Bocconi

IEFE

Centre for Research on Energy and Environmental Economics and Policy

Research Report Series – ISSN 2036-1785

Integrating European Electricity Markets

Giuseppe Buglione, Guido Cervigni,
Eileen Fumagalli, Elena Fumagalli and Clara Poletti

Research Report n. 2

October 2009

IEFE - The Center for Research on Energy and Environmental Economics and Policy at Bocconi University
via Guglielmo Röntgen 1, I-20136 Milan
tel. +39.02.5836.3820 – fax +39.02.5836.3890
www.iefe.unibocconi.it – iefe@unibocconi.it

Giuseppe Buglione (IEFE-Bocconi)

Guido Cervigni (IEFE-Bocconi and Lecg Italy s.r.l.)

Eileen Fumagalli (IEFE-Bocconi and Lecg Italy s.r.l.)

Elena Fumagalli (Politecnico di Milano)

Clara Poletti (IEFE-Bocconi)

We are very grateful to Fabrizio Carboni, Fabio Massaro and Giusi Squicciarini for comments, discussions and suggestions that substantially improved this research report.
Financial support from GME-Gestore del Mercato Elettrico S.p.A. and Terna S.p.A. is gratefully acknowledged.

Introduction

The creation of a single Internal Electricity Market is a prominent item in the European agenda. This research provides a comprehensive and up-to-date analysis of the issues surrounding theoretical and policy developments in this area.

The report is structured as follows. The first chapter discusses the economic foundations of the mechanisms being developed to integrate national electricity marketing. We cast the market integration issue within the more general context of congestion management. We analyze the impact of specific technical features of electricity on the efficient prices of electricity, on the format and content of the rights to use transmission resources, as well as on the features of alternative congestion management mechanisms.

The second chapter deals with market power in electricity markets. The section surveys the main contributions in the economic literature in a common analytical framework. It also provides several examples in order to show how different elements of market design might be crucial in affecting the presence of market power and the way it might be exercised. In particular, the fact that firms might exercise market power by implementing strategies that in other industries would not be interpreted as reflecting market power calls for specific market monitoring in wholesale electricity markets.

The third chapter links the theoretical findings of the first chapter with the current political debate on the most efficient mechanisms for coupling national electricity markets. We review the alternative options being discussed to address the main design issues and we cast them in a unifying framework.

The last chapter reports on the state of the different regional electricity initiatives (ERI). It also provides a detailed description of the main features of the national markets belonging to the Central-South Region, identifying the specific harmonisation needs in the area.

Chapter I

A primer on electricity transmission rights

1. Networks effects, the value of electricity, the value of transmission

1.1 Loop flow

1.1.1 Basic definitions

A transmission network is a set of power lines connecting different locations. The intersections of these power lines are called network "nodes" or "buses". At each network node there might be generators (who inject power in the system) and loads (who withdraw power). We will call g the amount of power injected by generators and d the amount of power withdrawn from loads. g and d are expressed in MW.

We call **net injection** at node i, y_i, the difference between total injections and total withdrawals at node i:

$$y_i = g_i - d_i$$

Positive and negative net flows in different nodes cause power flows along the network. A node whose net injection is positive, i.e. a node whose power generation exceeds demand, will be called an **injection node**. On the contrary, a node characterized by a negative net injection, i.e. a node whose power demand exceeds generation, will be a **delivery node**.

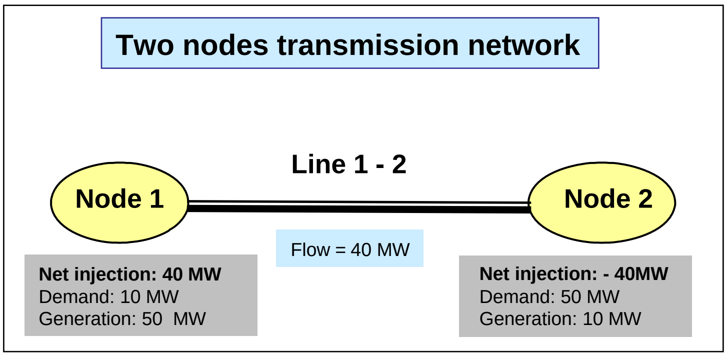

Figure 1 – Two nodes transmission network

In this work we will not investigate in detail the causes of transmission capacity limits. We will generally assume that thermal limits constraint the capacity of one or more lines.

1.1.2 Physical properties of power networks

In analyzing power networks, we adopt the standard **DC** approximation assumption: we will consider only real power flows. As a second simplifying assumption we will assume zero **transmission losses**[1].

For our tasks it is important to introduce some basic physical properties of power networks.

Firstly, when power moves between nodes of the transmission network, it does not follow a unique path; it rather follows every available parallel path between the injection node and the delivery point.

This phenomenon is called **loop flow.**

There are two principles that rule the way power flows along the network:

- **PRINCIPLE OF INVERSE RESISTANCE**.

  The share of power that flows along each path is inversely proportional to the relative resistance of such a path: the lower the resistance of a line, the larger share of power flows on that path.

- **PRINCIPLE OF SUPERPOSITION.**

  If traders are grouped so that each group is balanced (i.e. inflow equals outflow) the flow from all the groups trading at once is the sum of the flows the balanced groups would produce individually.

We now develop some examples to show how these properties apply in our stylized networks.

[1] This is a standard assumption always adopted in the literature in order to simplify computation of equilibrium outcomes.

EXAMPLE 1

(PRINCIPLE OF INVERSE RESISTANCE)

Assume we have a network composed by two nodes: node 1, the injection node, i.e. a node characterized by a positive net injection ($y_1 = g_1 - d_1 > 0$) and node 2, the delivery node, i.e. a node characterized by a negative net injection ($y_2 = g_2 - d_2 < 0$).

These two nodes are connected by two lines (A and B). Suppose also that line A has twice the resistance of line B (i.e. $R_A = 2\ R_B$).

Assume, for the time being, that there is no limit on available transmission capacity.[2]

Because of the loop flow phenomenon, power will flow from node 1 to node 2 on every parallel path, which means – in our example – both on line A and on line B.

We can be more precise: applying the Principle of Inverse Resistance, we can compute the share of power injected in node 1 that flows along each of the two lines. The share of power that flows along each path is inversely proportional to the relative resistance of such a path.

This share is called the Power Transmission Distribution Factor (PTDF) and it can be computed as:

$$PTDF_A = R_B \,/\, (R_A + R_B)$$

$$= R_B \,/\, 3\ R_B = 1 \,/\, 3$$

$$PTDF_B = R_A \,/\, (R_A + R_B)$$

$$= 2\ R_B \,/\, 3\ R_B = 2 \,/\, 3$$

Hence, we can conclude that one third of the power injected at node 1 will flow along line A and two thirds of that power will flow along line B.

[2] We will relax this assumption in the next example

Once we know the entity of the injection in node 1, say $y_1 = 75$ MW, we can then compute the exact amount of power that flows (PF) on each of the two lines.

$$PF_A = y_1\ PTDF_A = y_1 / 3 = 25MW$$

$$PF_B = y_1\ PTDF_B = 2\ y_1 / 3 = 50MW$$

The following picture summarizes this first example:

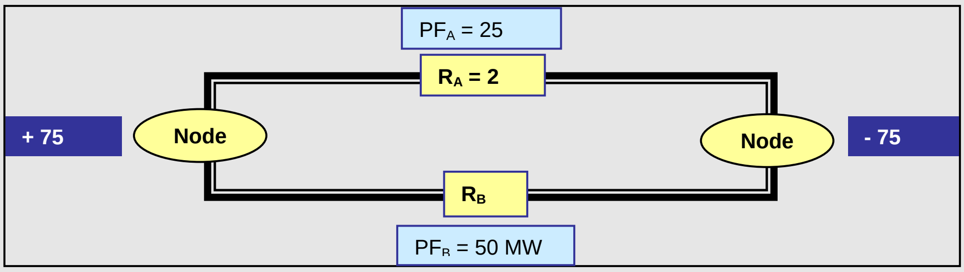

Figure 2 – Example 1 – The principle of inverse resistance

EXAMPLE 2

(CAPACITY CONSTRAINTS)

We now relax the assumption of unlimited capacity, introducing a thermal constraint of 50 MW on line A.

The introduction of such a constraint has an impact on the maximum amount of power that can be moved from node 1 to node 2, even if the capacity of line B is unlimited.

Applying the Principle of Inverse Resistance, we can compute the power transmission distribution factors: we know that $PTDF_A = 1 / 3$, that means that one third of the power injected at node 1 has to flow along line A, i.e. $PF_A = y_1 / 3$.

Hence, we can write explicitly the thermal constraint on line A:

$$PF_A \leq 50\ MW$$

$$y_1 / 3 \leq 50\ MW$$

$$y_1 \leq 150\ MW$$

This means that – due to the presence of a transmission constraint on line A – the maximum amount of power that can be transferred from node 1 to node 2 is 150 MW (even if the capacity of line B is unlimited).

The following picture provides a graphical representation of this second example.

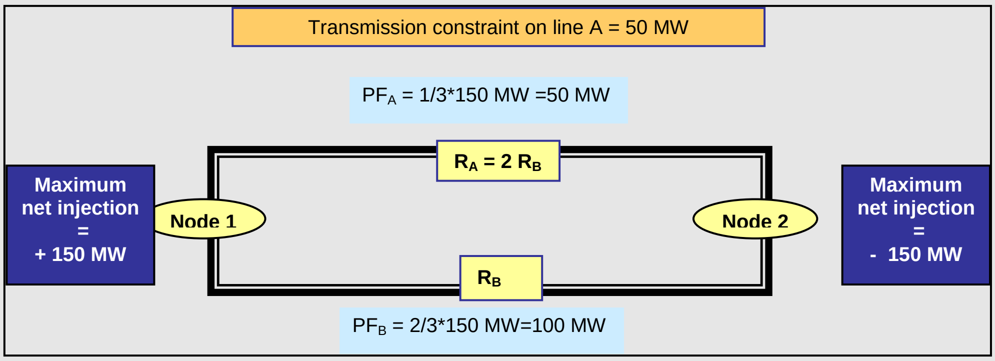

Figure 3 – Example 2 – Capacity constraints

EXAMPLE 3

(PRINCIPLE OF SUPERPOSITION)

In order to determine the implications of the principle of superposition, we consider a three-node network.

Assume that

$$y_1 = y_2 = 150\ MW$$

$$y_3 = -300$$

This means that node 1 and node 2 are (net-)injecting 150 MW and node 3 is (net-)loading 300 MW.

We also assume that the three transmission lines have equal resistance and unlimited transmission capacity.

Figure 4 shows the setup of this example.

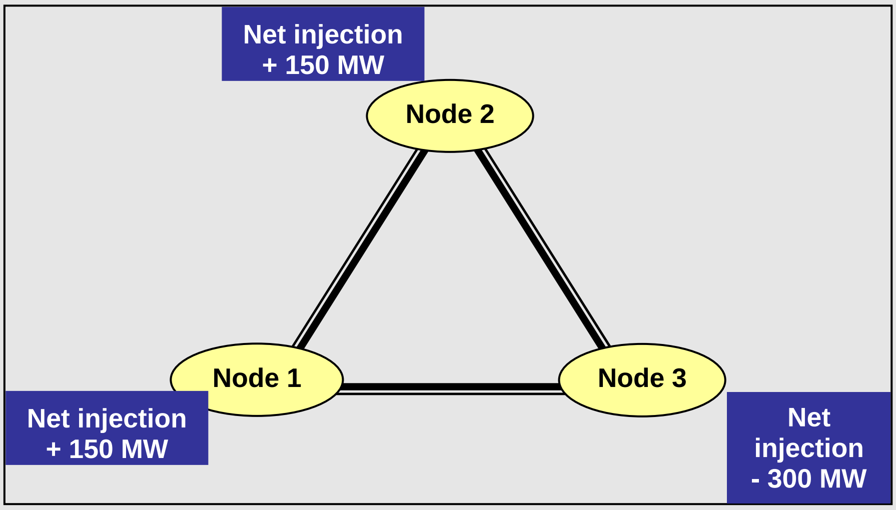

Figure 4 – Example 3 – The setup

Thanks to the principle of superposition we can determine the overall power flows on the network as the sum of the flows that would be caused independently by each node's net-injection.

