

Unfortunately, computing the electrical utilization of lines by agents is not a simple task either, since there is no indisputable method to do it. Several methods to determine the use of the network by agents have been proposed and applied, with results that vary significantly from one another. It is important to keep in mind that the final objective is not computing the use of the network by each agent, but determining the responsibility of this agent in the construction of the line.

Transmission tariffs in most countries do not contain any locational signal. They disregard the need to allocate efficiently line costs (see for instance ETSO 2008; Lusztig et al. 2006). Regulators have settled for simple transmission charges that socialize the cost of the network to its users. However, in our view, as time passes and all kinds of new generation compete to enter into the system, sending clear locational signals – including transmission tariffs – will become more relevant.

2.6.2.1 Computing the Responsibility of Agents in Network Costs

Whenever computing the benefits that network users obtain from transmission lines is not possible, the responsibility of these users in network costs should be determined taking as a reference the best estimate possible of their use of the grid. Olmos and Pérez-Arriaga (2007) point out that methods to be used to compute the use of the grid by generators and loads shall be in agreement with the underlying technical and economic principles of the functioning of power systems. Even when there is no indisputable method to compute the utilization of lines by agents, some proposed in the literature, like the method of Average Participations (AP) described first in Bialek (1996) and Kirschen et al. (1997), or the Aumann-Shapley method, whose application for the computation of transmission tariffs is analyzed in Junqueira et al. (2007), seem to be sensible options.

Most usage based network cost allocation methods providing sensible results (like AP or Aumann-Shapley) aim at determining the “average” use of the grid by each generator or load as if the latter had always been in place. However, the responsibility of agents in network reinforcements is directly related to the incremental flows produced by the decisions of these agents to install new generators or loads in specific places. Hence, usage based cost allocation factors produced by methods like AP or Aummann-Shapley should be modified to take account of the different possible patterns of change of the flows in the system caused the installation of each generator or load and the time when this generators or load and the lines in the system were built. The application of these principles to the process of computation of transmission charges is discussed in detail in Olmos and Pérez-Arriaga (2009).

Olmos and Pérez-Arriaga (2009) point out, the loading rate of each transmission line and the desired split of total transmission costs between generation and load in the system should also condition the level of transmission tariffs (complementary charges) paid by each network user. The fraction of the total cost of a line to be allocated to agents according to their responsibility in the construction of the line should probably be limited to the ratio of the loading rate of the line to that of other similar lines in the system. The remainder of the cost of this line should probably be socialized, since current users of the grid cannot be deemed responsible for the construction of the fraction of the capacity of a line that is expected not to be used until long time in the future (for lines that are underutilized in the present).

As already mentioned, the split of total transmission charges between generation and load should probably take place according to the total benefits that generation on the one hand, and load, on the other, will obtain from the grid. However, given that estimating these benefits may turn out to be very difficult in most cases, a 50/50 split of costs between the two groups may be adopted unless system authorities have sound arguments to set a lower limit to the overall fraction of costs to be paid by generators (operation decisions by generators may be more sensitive to the level of transmission charges than those by loads).

2.6.3 Designing UoS Charges

Designing transmission charges involves not only developing the methodology for computing the responsibility of agents in the cost of the transmission grid, but also providing adequate answers to many implementation issues. We now focus on the most relevant aspects of the implementation of locational transmission grid charges that are not directly related to the cost allocation algorithm applied. These include computing the number of operation scenarios to be considered; defining the structure of charges and their updating procedure; and deciding the way to deal with grandfathering issues arising in the process of implementation of these charges.

As Olmos and Pérez-Arriaga (2009) point out, tariffs should be published based on the expected future operation of the system over a set of scenarios that are representative of the different set of situations that may exist in the future once the considered generator or load has entered into operation. The relative weight given to each scenario in the computation of the allocation of the cost of a line should be in accordance with the reasons justifying the construction of this line. The total cost of the line should be apportioned into two parts: one representative of the weight that the reduction of transmission losses had on the decision to build the line and another one representative of the weight of the decrease in congestion costs. Then, the relative weight given to each scenario in the process of allocation of the cost of the fraction of the line deemed to be built to reduce losses should be proportional to the system losses in this scenario. The relative weight given to each scenario in the process of allocation of the cost of the fraction of the line attributable to the reduction of congestion costs should be proportional to the level of congestion costs in this scenario, which, as a proxy, can be deemed proportional to the load level.