In order to assess the flow caused by the net-injection at a node, we need to identify a reference node (or "**swing node**") in the network; we assume that a net-load of the same value to the net-injection under examination takes place at the swing node. The choice of swing node is arbitrary and does change the result. In this example we take node 3 as the swing bus.

Let's first consider the power injection in node 1; the power flow from node 1 to node 3 can move along two possible paths: the direct path, (1-3), or the indirect one, over lines (1-2) and (2-3).

Since all the lines have the same resistance, we have that the resistance of a path depends solely on its length. So we have that the resistance associated to the direct path (1-3) – which is twice as short as the indirect one – is exactly half of the resistance of the indirect one, (1-2) and (2-3).

So the principle of inverse resistance gives us the shares of power injected in node 1 that reaches node 3 along any specific path (i.e. PTDFs): [3]

$$PTDF^{1}_{1\text{-}3} = (R_{1\text{-}2} + R_{2\text{-}3}) / (R_{1\text{-}2} + R_{2\text{-}3} + R_{1\text{-}3})$$

$$= 2\,R / 3R = 2/3$$

$$PTDF^{1}_{1\text{-}2,2\text{-}3} = (R_{1\text{-}3}) / (R_{1\text{-}2} + R_{2\text{-}3} + R_{1\text{-}3})$$

$$= R / 3R = 1/3$$

Concluding, we have 100 MW flowing along line (1-3) and 50 MW along lines (1-2) and (2-3).[4]

$$PF^{1}_{1\text{-}3} = 2\,y_1 / 3 = 100\ MW$$

$$PF^{1}_{1\text{-}2} = y_1 / 3 = 50\ MW$$

$$PF^{1}_{2\text{-}3} = y_1 / 3 = 50\ MW$$

These power flows resulting from the injection of 150 MW in node 1 are represented in Figure 5.

[3] We define $R_{i\text{-}j}$ the resistance of link *(i-j)*

[4] Conventionally, we represent the power flow along line (**i-j**) with a positive number in case it has the same direction as the line (i.e. it flows from I to j in our example) and negative otherwise.

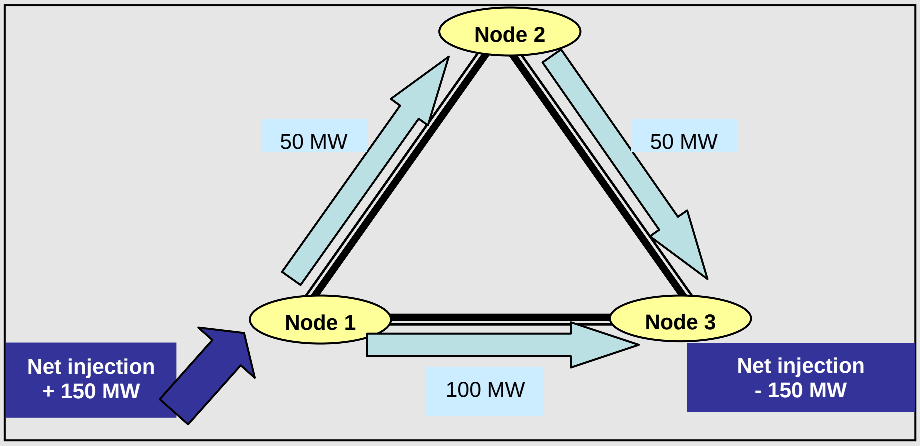

Figure 5 – Example 3 – Power flows from the injection in node **1**

Let's now consider node 2; we have now two possible paths for going from 2 to 3: (2-3) and (1-2), (1-3). So we have that the resistance associated to the direct path (2-3) – which is twice as short as the indirect one – is exactly half of the resistance of the indirect one, (1-2) and (1-3).

$$PTDF^2_{2\text{-}3} = (R_{1\text{-}2} + R_{1\text{-}3}) / (R_{1\text{-}2} + R_{2\text{-}3} + R_{1\text{-}3})$$

$$= 2\,R / 3R = 2/3$$

$$PTDF^2_{1\text{-}2} = -\ (R_{2\text{-}3}) / (R_{1\text{-}2} + R_{2\text{-}3} + R_{1\text{-}3})$$

$$= -R / 3R = -\ 1/3$$

$$PTDF^2_{1\text{-}3} = (R_{2\text{-}3}) / (R_{1\text{-}2} + R_{2\text{-}3} + R_{1\text{-}3})$$

$$= R / 3R = 1/3$$

So we have:

$$PF^2_{2\text{-}3} = 2\,y_2 / 3 = 100\ MW$$

$$PF^2_{1\text{-}2} = -\ y_2 / 3 = -\ 50\ MW$$

$$PF^2_{1\text{-}3} = y_2 / 3 = 50\ MW$$

These are the power flows following an injection of 150 MW in node 2; they are also represented in the following figure.

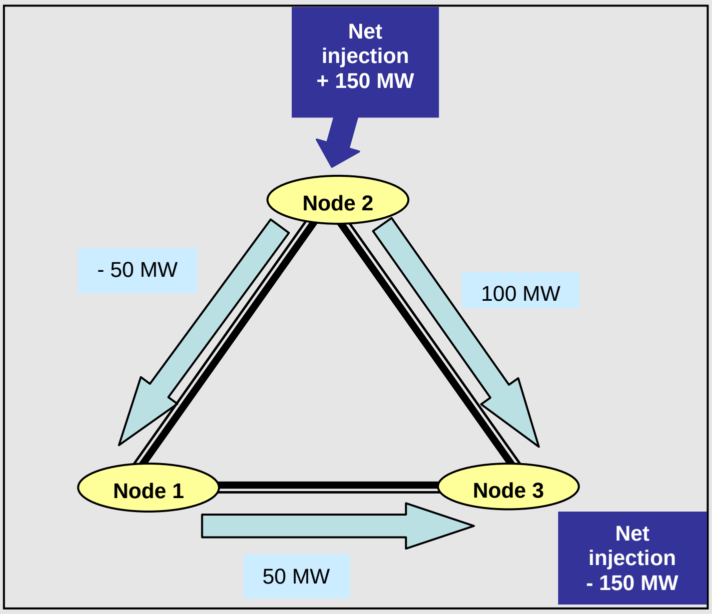

Figure 6 – Example 3 – Power flows from the injection in node 2

Since Node 3 is the swing bus, its net-load does not cause any flow on the network

Applying the principle of superposition, we can determine the overall flows on the network by adding the flows obtained considering separately each node of the grid:

$$PF_{1\text{-}3} = PF^{1}_{1\text{-}3} + PF^{2}_{1\text{-}3}$$

$$= 100\ MW + 50\ MW = 150\ MW$$

$$PF_{1\text{-}2} = PF^{1}_{1\text{-}2} + PF^{2}_{1\text{-}2}$$

$$= 50\ MW - 50\ MW = 0$$

$$PF_{2\text{-}3} = PF^{1}_{2\text{-}3} + PF^{2}_{2\text{-}3}$$

$$= 50\ MW + 100\ MW = 150\ MW$$

These net flows are represented in Figure 7.

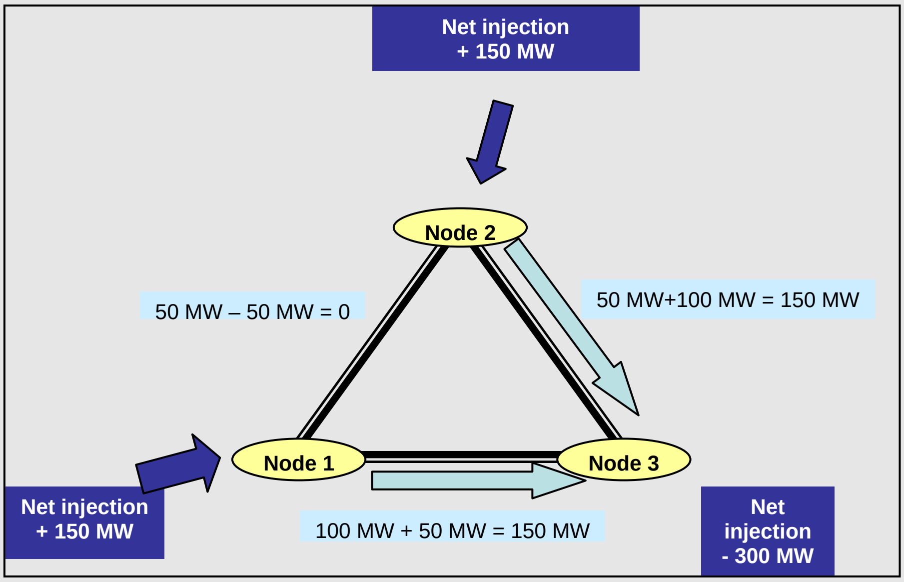

Figure 7 Example 3 – The principle of superposition

This simple example shows that when an agent causes a flow on an the network, there are (positive or negative) externalities on other agents operating on the same grid as the possibility for the network to support some flows is increased along some paths and decreased along some others.

1.2 The definition of transmission rights

In this section we are going to provide different definitions of transmission rights; in order to obtain the efficient allocation of the right to use the transmission system, the network effects highlighted in the previous section have to be internalized. Alternative market-based allocation mechanism may be built around different designs of transmission rights.

In the following Chapters we will deal in detail with the properties of each solution. Here we limit ourselves to surveying the main approaches. We do this in the simple network setting shown in the following picture:

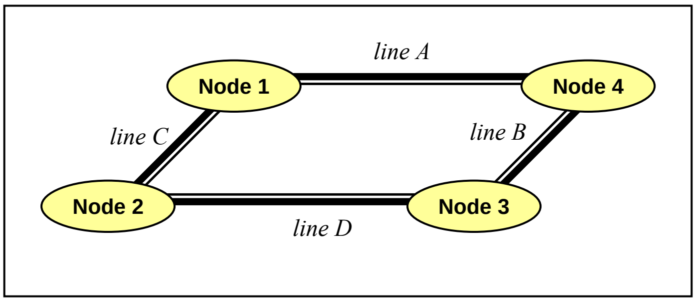

Figure 8 – The transmission grid

We assume that lines A and B have the same resistance equal to half the resistance of lines C and D. We consider the situation in which two agents would like to implement a transaction involving 300 MW injection in node 1 and 300 MW load in node 3.

1.2.1 Point-to-point approach

The first approach we will examine is the so called "Point-to-point" method. According to the US Federal Energy Regulator Commission (FERC):

> "Reservation of transmission capacity should permit the customer to receive up to a specific amount of power into the grid at the specified [Points Of Receipt], and to deliver up to a specific amount of power from the grid at specified [Points Of Delivery], on a firm basis."

Hence, a transmission right is designed as the right to inject power in a node and withdraw it from a different node. In our example, in order to set up their program, the parties involved in the transactions must secure 300 MW of an entitlement defined as the right to inject in node 1 and withdraw in node 3.

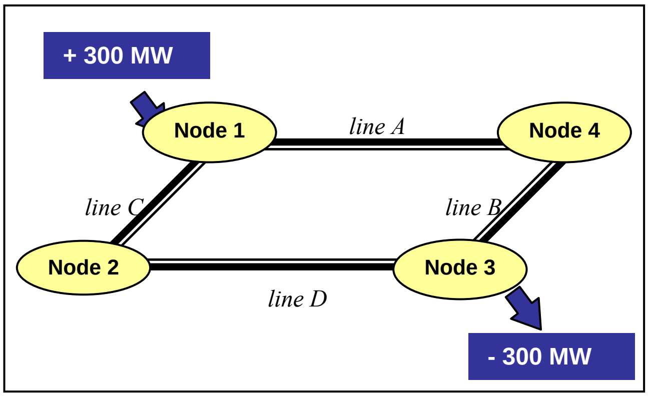

Figure 9 – The point-to-point approach

A crucial feature of this definition of transmission right is that this right is not directly related with actual flows. As it will be discussed more in detail in Chapter XX, actual flows will in general depart from those corresponding to the quantity of transmission rights allocated.

1.2.2 Flow-gate approach

An alternative approach defines transmission rights in terms of the right to use network elements (lines, transformers….). This implies that a party willing to transfer power from one node to another will have to purchase rights to use all the network elements involved.[5]

In our example, the parties involved in the transaction would have to buy the right to use all the four lines of the transmission grid.
Since one third of the 300 MW injected in node 1 will flow along lines C and D and two thirds will pass through lines A and B[6], in order to set up his transmission program, the agent has to buy the right to use 200 MW of capacity for lines A and B and 100 MW for lines C and D.

We can see the representation of our example's actual flows in the following figure.

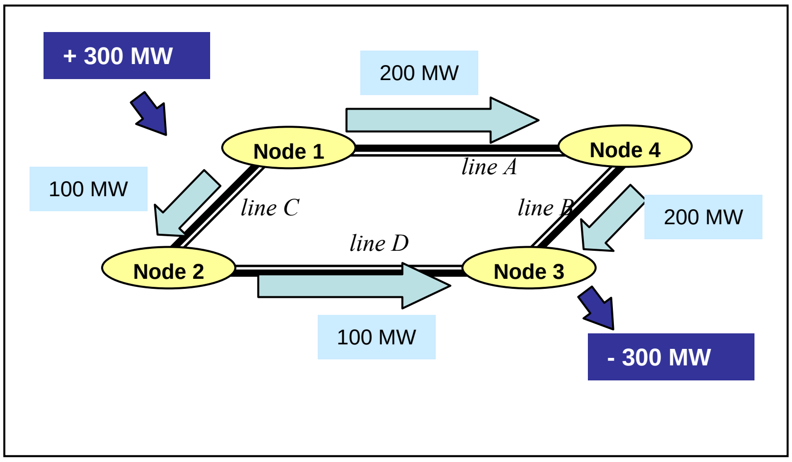

Figure 10 – The Flow-Gate approach

A thorough analysis of this design, within the framework of the more general "**flowgate**" approach, is carried out in Chapter 3.4.

---

[5] Proponents of this approach actually suggest that those rights should be defined only for potentially congested network elements (Chao Peck 1996)
[6] Given the assumptions about resistances is $R_A = R_B = R_C/2 = R_D/2 = R$, PTDFs are:
$PTDF^1_{A,B} = (R_C + R_D) / (R_A + R_B + R_C + R_D) = 4R / 6R = 2/3$
$PTDF^1_{C,D} = (R_A + R_B) / (R_A + R_B + R_C + R_D) = 2R / 6R = 1/3$

1.2.3 Interface approach

We finally consider the zonal or "**interface**" approach, which has been widely used also in the pre-liberalization era.[7] This approach is based on a highly simplified representation of the transmission grid. It splits the network in wide areas by identifying virtual interfaces. The transmission rights are defined as the right to inject power in a node at one side of the interface and withdraw it from any node at the other side of the interface.

Keeping the example previously introduced, assume that generators are mainly located at nodes 1 and 2, while loads are located mainly at nodes 3 and 4. A "natural" interface could then be set-up as shown in the following picture.

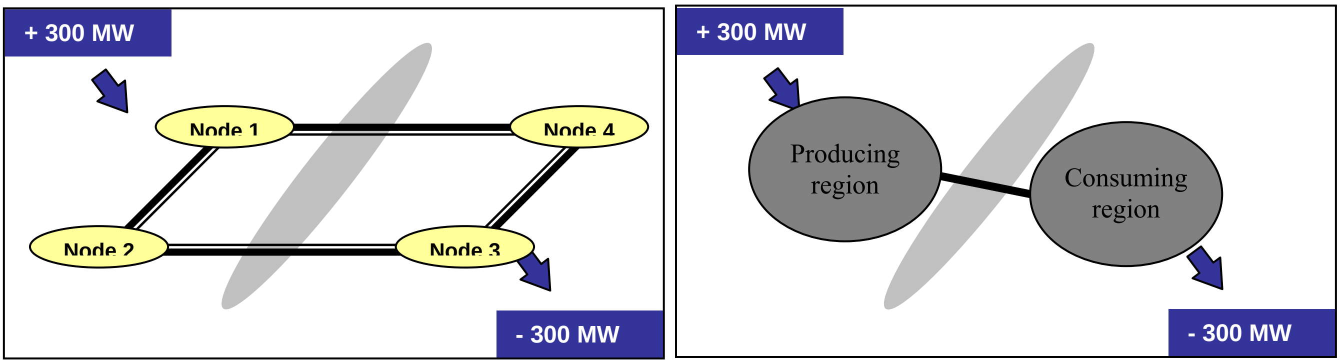

Figure 11 – The interfaces approach

Under this approach, a transmission right is then the right to inject power in one region and load it in another region.

We will discuss in Chapter XX pros and cons of this approach.
Here we just show a drawback of this method, showing when the transmission operator is to define the available level of transmission capacity.

[7] In the industry jargon, this approach is often labeled as contract path fiction.

Let's go back to our example and assume that the line connecting node 2 and node 3 (*line D* in Figure 8) has a thermal limit of 100 MW.
Suppose we should now estimate the interface transfer capacity between the two regions.

In case (i) we have an injection of 300 MW in node 1 and a 300 MW load in node 3; this represents the maximum amount that can be transferred from 1 to 3, given the thermal constraint on line (2-3), so we could say that the interface transfer capacity is 300 MW.

Let us now consider case (ii), where we have injection at node 2: given the presence of the thermal constraint[8], we know that the maximum feasible injection in 2 is 150 MW. Hence the interface transmission capacity would be in this case 150 MW.

Figure 12 summarizes the two cases introduced above.

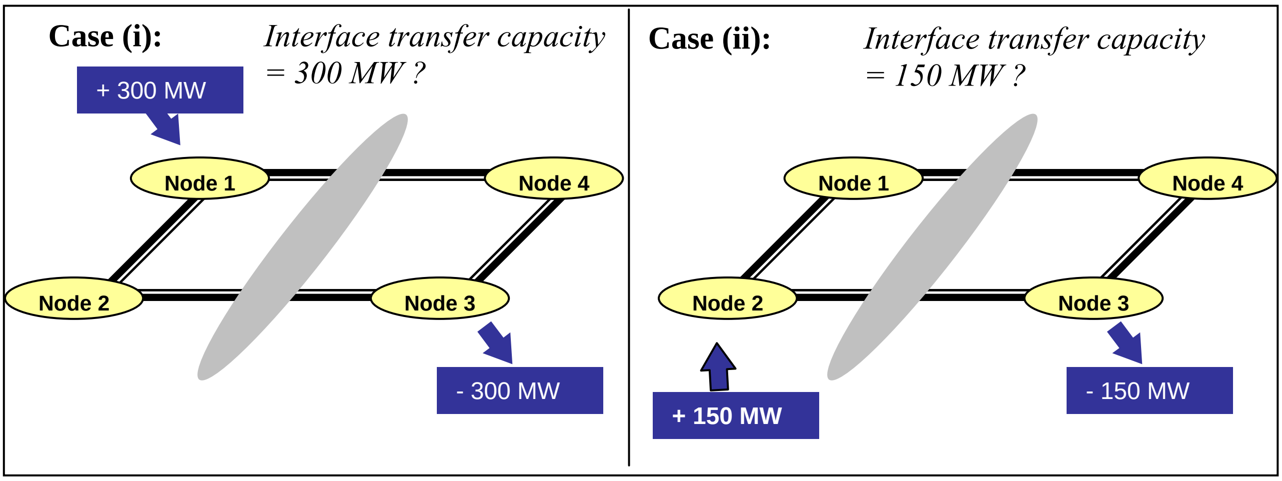

Figure 12 – The interfaces approach

[8] In this case we have to consider the fact that - given the relative resistances of the lines (see the assumptions we made in the beginning) - we know that two thirds of the power injected in node **3** will flow along the line (**2-3**); this and the thermal limit of 100 MW on that line give us the maximum feasible amount of injections in node 3.

1.3 The value of power along the network

The presence of transmission constraints causes electricity market equilibrium prices to differ among nodes of the network. Each set of market equilibrium prices implicitly determines the value for the transmission grid itself, or better, the value for the transmission rights.

The economics of electricity transmission rights may be analyzed either in a "market equilibrium" framework or in a "welfare maximization" setting.
In a market equilibrium framework, nodal prices and the value of transmission rights are shown to result, in case of congestion, by the interactions of market participants (generators, loads, arbitrageurs …) in the energy market. This framework is appealing since energy and transmission prices result from the usual decentralized market process. We show this approach in section 2.3.1.
In practice, network effects are such that a fully decentralized mechanism for the reallocation of transmission rights among market participants may generate inefficiencies. High transactions costs are necessary to match, on the one hand, generators and customers who give the highest value to the transmission capacity and, on the other hand, transmission-right holders, all the more so since the efficient allocation of transmission rights changes very frequently, commonly on a hour-by-hour basis.
For those reasons, institutions and trading arrangements implemented in most electricity markets are based on some degree of centralization. An auctioneer is typically set up to allocate the rights to use the transmission system. The logic underlying the algorithms implemented in centralized auctions is the same as the one that would drive a welfare maximizing social planner controlling generation, loads and the transmission system. This leads us to section 2.3.2, in which electricity nodal values are shown to be the result of the problem of maximizing the net surplus generated power transactions. Standard welfare economic theory guarantees that the outcome characterized as the solution of the surplus maximization problem may be obtained as a decentralized market outcome, as long as there are no transaction costs. Institutions developed to reduce transactions costs in the industry are discussed in Ch. 3.

1.3.1 The decentralized approach

In the decentralized – or market – approach, trading takes place at the same time between buyers and sellers of electricity and between buyers and sellers of transmission rights, which are necessary to carry out transactions involving injection in one node and withdrawal in a different node.

Assume we have the following very simple setup: there are two trading locations connected by a single 500 MW transmission line.
The first location – Node 1 – is a production bus with a relatively small local demand of 100 MW.
Costs of production are defined as follows: the costs of power generation starts at 20 €/MWh and increases at a rate of 2 €/MWh for each additional 100 MW of output (in analytical terms, this means that the marginal costs of production for the producer located in node 1 are $MC_1=20+Q/50$)
The second location – Node 2 – is a consumption bus having a local demand equal to 800 MW.
The costs of power generation in node 2 are similar to the costs defined for producers in node 1, the only difference being in the starting cost, which is higher: $MC_2=40+Q/50$.

The following picture summarizes the assumption we made so far:

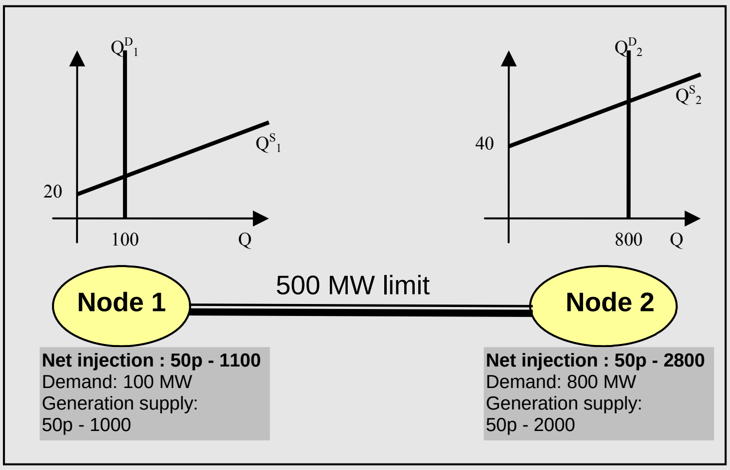

Figure 13 . A two-node transmission grid

(i) SEPARATED LOCATIONS (AUTARCHY SOLUTION)

Let us first analyze what would happen in case the two nodes were not connected. In such a case, each generator would serve its domestic demand, that is:

$$g_i^A = d_i \qquad \text{for } I = 1,2$$

and prices would be

$$p_1^A=20+100/50=22\ €/MWh$$

$$p_2^A=40+800/50=80\ €/MWh$$

where the superscript "A" stands for the Autarchy solution.