As aforementioned, operation decisions by network users, which are short-term decisions, should not be conditioned by the level of the transmission charge paid by these agents to recover the total network costs, which should be a long term signal. Short-term locational signals can be sent via nodal energy prices (locational marginal prices, LMP in the US terminology). If transmission tariffs are applied in the form of energy charges (€/MWh), i.e. a charge that depends on the amount of energy produced or consumed by the corresponding agent, network users will internalize these charges in their energy bids to the Power Exchange or in their bilateral contracts, therefore causing a distortion in the original market behaviour of these agents and the outcome of the wholesale market. It is then concluded that the transmission charge should have the format of a capacity charge (€/MW. year) or of just an annual charge (€/year). The first option runs into the problem of applying the same charge to all generation units with the same maximum capacity, which may have quite differing operation profiles. (the same occurs with demands that have widely different utilization factors and the same contracted capacity). The transmission charge should therefore be an annual charge (€/year) or a capacity charge computed separately for each type of generator or demand in each type of area in the system (see Olmos and Pérez-Arriaga 2009).

Olmos and Pérez-Arriaga also argue that the transmission tariff to be applied to each generator or load must be computed once and for all before its installation, since the level of this tariff should be based on the expected incremental contribution of this generator or load to the use of the grid (this is the driver of transmission investments). This means that the transmission charge to be paid by a network user should not be modified after its installation. Otherwise, the locational signal sent through this charge would be severely weakened.

Lastly, the process of implementation of new tariffs must be thought carefully. In order to avoid making big changes to the level of tariffs paid by already existing network users when introducing a new tariff scheme, the application of charges computed according to the new scheme could be limited to new network users. Alternatively, charges paid by already existing users could gradually evolve from the old tariff regime to the new one. In any case, the difference between the total cost of the grid and revenues from the application of tariffs should be socialized (preferably to demand).

2.7 Conclusions

Chapter 2 has analysed the effect that the grid should have on prices paid and earned by network users. Prices set should send both efficient short term signals driving operation decisions and long term ones driving the development of the system. Additionally, prices should provide an adequate remuneration of the transmission service guaranteeing its economic viability. Therefore, prices applied should be able to recover 100% of the regulated cost of the grid. No single set of prices seems to be able to meet all the aforementioned requirements, nor the sale of FTRs aimed at hedging the corresponding energy price differences. Thus, at least two set of transmission related prices must be applied.

Energy prices are aimed at driving operation decisions. Nodal prices, also called locational marginal prices, are deemed to be optimal energy prices because, assuming perfect information and competition, they encourage market agents to make socially optimal short-term decisions. Nodal prices internalize the effect of network losses and congestion on operation costs. However, in many real life systems, differences among nodal prices are small. Then, applying a single energy price (Single Pricing) or a price common to all the nodes within each of a set areas (zonal pricing) is considered to be preferable.

Net revenues resulting from the application of locationally differentiated energy prices, or from the sale of FTRs corresponding to commercial power transactions taking place, fall short of those needed to recover the whole cost of the grid. Then, additional charges, normally called transmission charges, or complementary charges, must be applied to complete the recovery of the grid cost. Complementary charges applied should allocate the cost of lines to those network users responsible for their construction. The electrical usage of lines by agents may be used as a proxy to network cost causality. However, it is the incremental usage made of new lines by new agents what determines the network reinforcements to be made. Therefore, network usage factors produced by most network cost allocation methods are useless, while average network usage factors produced by other methods like Average Participations or the Aumann-Shapley method must be modified to reflect the incremental nature of flows driving the development of the grid. Last but not least, in order for transmission charges not to interfere with the short term decisions by network users (to be driven by energy prices), they should be computed, once and for all, before the corresponding generators or loads are installed, and should take into account the expected increase in network flows that may result from the installation of the latter over all the set of possible operation situations that may occur along the economic life of these generators or loads. Besides, network tariffs should be applied as a fixed annual charge or a capacity charge computed separately for each type of generator or demand in each area in the system.