(II) LOCATIONS CONNECTED BY A LINE WITHOUT CAPACITY LIMITS

As a transmission line is available between the two nodes (suppose without any transmission capacity constraint for the moment), the two markets can be substantially considered as one: all consumers located in node 2 will be able to buy from both firms.

They will obviously choose to buy the cheaper power produced by the firm located in node 1.[9]

So we will have that in the unconstrained (UC) solution

$$g_1^{UC}= 900\ MW$$

$$g_2^{UC}= 0\ MW$$

[9] Notice that we have no capacity constraint on the production side. If there was one, this could change the allocation of production even without constraints on the line. Anyway this additive assumption would not introduce any improvement in the understanding neither of the way the efficient allocation of production might be found, nor in the evaluation of the transmission rights. This is why we adopt the simplification of infinite producing capacity in what follows.

The price will be equal to the marginal cost of producing one extra unit of power, that is, in this case, the marginal cost of the firm located in node 1.

So the prices in the two nodes will be the same (the producer of the marginal unit is firm 1 in both cases):

$$p_1^{UC} = p_2^{UC} =$$

$$=20+900/50$$

$$=38 \text{ €/MWh}$$

Thus, the firm located in node 2 cannot sell anything as its marginal cost of production would be higher than the equilibrium market price.

It is now interesting to notice that in case there is no binding limit on transmission capacity, the price in the two nodes is the same (as we pointed out before, the price in one node is equal to the marginal cost of producing one extra unit of consumption in that node; this is equal to the marginal cost of the most efficient firm).

The fact that there is a unique price means that the value of the transmission right is null (nobody would be willing to pay for the right to transmit capacity from one node to the other). This will be clearer in what follows.

(III) LOCATIONS CONNECTED BY A 500-MW LINE

Introducing the limited capacity assumption (in particular, assuming that the line connecting the two nodes has a thermal limit of 500 MW) we change completely the result: not all consumers located in node 1 will be able to buy power from the producer located in node 2; the limited capacity of the line will make the producers located in node 1 produce for the domestic consumption (100MW) and produce the maximum that the line can sustain (500MW) for the consumers located in node 2.

This means that the equilibrium production levels, in this constrained (C) setting, will be:

$$g_1^C = 600\ MW$$

$$g_2^C = 300\ MW$$

so two different prices arise in the two nodes:

$$p_1^C = 20+600/50=32\ €/MWh$$

$$p_2^C = 40+300/50=46\ €/MWh$$

Note that these prices have been determined by a perfectly competitive, fully decentralized bilateral market and are not the result of a centralized optimization process.

Transmission rights are not bundled with energy, but are traded separately.
Which is the market value of such rights?
To sell and transmit power from node 1 to node 2, either the buyer or the seller must own a transmission right in the amount of the power sold.
On the one hand, as long as the price in node 1 is smaller than the price in node 2, consumers located in node 2 will be willing to pay for the transmission right in order to be able to buy cheaper power; in our specific case, power is cheaper by 14 €/MWh in node 1, so consumers located in node 2 are willing to pay up to 14 €/MWh for a transmission right, to gain the access to the cheaper power.
On the other hand, as long as the price in node 2 is higher than the price in node 1 the producers located in node 1 are willing to pay up to the price difference (14 €/MWh in the specific case we are analyzing) to gain access to node-2 customers.
If the transmission right price was less than 14 €/MWh, then node-2 producer and node-1 consumers would buy those rights.

So market forces would drive up the price of transmission rights. Competitive pressures in the bilateral market will drive the price of transmission rights to the energy price difference between the two nodes.

We can generalize this result by saying that, if the market of transmission rights is competitive, then the equilibrium price of a right to transmit power from node *i* to node *j*, P_{ij}, is equal to the price of power in the final node, P_j, minus the price of power in the injection node, P_i.

$$P_{ij} = P_j - P_i$$

As we pointed out before, this is the result of a competitive decentralized market. Since these are competitive prices, they imply that suppliers are minimizing their production costs.

It is easy to check that a social planner pursuing the objective of minimizing production costs[10] would allocate the production in the same way as the competitive market. This would then lead to the same nodal equilibrium prices and to the same evaluation of transmission rights.

Let's consider the cheapest way to supply the load: since the most efficient producers are the ones located in node 2, these are the first to be used.
Given the 500MW transmission limit, the generators in node 2 can only provide part of the load on node 1: they will produce 500MW, with a marginal cost of 32 €/MWh (which is well below the marginal cost of production of the producers located in node 1; this implying that it is convenient to "import" from node 2 all the 500MW).
The 300MW that are missing to satisfy the load have to be produced by the most efficient generators located in node 2.

[10] In this specific setup with inelastic demand, social welfare maximization coincides with production costs minimization.

Let's now compare this allocation of production with the one induced by the competitive locational prices; if the locational price in node 1 is 46 €/MWh, then only the producers with marginal costs below this value will produce, and this is exactly what we had before: the 300MW are produced by the most efficient firms in node 2 (all other producers would lose money by producing at this price).
The same considerations can be done for node 1.

We can conclude that, if the target is to impose the cost minimizing allocation of production in the example network through free market, one should use competitive locational prices.

This result is very appealing: market forces can drive to socially optimal production allocations and transmission right pricing. The problem is that this is easily computed in the two-node single-line network, but once we leave the single-line transmission network, things start to be very complicated.

Besides the computational complications, there is an even more relevant problem of coordination among agents. Assume the transmission rights have been initially allocated to a set of agents that does not necessarily coincide with the set of generators and consumers. In order for the efficient outcome to occur, it is necessary that producers perfectly forecast the transmission value, but this means that they should be able to anticipate the whole market outcomes. These generators would then meet the transmission rights owners who have the same prediction. All this procedure should be replicated for all the hours in a year… That is why it is unrealistic to expect an efficient outcome from the decentralized approach; in the next subsection we present a different approach.

1.3.2 The centralized approach

We now introduce the analysis of a more complex setting assuming that the equilibrium is not going to be found by market participants through a decentralized approach, but through the intervention of a superior decision maker, the so-called benevolent planner.

In the standard economic theory a benevolent planner is a decision maker that maximizes social welfare; in our case, this means maximizing the net surplus generated by transactions (more specifically, the difference between the willingness to pay of the demand side and the variable costs of supply).

This setting is relevant for a number of reasons.

First, from a theoretical point of view, we know that the allocation outcomes resulting from the social planner's choice can be the equilibrium outcomes of a decentralized market.

Second, in an historical perspective, the maximization problem solved by the social planner is analogous to the one faced by the former monopolists acting in the power industry. Considering the very low level of elasticity of demand, the monopolist's cost minimization problem coincides with the total welfare maximization set up by a social planner.

Last, from the point of view of current institutions, market mechanisms (already in place or under construction) shaped as auction in most electricity markets have total welfare as a target function.

> Assume we have the following setup: there are three trading locations connected by a transmission network in which one of the lines has a 100 MW- thermal constraint.
> The first location – Node 1 – is characterized by a local demand of 300 MW. Costs of production are defined as follows: the costs of power generation starts at 20 €/MWh and increases at a rate of 2 €/MWh for each additional 100 MW of output (in analytical terms, this means that the marginal costs of production for the producer located in node 1 are $MC_1=20+Q/50$).

The second location – Node 2 – shows a local demand equal to 600 MW. The costs of power generation in node 2 are similar to the costs defined for producers in node 1, the only difference being in the starting cost, which is higher: $MC_2=30+Q/50$.

The third location – Node 3 – shows a local demand equal to 600 MW. The costs of power generation in node 3 are similar to the costs defined for producers in the other nodes, the only difference being in the starting cost, which is the highest: $MC_2=50+Q/50$.

The following picture summarizes the assumptions we did so far:

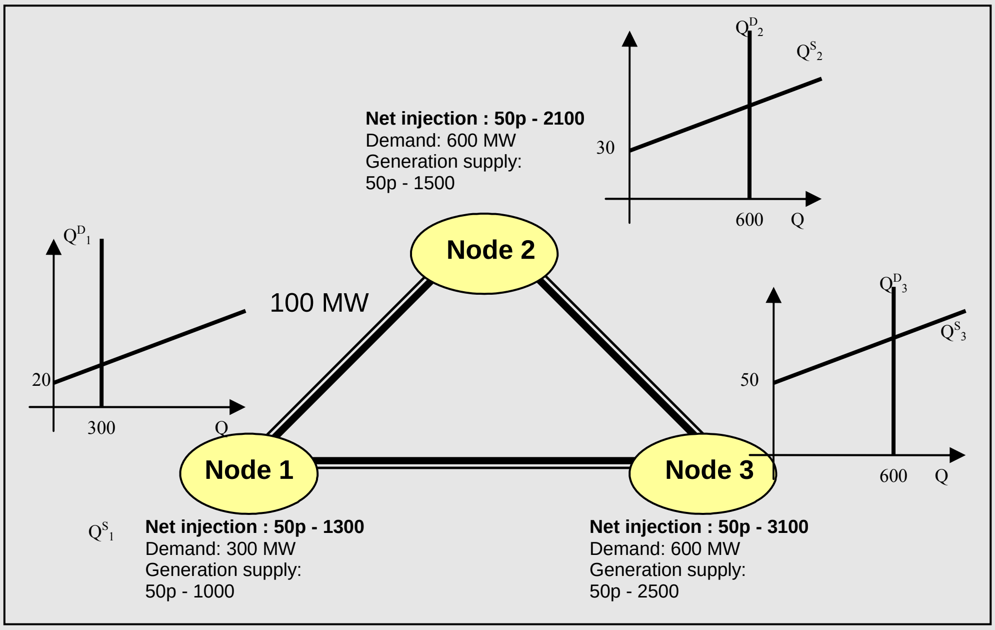

Figure 14 . A three-node transmission grid

As in the example developed in the previous paragraph, we analyze three possible equilibria: first in case of complete separation of the three nodes, then we introduce the network and finally we introduce the thermal limit on one line.

(I) SEPARATED LOCATIONS (AUTARCHY SOLUTION)

Let us first analyze what would happen in case there was complete separation of the three locations (i.e. no line among the nodes).
In that case, each generator would serve its local demand, that is:

$$g_i^A = d_i \qquad \text{for } i=1,2,3$$

and prices would be

$$p_1^A = 20+300/50 = 26 \text{ €/MWh}$$

$$p_2^A = 30+600/50 = 42 \text{ €/MWh}$$

$$p_3^A = 50+600/50 = 62 \text{ €/MWh}$$

(II) LOCATIONS CONNECTED BY LINES WITHOUT CAPACITY LIMITS

As we introduce a transmission grid (suppose without any limit for the moment), the three markets can be substantially considered as one: it is now possible to establish that production is done by the most efficient producers.
Up to 500MW, the producer located in node 1 is the most efficient, so the first 500MW have to be produced by it. Then producers in 1 and in 2 have the same level of efficiency, so it is optimal (i.e. the most efficient production allocation) that they produce the same quantity.
So the remaining 1000 MW will be split equally among the two most efficient producers, (as long as their marginal costs do not exceed the ones by the producers located in node 3, we will show that this is not the case in the present example).
This means that the level of production is:

$$g_1^{UC} = (500+500) \text{ MW} = 1000 \text{ MW}$$

$$g_2^{UC} = 500\ MW$$

$$g_3^{UC} = 0\ MW$$

The price in each node is the same. Remember that the nodal price equals the marginal cost of production of an extra unit of consumption in that node.

$$p_1^{UC} = 20+1000/50=40\ €/MWh$$

$$p_2^{UC} = 30+500/50=40\ €/MWh$$

$$p_3^{UC} = 1/2(40+40)=40\ €/MWh$$

As we pointed out before, it is not socially convenient to ask the producers located in node 3 to produce as they would be more inefficient (remember that the marginal cost function is above 50)

Moreover we can underline that in the present example, that is, in absence of constraint on capacity transmission, the price is the same in all the nodes and the value of transmission rights is null.