References

Aguado JA, Quintana VH, Madrigal M, Rosehart WD (2004) Coordinated spot market for congestion management of inter-regional electricity markets. IEEE Trans Power Syst 19(1): 180–187

Alomoush MI, Shahidehpour SM (2000) Contingency-constrained congestion management with a minimum number of adjustments in preferred schedules. Electr Power Energy Syst 22: 277–290

Baran ME, Babunarayanan V, Garren KE (2000) Equitable alloation of congestion relief cost to transactions. IEEE Trans Power Syst 15(2):579–585

Bialek J (1996) Tracing the flow of electricity. IEE Proc Gener Trans Distrib 143(4):313–320

Biskas PN, Bakirtzis AG (2002) Decentralised congestion management of interconnected power systems. In: IEE Proc Gener Trans Distrib, pp 432–438

Boucher J, Smeers Y (2001) Towards a common European electricity market-paths in the right direction... still far from an effective design. Harvard Electricity Policy Group. http://www.ksg.harvard.edu/hepg/Standard_Mkt_dsgn/Smeers_Interconnections1_4jni_3.do1.pdf

Chao HP, Peck S (1996) A market mechanism for electric power transmission. J Regul Econ 10: 25–29

Chao HP, Peck S (2000) Flow-based transmission rights and congestion management. Electr J 13(8):38–59

Cheng X, Overbye TJ (2006) An energy reference bus independent LMP decomposition algorithm. IEEE Trans Power Syst 21(3):1041–1049

Christie RD, Wangensteen I (1998) The energy market in Norway and Sweden: congestion management. IEEE Power Eng Rev 18:61–63

Consentec/Frontier (2004) Analysis of cross-border congestion management methods for the EU internal electricity market, consentec and frontier economics. Prepared for the European Commission. Directorate-General Energy and Transport. Web page: http://europa.eu.int/comm/energy/electricity/publications/index_en.htm

Dismukes DE, Cope RF III, Mesyanzhinov D (1998) Capacity and economies of scale in electric power transmission. Utilities Policy 7(3):155–162

ETSO (2001) Co-ordinated auctioning: a market-based method for transmission capacity allocation in meshed networks. Association of European Transmission System Operators (ETSO). Web address: http://www.etso-net.org/upload/documents/Coordinated%20Auctioning.pdf, p 22

ETSO (2004) An overview of current cross-border congestion management methods in Europe. European association of Transmission System Operators (ETSO). Web address: http://www.etso-net.org/upload/documents/Current_CM_methods_final_20040908.pdf, p 31

ETSO (2006) Transmission risk hedging products. Solutions for the market and consequences for the TSOs. ETSO Background Paper, 20 April 2006. http://www.entsoe.eu

ETSO/EuroPEX (2004) Flow-based market coupling: a joint ETSO-EuroPEX proposal for cross-bcongestion management and integration of electricity markets in Europe. Association of European Transmission System Operators (ETSO) and association of European Power Exchanges (EuroPEX)address: http://www.etso-net.org/upload/documents/ETSO-EuroPEX_Interimreport_Sept-2004-.pdf, p 26

European Transmission System Operators (ETSO) (1999) Evaluation of congestion management methods for cross-border transmission. http://www.etso-net.org/, p 22

European Transmission System Operators (ETSO) (2008) ETSO overview of transmission tariffs in Europe: synthesis 2007. ETSO tariffs task force. Web page: http://www.etso-net.org/upload/documents/11.a.%20Final_Synthesis_2007_18-06-08.pdf, p 27

Fang RS, David AK (1999) Optimal dispatch under transmission contracts. IEEE Trans Power Syst 14(2):732–737

Galiana FD, Ilic M (1998) A mathematical framework for the analysis and management of power transactions under open access. IEEE Trans Power Syst 13(2):681–687

Galiana FD, Conejo AJ, et al. (2003) Transmission network cost allocation based on equivalent bilateral exchanges. IEEE Trans Power Syst 18(4):1425–1431