(III) LOCATIONS CONNECTED BY A 100-MW LINE

We now introduce a 100 MW thermal constraint on the line connecting node 1 and node 2. This implies that the scheme proposed in the previous section is no more feasible: given the assumption on equal resistance of all lines, we know that the flow on the line connecting nodes 1 and 2 would be equal to 266.6 MW.

This exceeds the thermal constraint on that line, so the social planner would have to revise the scheme proposed in the following direction: the production from the most efficient node has to be partially substituted by the production of the less efficient generators located in node 2.

First of all, the first 300MW produced are going to be loaded in the same node, so this first quantity does not even enter the network and consequently has no impact on the line congestion.

Given this initial production in node 1, the marginal costs of production raise to 26. If the production in node 1 gets larger than 500MW, then the efficiency level of this production node is equal to the efficiency level of node 2; meaning that there will be production in node 2 as well.

If the injection in node 1 exceeds the 300MW, the thermal constraint gets binding and it is then necessary to have 1MW injected in node 2 per MW injected in node 1.

So the generators in node 2 will serve their domestic market (thus, without affecting the network); on top of that, generators in node 1 and 2 will split equally the residual demand (1500-600-600=300 MW).

So the optimal allocation of production is the following:

$$g_1^C = (300+300+150)\ MW$$

$$= 750\ MW$$

$$g_2^C = (600+150)\ MW$$

$$= 750\ MW$$

$$g_3^C = 0\ MW$$

and the associated local prices are

$$p_1^C = 20+750/50 = 35\ €/MWh$$

$$p_2^C = 30+750/50 = 45\ €/MWh$$

$$p_3^C = 1/2\ (35+45) = 40\ €/MWh$$

We can finally check that it is not convenient from an efficiency point of view, to allocate any production to firm 3 as the marginal cost of production using the two other plants (which is the average of the marginal costs of the two generators) is smaller than the marginal cost of producers in node 3.

Notice that the nodal price is equal to the variable cost of production of one extra unit to be sold in that node: in the case of the two efficient nodes (node 1 and node 2), this means that the price equals the local marginal cost of production, while in node 3 this is equal to the average of the marginal costs of production in the two efficient nodes (in fact, if the demand in node 3 grew by 1 MW, given the constraints, the extra production should be allocated half in node 1 and half in node 2).

Concluding, the presence of network externalities and capacity constraints implies a different value for electric power in the different nodes, and thus a different value for the transportation system. We have seen that the decentralized approach would yield the same production allocation as the centralized one, but we have not shown **how** such allocation (and the consequent transmission rights pricing) can be implemented. The next paragraph addresses exactly this question.

2. Mechanism design for allocating transmission rights

In this chapter we will consider the main features of the institutional arrangements implemented to allocate transmission rights: implicit auctions and explicit auctions. In both cases a point-to-point definition of transmission right is adopted (see section 2.2): transmission capacity is (either implicitly or explicitly) defined in terms of possibility to inject in one node and withdraw power from another node.[11]

In section 3.1 we will deal with explicit auctions; in section 3.2 we introduce implicit auctions.

2.1 Transmission rights allocation through explicit auctions

With explicit auctions a central auctioneer receives bids for transmission rights, defined as the right to schedule injections in a node and (simultaneous) withdrawals in a different node. The auctioneer allocates the available transmission rights to the agents who value them the most.[12] Auction proceedings characterize as a rent generated by a scarce resource (transmission capacity).[13]

Explicit auctions have been extensively used to allocate cross-border transmission capacity.

Adopting explicit auctions, the allocation of transmission rights is carried out independently from the process that leads to the selection of the generators who will inject power in and of the customers who will withdraw power from the network. As a result of the transmission capacity allocation some agents are entitled to the right to inject a certain quantity in one location and withdraw the same quantity at a different location. It is then up to each of those agents to select the generators – among those located in the

[11] We are not aware of system in which transmission rights are defined in terms of right to use flowgates, according to the approach sketched in section 3.4.
[12] Auction design issues are beyond the scope of this work.
[13] We will discuss issues about the allocation of the auction proceedings in section XX.

first location – and the customers – among those located in the other location – who will inject and withdraw power. [14]

The separation between transmission rights allocation and injection/withdrawal decisions may cause inefficiencies, if the transmission right holder is not able to procure power in the injection zone and sell it in the withdrawal zone in each moment at market prices. If for example the owner of a transmission right from node 1 to 2 procures and supplies power for the entire duration of the transmission right (say a month) in one go, the corresponding flow might turn out to be inefficient given actual spot prices: in fact in some hours prices in 1 might be higher than in 2, making it profitable (and efficient) for the transmission-right holder to purchase power in 2 (to honor his supply contact) instead of injecting it in 1. [15]

Transmission rights can be traded in secondary markets. The transmission capacity corresponding to un-exercised rights may either remain unused or assigned to other market participants (or allocated via implicit auctions, see next section). The owner of unexercised rights that are reallocated may be entitled either to the value of that capacity (so called *use-it-get-paid-for-it* method) or to no compensation (so called *use-it-or-lose-it* method).

Transmission rights are exercised by notifying the transmission system operator(s). In the France-Italy border, for example, the right-holder notifies the Italian and the French transmission system operators; following notification of, say, an import program of 100 MW in hour t from France into Italy, the French system operator debits the right-holder 100 MW in hour t, for the purpose of the French settlement system; the Italian system operator credits the right-holder 100 MW in hour t, for the purpose of the Italian settlement system. The French "debit" is to be matched by the right holder by

[14] Here we assume for simplicity that injection/withdrawal commitments corresponding to market transactions are always respected, which allows us to disregard the ex-post settlement system, that is a feature of all power markets.

[15] Incidentally observe that the right (for the contracted generator) to inject power in location A and a right (for the contracted customers) to withdraw power from location B result from the purchase and sale transactions, and not from the allocation of transmission rights. To show this, it is sufficient to note that a transaction between a customer and a generator both located at A would also result in the right to inject and to withdraw power in node A, without involving the exercise of any transmission rights.

scheduling 100 MW injections in France; the Italian credit allow the right-holder to schedule 100 MW withdrawals in Italy.[16]

2.2 Transmission rights allocation through implicit auctions

As shown in the previous section in an explicit auction mechanism the allocation of transmission rights is carried out independently from the process that leads to the selection of the generators who will inject power in and of the customers who will withdraw power from the network.

The main feature of implicit allocation mechanisms is that the allocation of transmission rights and the selection of the generators and loads that will be entitled to inject and withdraw power are carried out simultaneously.

An auctioneer[17] collects bids and offers from generators and loads at each location. Bids and offers refer to power injections and withdrawals[18]. A bid (offer) commits the generator (customer) to inject (withdraw) the amount of power specified in exchange for a price greater (smaller) or equal to the bid price. The auctioneer accepts bids and offers in order to maximize the net value of the resulting transactions, under the constraints that the corresponding power flows are compatible with the available transmission capacity. Incidentally we note that the problem solved by the auctioneer is identical to the one solved by the benevolent social planner in Sect. XX. Under the perfect competition assumptions (generators' bids reveal the true variable costs and consumers' offers represent their true willingness to pay), maximization of the net value of transactions by the auctioneer coincides with the maximization of total welfare; this means that the auctioneer is indeed providing the efficient solution.

[16] Failure by the right-holder to schedule injections in France result in a negative imbalance in the French settlement system. Failure by the right-holder to schedule withdrawals in Italy result in a positive imbalance, that is settled according to the Italian settlement rules.

[17] Typically the system operator or a power exchange.

[18] And not, as in the explicit auction mechanism, to the right to schedule a balanced injection/withdrawal program in two locations.

The electricity price at each node – paid to all accepted bids and by all accepted offers[19] – equals the minimum cost to the auctioneer for matching a marginal withdrawal increase in that node[20].

In case there is no binding transmission constraint in equilibrium, a unit increase in the net load in one node can be offset by an increased injection (or load decrease) in any other node; so the minimum cost to compensate a net load increase is the same in all nodes. Hence the market equilibrium electricity price is the same in all nodes.

In the opposite case, the presence of at least one binding transmission constraint introduces a differentiation in the nodal prices. Because of the constraint, a net-load increment in one or more nodes cannot be off-set by a unit injection increase in any other node. Increasing injections of the least cost generator available in the system may generate flows that are incompatible with the network constraint. Because of the transmission constraint, the cheapest way to match the marginal increase in net load in a node could then be by increasing production of a more expensive generator. More generally the least cost way to match a net-load increase in case of congestions involves varying injection levels by a combination of units located in different nodes.

As a byproduct of the auction, the (efficient) use of the network is determined. Since power exported from positive-net-injection nodes receives a lower price than the one paid for power imported into negative-net-injection nodes, the auctioneer collects a rent, which is typically referred to as "congestion-rent".

Next we illustrate the implicit auction mechanism with a very simple two-node example.

[19] Since it is the most widely implemented, we consider a uniform price auction, meaning that all the generators whose offers are accepted will receive the same price: the price of the marginal offer.
[20] That is, the minimum between the as-bid cost of increasing nodal injection by one unit and the cost of decreasing nodal withdrawal by one unit.

EXAMPLE I

Assume that we have two nodes – node 1 and node 2 – that are characterized by the bids and offers represented by the following picture[21]:

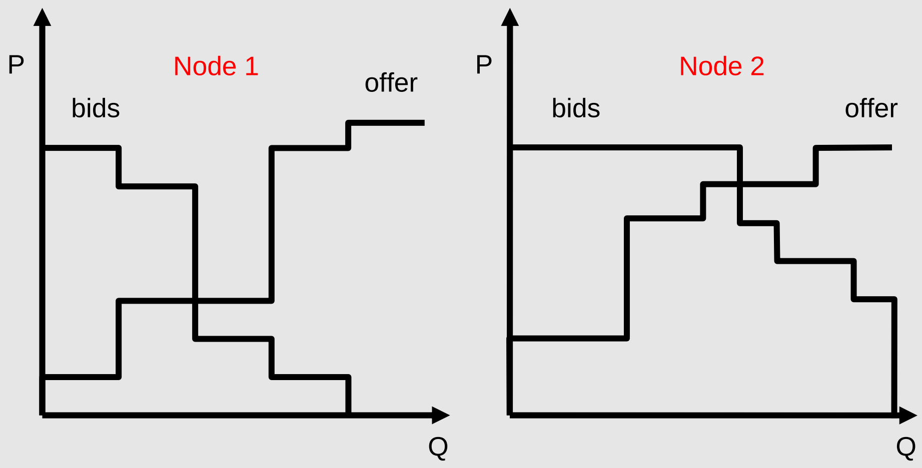

Figure 15 . Bids and offers in the two nodes

Assuming that these are the bids and offers relative to the two nodes, we analyze the three possible cases we have identified in the previous section: completely separated nodes /what we have also denoted as the autarchy solution), nodes connected without (binding) capacity constraints and nodes connected with a (binding) constraint on the line.

(I) SEPARATED NODES (AUTARCHY SOLUTION)

In case the nodes are separated we will have an equilibrium with different quantity and prices:

[21] Notice that we are now relaxing the simplifying assumption of continuous and differentiable demand and supply functions: since we want to deal with the power market where the bids specify a certain quantity at a certain price, demand and supply will be from now on piece-wise functions.

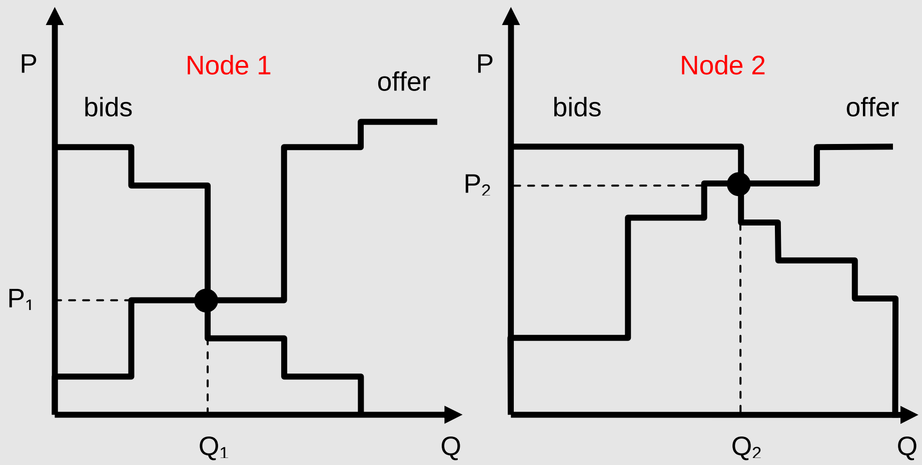

Figure 16 . Autarchy solution

Given the assumptions we did on the shape of demand and supply in the two nodes, power is going to be cheaper in node 1 than in node 2, due to the relative higher efficient of the producers in this node.
This means that the production in node 2 is too expensive for consumers and too costly for producers, with respect to what it would be socially optimum if there was the option to move power from one node to the other.