Henney A (2002) What the US could learn from Western Europe and elsewhere. Electr J 15(10): 53–64

Hogan WW (1992) Contract networks for electric power transmission. J Regul Econ 4:211–242

Hogan WW (2002) Financial transmission right formulations. Cambridge. Center for Business and Government, John F. Kennedy School of Government, Harvard University. http://www.ksg.harvard.edu/hepg

Junqueira M, da Costa LC, Barroso OGC, Thomé LM, Pereira MW (2007) An aumann-shapley approach to allocate transmission service cost among network users in electricity markets. IEEE Trans Power Syst 22(4):1532–1546

Kirschen D, Allan R, Strbac G (1997) Contributions of individual generators to loads and flows. IEEE Trans Power Syst 12(1):52–60

Lusztig C, Feldberg P, Orans R, Olsonet A (2006) A survey of transmission tariffs in North America. Energy 31(6–7):1017–1039

Newbery DV, Damme E, von der Fehr NM (2003) Benelux market integration: market power concerns. Dutch energy regulator (DTe). Web page: http://www.dte.nl/nl/Images/12_13216.pdf

O'Neill R, Helman PU, Hobbs BF, Stewart WR Jr, Rothkopf MH (2002) A joint energy and transmission rights auction: proposal and properties. IEEE Trans Power Syst 17(4):1058–1067

Olmos (2006) Regulatory design of the transmission activity in regional electricity markets. Ph.D. dissertation, Universidad Pontificia Comillas

Olmos L, Neuhoff K (2006) Identifying a balancing point for electricity transmission contracts. IEEE Trans Power Syst 21(1):91–98

Olmos L, Pérez-Arriaga IJ (2007) Evaluation of three methods proposed for the computation of inter-TSO payments in the internal electricity market of the European Union. IEEE Trans Power Syst 22(4):1507–1522

Olmos L, Pérez-Arriaga IJ (2009) A comprehensive approach for computation and implementation of efficient electricity transmission network charges. Energy Policy 37(12):5285–5295

Oren SS, Ross AM (2002) Economic congestion relief across multiple regions requires tradable physical flow-gate rights. IEEE Trans Power Syst 17(1):159–165

Pérez-Arriaga IJ, Rubio FJ, Puerta JF, Arceluz J, Marín J (1995) Marginal pricing of transmission services: an analysis of cost recovery. IEEE Trans Power Syst 10(1):546–553

Rau NS (2000) Transmission loss and congestion cost allocation: an approach based on responsibility. IEEE Trans Power Syst 15(4):1401–1409

Rivier M, Pérez-Arriaga IJ (1993) Computation and decomposition of spot prices for transmission pricing. In: Proceedings of the power systems computation conference (PSCC), Avignon

Rivier M, Pérez-Arriaga IJ, Luengo G (1990) JUANAC: a model for computation of spot prices in interconnected power systems. In: Proceedings of the 10th PSCC conference, Graz, 19–24 Aug 1990

Rubio FJ, Pérez-Arriaga IJ (2000) Marginal pricing of transmission services: a comparative analysis of network cost allocation methods. IEEE Trans Power Syst 15(1):448–454

Ruff LE (1999) Competitive Electricity markets: why they are working and how to improve them. http://www.ksg.harvard.edu/hepg,n/e/r/a

Schweppe FC, Caramanis M, Tabors RD, Bohn RE (1988) Spot pricing of electricity. Kluwer, Boston

Singh H, Hao S et al (1998) Transmission congestion management in competitive electricity markets. IEEE Trans Power Syst 13(2):672–680

Stoft S (1998) Congestion pricing with fewer prices than zones. Electr J 11(4):23–31

Tabors R, Caramanis M (2000) Real flow, a preliminary proposal for a flow-based congestion management system. Cambridge, MA. http://www.ksg.harvard.edu/hepg/flowgate/Real%20Flow.pdf

Tao S, Gross G (2002) A congestion management allocation mechanism for multiple transaction networks. IEEE Trans Power Syst 17(3):826–833