We note incidentally that net-surplus maximization, in the context of the autarchy solution, is immediately equivalent to the usual representation of the market equilibrium in terms intersection between the demand and the supply functions.

(II) CONNECTED NODES (WITHOUT ANY (BINDING) CONSTRAINT ON THE TRANSMISSION CAPACITY)

In case the two nodes are connected without any (binding) constraint on the transmission lines, it turns out to be efficient to have the firm located in node 1 expanding its production in order to serve part of the demand generated in node 2 which is cheaper if produced by 1 than by 2.

This situation can be represented as an increase in demand in the most efficient node (node 1) and a decrease (for the same quantity) in the demand located in the least efficient node (node 2).

This change in the demand levels is then the part of node 2's demand that is satisfied by node 1's producers. From another perspective, it represents the power flow from node 1 to node 2.

The following picture shows the graphical representation of what we have described:

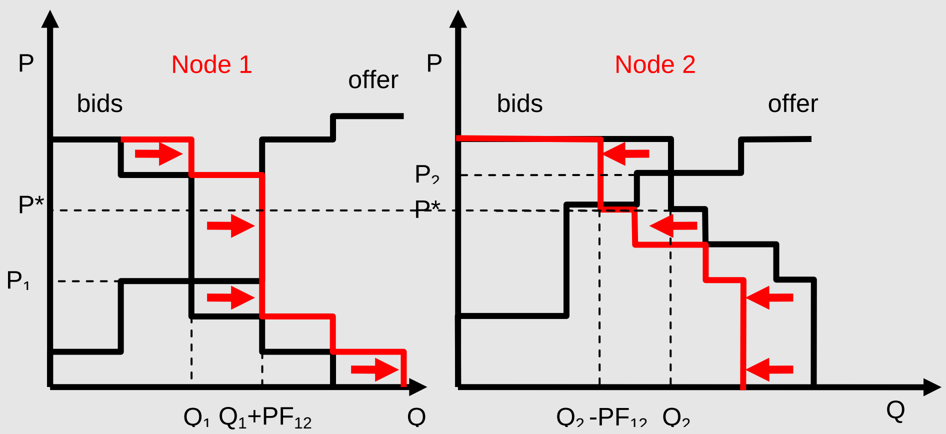

Figure 17 . Equilibrium without binding constraints

In absence of constraints on the line connecting the two nodes (or in presence of a non binding constraint), the equilibrium power flow is the one that makes the prices in the two nodes equal.

In efficiency terms this means that part of the production to satisfy the demand in 2 is allocated in node 1 as in such node we have the most efficient producers.

The criterion, then, is to move production from node 2 to the other node as long as node 1 is more efficient than node 2, i.e. as long as the marginal cost of producing in node 1 is smaller than the marginal cost of producing in node 2.

So production is moved away from 2 to 1 up to the point in which the two marginal costs are equal, or – in other terms – the two prices are the same in the two nodes.

***

The equilibrium can be represented also as the intersection between the net supply function of one node and the net demand function of the other.

Let's consider node 1's net supply function, i.e. the horizontal difference between node 1's supply and demand (i.e. for each price, the difference between supply and demand, in terms of quantity).

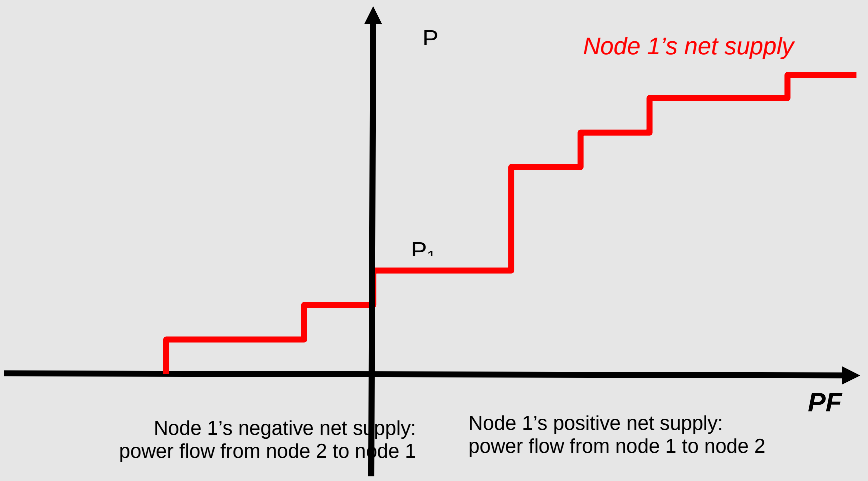

Figure 18 . Node 1's net supply function

Note that the variable on the horizontal axis is power flows from one node to the other; since we are considering node 1's net supply, when this value is positive we are going to have a flow from node 1 to node 2, on the contrary, when it is negative, the power flows from 2 to 1. When it is zero, we are in autarchy (no flow between the two nodes).

It should be clear at this point that the vertical intercept of node 1's net supply function is the autarchy equilibrium price, P_1.

Analogously, we derive the net demand function for node 2 as the horizontal difference between node 2's demand and supply:

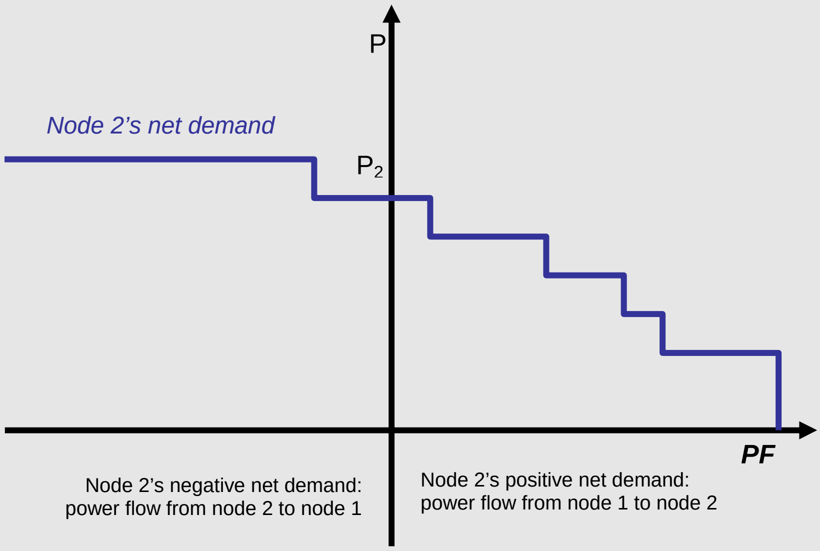

Figure 19 . Node 2's net demand function

In figure 19, the vertical intercept of node 2's net demand function is the autarchy equilibrium price, P_2.

The equilibrium power flow between the two nodes and the resulting price in case there is no binding constraint on transmission capacity, is represented by the intersection of these two curves, as shown in figure 20.

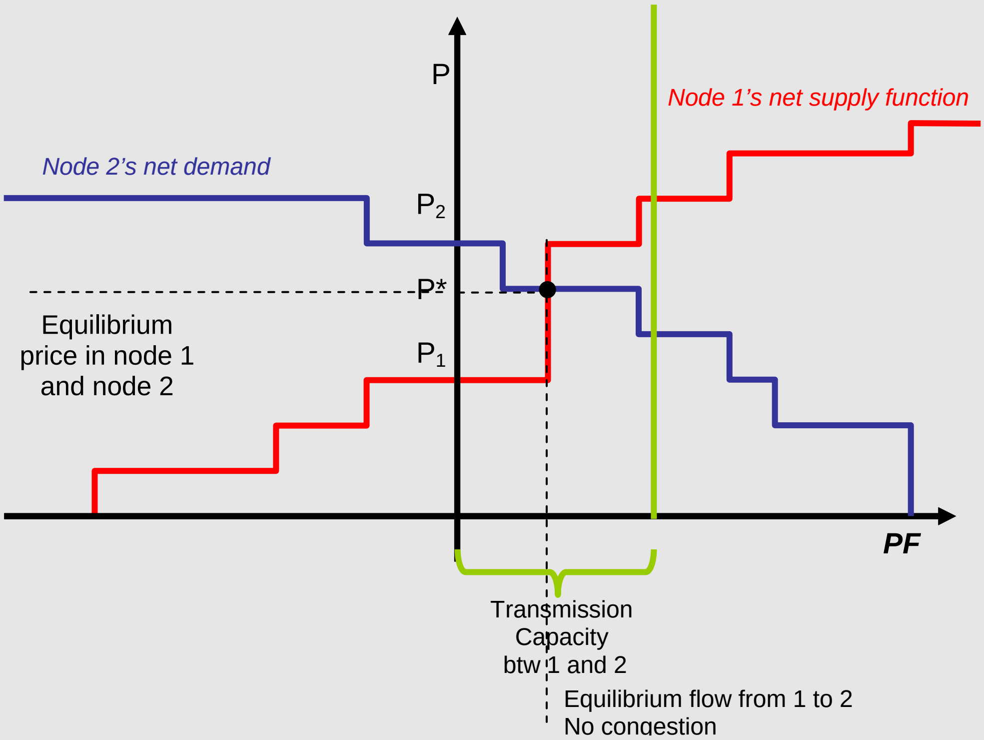

Figure 20 . Equilibrium outcome without congestion

(III) CONNECTED NODES (WITH A (BINDING) CONSTRAINT ON THE TRANSMISSION CAPACITY)

We now modify our example in order to account for possible limits on the transmission capacity: we assume that the transmission capacity between the two nodes is now smaller than the level of power flows that would be optimal to have.

This means that there is an upper bound on the power flows and the allocation of production cannot be the one presented above (which is the most efficient), so in equilibrium we will have two different prices in the two nodes.

Figure 21 shows graphically how the two prices are determined.

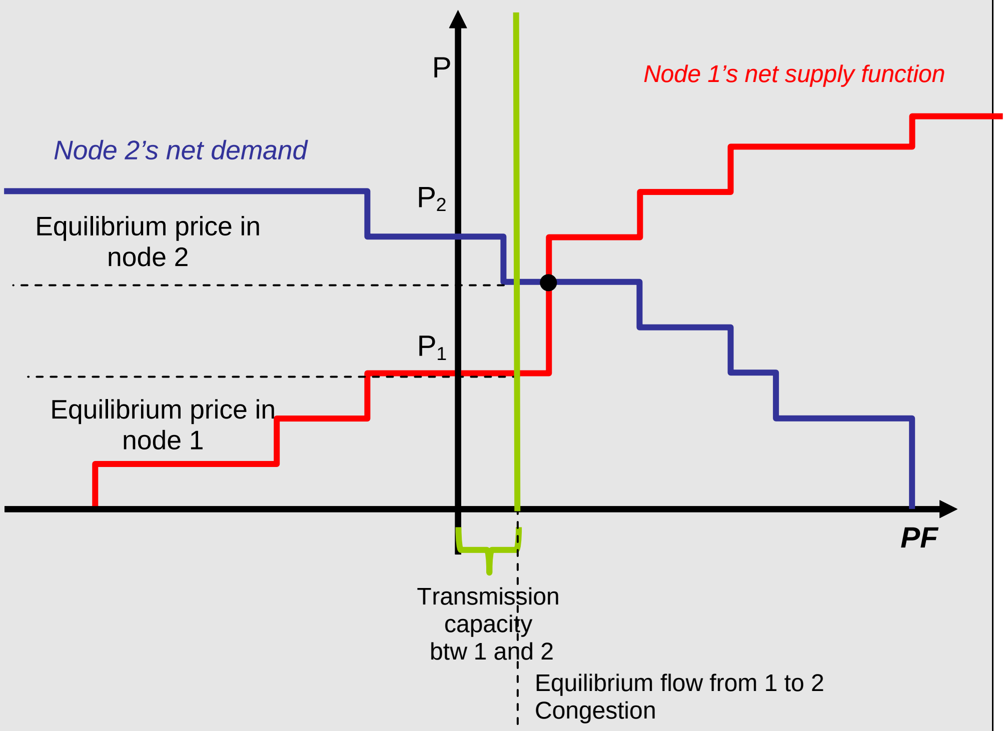

Figure 21 . Equilibrium outcome with congested line

The equilibrium reached in case of a binding constraint on transmission capacity is characterized by a power flow between the two nodes equal to the transmission capacity (by assumption the constraint is binding) and by the presence of two different prices in the two nodes.

In an implicit mechanism for the allocation of transmission rights, the transmission capacity is not actually allocated to any market participants, but it is (efficiently) used by the auctioneer.

The auctioneer collects from the operators the value associated to the use of the available transmission capacity.

To see this, we can come back to the last example:

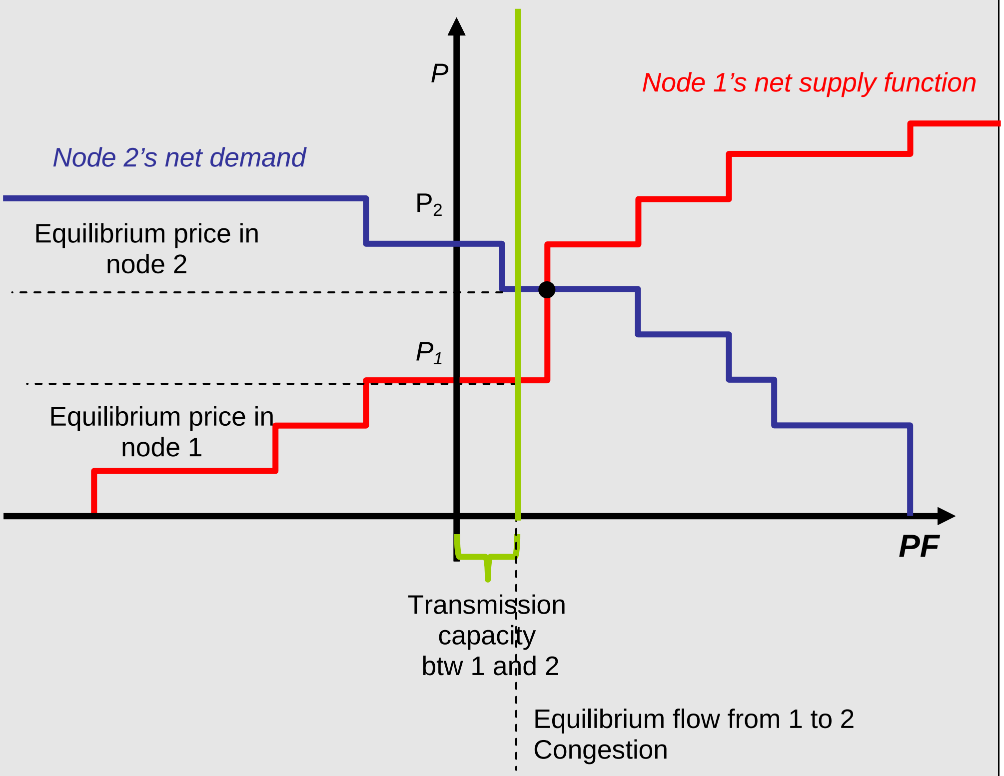

Figure 22 . Equilibrium outcome with congested line

In such an equilibrium we have the producers located in node 1 that receive a price P_1 lower than the price (P_2) paid by the buyers in node 2.
The auctioneer is then collecting P_2 from the buyers located in node 2 for the quantity transmitted from node 1 to node 2 and is paying a price P_1 for buying the same quantity in node 1. Thus, the auctioneer is earning more than what he pays; namely, he has positive profits equal to the difference between the two nodal prices (P_2 - P_1) times the transmission capacity.

This is exactly the value generated by the available capacity, called the congestion rent.

In section 3.3, we will show financial instruments aimed at providing – through an explicit auction – the same (or better) outcomes as the implicit allocation of transmission rights.

***

In the stylized institutional setting we described above, the auctioneer acts (also) as a power exchange. Market participants bid to inject or to withdraw power independently at each node; the auctioneer buys and sells all the power scheduled for injection (and withdrawal). All physical transactions – i.e. all transactions involving a commitment to injecting/withdrawing capacity – are carried out through (only) one power exchange, which also determines the use of the available transmission capacity.

In some countries more power exchanges or scheduling coordinators are active[22]. When multiple scheduling coordinators operate in same the market, they compete for the use of the available transmission capacity.

In a pure implicit-auction setting the efficient allocation of available capacity with multiple scheduling coordinators can be achieved by allowing each schedule coordinator to submit to the auctioneer a net supply curve at each node; exactly as if each schedule coordinator were a generator or a load.

In a mixed implicit-and-explicit auction setting the auctioneer accepts both *i)* bids/offers for net injections at each node and *ii)* bids for point-to-point transmission rights; the latter are bids that, if accepted by the auctioneer, entitle the owner to submit a balanced injection/withdrawal schedule. A rough version of this solution was implemented in the Californian market and is currently implemented in the Italian market. Market participants engaged in a bilateral transaction involving injections and withdrawals at two different nodes are granted priority scheduling right and charged the difference between the price in the withdrawal zone and the price in the injection zone, as resulting in the implicit auction run by the power exchange. This arrangement amounts to allowing only infinite-price bids for point-to-point transmission rights in the mixed implicit-and-explicit auction system.

[22] For the purpose of this analysis a scheduling coordinator is any entity that has the right to submit injections and withdrawal schedules to the system operator; a seller-buyer couple agreeing bilaterally on the price of a power transaction and notifying the system operator the corresponding injections and withdrawal schedules is the smallest conceivable scheduling coordinator; a power exchange is a “large” scheduling coordinator.

2.3 Financial transmission rights

As we pointed out above, the implicit auction mechanism provides an efficient use of the network. As a by-product of the (implicit) auction the auctioneer collects the congestion rent[23].

In the implicit auction mechanism, market participants making transactions between two geographical areas are exposed to the risk associated to the variability of the difference in prices in the injection and load areas. The agents might decide to hedge this risk by buying appropriate financial products.

In order to hedge the risk of volatile power nodal prices, the financial product should entitle its owner to achieve the same economic outcome (or a better one) they would have with physical transmission rights; this is why these financial products are typically called financial transmission rights.

The difference between physical and financial transmission rights stands in the fact that, on the one hand, the physical transmission right entitles its owner to use a fraction of line capacity, by scheduling injection and withdrawals. On the other hand, the financial right is the right to appropriate the congestion rents associated to a certain line. In other words, the physical transmission right is the right to *use* the line, while the financial transmission right is not linked to any actual transmission; there is a separation between the use of the system and the financial benefits associated to this use.

2.3.1 Financial and physical transmission rights

This section is aimed at showing the relationship between a financial transmission right defined in an *implicit-auction setting* and a physical transmission right defined in a *bilateral-contracts setting*.

In the former institutional setting, market operators typically write long term contracts that can be changed in the hourly market on the power exchange. So the program associated to each contract can be reviewed in the day-ahead market.

[23] Already introduced in the previous paragraph, on page 41.

On the contrary, in a bilateral contract setting, long term contracts provide a program that is going to be implemented, whatever happens in the market.

Let's consider an example to underline this difference between these two settings. A long term contract between a generator G located in one node and a consumer C located in different nodes on the provision of Q MW at price P_C assumes different specifications in the two settings. On the one side, in a bilateral-contracts setting, such an agreement prescribes a mandatory production (and injection) by G and a mandatory load (and consumption) by C. On the other side, in an implicit-auction setting, the contract does not imply any obligation to produce or load power for the two parties. What the two parties can do in this setting, is to decide in each hour whether it is convenient to produce (consume) or it is more efficient to let another generator (consumer) produce (load).

This problem is represented in the following picture[24]

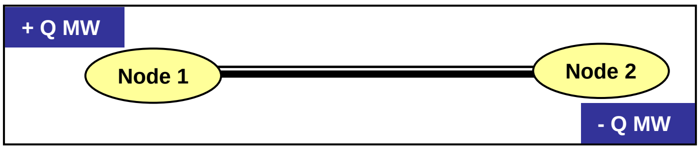

Figure 23 . Definition of transmission rights

It is worth to notice, at this point, that such a contract implicitly implies a problem of efficiency of production that can be seen in two specifications: within the node (is G efficient in producing or is there any other generator who could produce at a lower cost in the same node?) or between the nodes (is there any generator in the other node who could produce at a lower cost?).

In a bilateral-contracts setting, we assume there is no power exchange where power can be traded, so all transactions are physical. Moreover, we assume that there are high transaction costs, so that it is not possible for generator G to look for a more efficient generator to produce in his place. This implies that the injection/load programs defined by the contract are actually implemented, even if they are not the most efficient (and hence, could be renegotiated).

[24] We label node 1, the node where the generator is located and node 2 the consumer's node.

In this setting we assume that generator G holds a physical transmission right, i.e. the right to implement the specific trade associated with the contract without being charged for the use of the line, even in case of congestion[25].

In this case the generator's profit[26] associated to the contract described above is:

$$Profit^{PTR} = \ (P_C - c)\ Q$$

Note that the presence of a physical transmission right allows the generator to use the line needed to implement his program without having to pay for it. As a consequence the generator's profit – gross of the physical transmission right price - depends on the contract variables only.

We now turn to the analysis of the profits associated to the implicit-auction setting. We then assume that the generator holds a financial transmission right for Q MW from node 1 to node 2.

Let us first analyze what generator G is going to make in order to implement the contract. He will make a bid on the power exchange for an injection of Q MW in node 1; this bid is typically assumed to be the variable cost of production, c. This means that, in case the equilibrium market price in node 1 exceeds the generator's bid (cost), then he will be allowed to inject power.

The generator will also make an offer to buy Q MW in node 2; since he has committed to provide Q MW to the consumer, G will make an offer to buy at a very high price, in order to be sure to be accepted on the market.

On the one hand, the generator is going to be sure to buy the power in node 2, on the other hand, it is market equilibrium price that determines the acceptance of the bid in node 1.

In case the equilibrium market price in node 1 is higher than the generator's bid, then the bid will be accepted and the generator will have to inject power in node 1.

[25] We also assume that this physical transmission right – and the following financial transmission right - is an obligation.

[26] Gross of the price for the physical transmission right.

In this case, the generator will produce at variable cost c ($-c$ in the profit function) and he will sell its production on the power exchange for price P_1 ($+P_1$ in the profit function). Then he will buy the power in node 2 at price P_2 ($-P_2$ in the profit function) and deliver it to the consumer in node 2 for the agreed price P_C ($+P_C$ in the profit function). On top of that, he is going to receive congestion rents as well.

So in this second case the generator's profit[27] is:

$$Profit^{FTR} = \underbrace{(P_C - c)\,Q}_{\text{Contract profits}} - \underbrace{(P_2 - P_1)\,Q}_{\text{(paid) congestion rents}} + \underbrace{(P_2 - P_1)\,Q}_{\text{(received) congestion rents}}$$

The generator's profit can be simplified as:

$$Profit^{FTR} = (P_C - c)\,Q$$

which is exactly equal to the generator's profits computed in the previous setting.
So in this first case in which it is efficient for G to produce, the profits associated to a financial transmission right in a implicit-auction setting are equivalent to the profits associated to a physical transmission right in a bilateral-contracts setting.

In case the generator is not efficient, it might happen that the bid made by G in node 1 is higher than the equilibrium price. As a consequence, G's bid is not accepted as it is too high (his efficiency level is too low); this means that G is going not to produce, but he will buy power in node 2 and sell it to the customer.

On top of that he will receive congestion rents associated to that line.

$$Profit^{FTR}\,(P_1 < c) = (P_C - P_2)\,Q + (P_2 - P_1)\,Q$$

[27] As before, not considering the price of the financial transmission right.

Notice that these profits exceed the profits associated with the physical transmission right (which, by the way, do not have the “no production” option); we can show this:

$$Profit^{FTR}\ (P_1 < c) > \ Profit^{PTR}$$

$$(P_C - P_2)\, Q \ + (P_2 - P_1)\, Q > (P_C - c)\ Q$$

$$P_C - P_1 > P_C - c$$

$$c > P_1$$

This last condition is true by assumption.

We can then conclude that financial transmission rights perform better than the physical transmission rights, both in terms of single firm’s profitability and in terms of efficiency of production allocation.

2.3.2 Transmission congestion contract

The typical example of financial transmission right is the transmission congestion contract (TCC)[28]. A TCC is a contract between the transmission system operator and a market operator; it prescribes the transfer of the congestion rents on a certain line on a certain time and on a certain amount of power, from the system operator to the agent.
In other words, it provides a disbursement of the congestion rents by defining a point-to-point contract to collect the difference in locational prices.

So a transmission congestion contract from node 1 to node 2, related to Q MW at time t, is prescribed to pay its owner as follows while it is in effect[29]:

$$TCC^t_{1,2} = Q * (P^t_2 - P^t_1)$$

[28] Introduced by Hogan (1992)
[29] Typically, TCCs are long term contracts and are settled in the day-ahead market; that is, P^t_1 and P^t_2 are day-ahead prices.

Typically a transmission congestion contract is defined over multiple hours, so, given nodal price volatility, the prices in the definition of the contract will change during the time interval defined by the transmission congestion contract.
This implies that the value associated to such a contract is computed as the (discounted) sum of a sequence of fluctuating values.

$$Value\ of\ TCC = \Sigma_t\, \delta^t * Q * (P^t_2 - P^t_1)$$

We remark once more that the value of a transmission congestion contract on a certain line is completely unrelated to any (eventual) power transmission along the path in question: the TCCs are directly connected only to prices and not to power flows. Transmission congestion contracts are not related to any particular wire but only to a starting and an ending node. The lack of connection to anything physical other than these two points gives the TCCs their simplicity.

Given the definition of TCCs, it is straightforward to derive that:

$$TCC_{1,2} = - TCC_{2,1}$$

Notice that the value of transmission congestion contracts might be positive or negative. In the latter case, the contract owner would incur in a loss and would have to pay the value congestion rents to the system operator. This is due to the obligation nature of this kind of contracts, which yields both profits and losses, according to what the market equilibrium outcome is. In order to avoid such (eventual) losses, the market has developed an option type of financial transmission right, which we will deal with in the next paragraph.

2.3.3 Options and obligations

Financial transmission rights can be divided into two categories: options and obligations (as for instance transmission congestion contracts).

What we have analyzed so far is the obligation-type of financial transmission rights, which prescribes a payment to the right holder equal to

$$Q * (P_2 - P_1)$$

In case equilibrium market prices are such that $P_2 < P_1$, we have that power flows from node 2 to node **1**, that is, in the opposite direction with respect to the one associated to the generator's contract. In this case, the right holder has to pay the transmission system operator the difference in prices times the quantity associated to the right.

This implies a loss for the transmission holder. In order to prevent the operator to incur in such a loss another financial instrument has been designed: an option-type financial transmission right. This other transmission right entitles its owner to receive the congestion rents if positive and not to pay anything otherwise. In analytical terms, the prescribed payment of an option-type financial transmission right related to Q units from node 1 to node 2, can be written as:

$$Max\ (\ Q * (P_2 - P_1)\ ,\ 0\ )$$

This way the transmission right holder would be paid in case of a positive difference, but not charged in case the difference is negative.

This second specification of the financial transmission right is obviously more favorable for market participants as it cancels the possibility to incur in losses.

2.4 Flow-gate rights

Chao and Peck 1996 propose a transmission right definition alternative to the point-to-point definition that we discussed in section 2.2.

The two authors define transmission rights in terms of use of key resources, labeled "flow-gates", over which congestion is likely to occur.

There is a substantial difference between flow-gate rights and point-to-point transmission rights: while the latter were defined as the right to inject power in a certain location and withdraw it in another one; the former are defined as the rights to use a physical scarce resource (line, transformer…).

Let's consider the simple example depicted in figure 24.

We have two heterogeneous producers, G_1 and G_2, located in node 1 and 2, respectively. Generator 1 has constant variable costs equal to 25, while generator 2 is less efficient than 1, having variable costs equal to 45. In node 3 there is a consumer, C_3.

Assume that line 1-2 is congested (in the direction from node 1 to node 2).

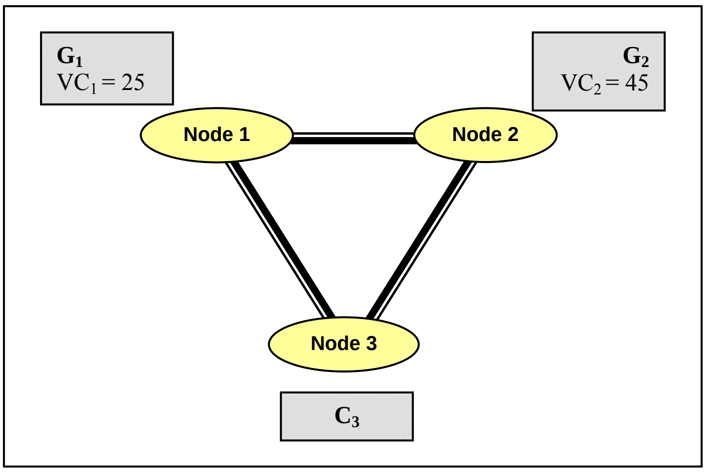

Figure 24 . Flow-gate transmission rights

As we showed in the previous section, according to the point-to-point approach, three nodal prices would be calculated; the two generators selling their production to customers located in node 3 would pay implicitly for the right to use transmission capacity by collecting the nodal price at the injection node and paying Node 3 price for withdrawals;

the system operator would collect the (corresponding) congestion rent and the owners of a transmission right from Node 1 (Node 2) to Node 3 would be entitled to receive (P3-P1)€/MW from the system operator.

Under the flowgate approach, the system operator would acknowledge the fact that line **1-**2 is frequently congested and would then define it as a flowgate. Market participants would then be required an amount of right to use that flowgate consistent with their injections/withdrawals.

In order to assess the relationship between an injection/withdrawal program and the use of a certain line, Power Transmission Distribution Factors (*PTDFs*) are calculated.

Each *PTDF* measures how much power would flow on a certain line if 1 KW were injected at a certain node. Once these factors are known, a market operator willing to implement a program will need to buy an amount of flowgate rights equal to product between his, say, injections at a node and the corresponding *PTDF*.

The three market operators, G_1, G_2 and C_3 will bargain over the price of power and over the price of the flowgate 1-2 (P_{FG12}).

Consider first G_1: in order for that generator to produce profitably, he must recover two different costs:

(i) the production costs (25 €/MWh)

(ii) the cost of the transmission right, the product between the flowgate price, P_{FG12}, times the share of the power injected in Node 1 flowing on Line 1-2, i.e. the *PTDF* of Node 1 for Line 1-2.[30]

The generator located in node 1 will thus produce as long as:

$$P_3 \geq 25 + 1/3\ P_{FG12}$$

Where P_3 is the price paid by the customer located in Node 3.

G_2 also must cover two different costs:

(i) the production costs (45)

[30] In this specific example, we assume equal resistance along all the lines. So the share of the injections in 1 loaded in 3 that flows over the line 1-2 is 1/3. For the derivation of this result, see the examples in section 2.1.2.

(ii) the cost of the transmission right. Unlike $G_{1,}$, the power injected in the system by G_2 relieves the congestion on line 1-2.

The generator injecting in node 2 will thus be entitled to sell flowgate 1-2 rights to generators located in Node 1[31]. The value of flowgate 1-2 rights issued by G_2 for each MW injected in Node 2 is $PTDF * P_{FG12} = (-1/3) * P_{FG12}$. We can conclude that the generator located in node 2 will produce as long as:

$$P_3 \geq 45 - 1/3\, P_{FG12}$$

Competition between the two generators, will drive the price in Node 3 down to marginal cost, so that the two previous conditions will hold with equality:

$$P_3 = 25 + 1/3\, P_{FG12}$$

$$P_3 = 45 - 1/3\, P_{FG12}$$

Thus, we can derive the price for the flow-gate right:

$$25 + 1/3\, P_{FG12} = 45 - 1/3\, P_{FG12}$$

$$2/3\, P_{FG12} = 20$$

$$P_{FG12} = 30$$

So the cost of power in each node is the sum of the value of energy and of the flowgate cost. The value of energy is then 25 €/MWh everywhere across the transmission network; the differences in prices across nodes are caused by the different use of the flowgate that injection/withdrawals at each Node require.

The virtues of the flowgate approach relate to the ability of the mechanism to produce efficient power and transmission prices, in principle in a purely decentralized institutional setting, while fostering liquidity of the forward markets.

[31] Notice that the *PTDF* (on line 1-2) associated with an injection in 2 and a withdrawal in 3 shows a negative value, reflecting the direction of the flow.

On the liquidity issue, compared to the point-to-point definition of transmission rights, the flowgate definition may dramatically reduce the number of rights that market participants need to exchange.

Consider a *n*-node transmission network where *k* lines are expected to be congested (i.e. a grid with *n* nodes and *k* flowgates). On the one hand, under the point-to-point approach, in order to make transactions along the grid, a market for each of the $n * (n-1) / 2$ transmission rights would be required. On the other hand, in the flowgate approach, only *k* flowgate markets would be required. By focusing on the market for the expected congested lines, the Chao-Peck approach concentrates the bargaining activity on few products (flowgates), thus potentially reducing transaction costs.

The merits of this point should be assessed empirically. Chao-Peck (1996) claim that

> *"[a]lthough there is potentially a large number of flowgate rights, the system operation can be simplified further by using a fixed but small (say, up to 10) floating flowgate rights to set scheduling priority."* [32]

Hogan (2000) underlines that the ability to predict the correct number and identity of the flowgates is actually very limited[33] and the number of flowgate rights needed to implement the proposed approach might be quite large.

On implementation, an attractive feature of the flowgate approach is that it could in principle be implemented through completely decentralized trading arrangements, which would make the transmission right market similar to standard input markets.

Without any centralized institution collecting and selecting proposals, several iterations of the bargaining stage would be needed to converge to the efficient outcome. This could give rise to high transaction costs, compared to the implicit auction system implementing point-to-point transmission-rights allocation.

[32] Chao, Peck, Oren, Wilson (2000) p.23

[33] Hogan (2000) reports a very interesting case to illustrate this: "*In one outside study intended to support the development of zonal pricing and decentralized congestion management through something like a flowgate model, a set of 28 constraints were identified as important and analyzed for the variations in the equivalent of a PTDF table. While 28 may seem a large number and difficult to deal with in assembling the capacity rights to use the transmission system, it turned out not to be large enough. In the event, the first six months of operation of locational pricing in PJM found 43 constraints actually binding. Most importantly, none of these actual constraints were in the list of 28 supposedly easy-to-identify flowgates*."

Considering that such bargaining process should be carried out for hourly products, either high transactions costs would arise for the market players or non trivial inefficiencies would result in the decentralized system[34].

Finally the point-to-point and the flowgate systems differ in the risk allocated to market participants. Uncertainty features not only in the determination of the future congested lines, but also in the assessment of *PTDF*s; both these elements are crucial to the assessment of the amount of flowgate rights needed to cover a certain injection/withdrawal schedule. In the point-to-point approach, market operators purchase the right to inject/withdraw power[35] independently of actual network capacity. This means that the cost of any change in transmission capacity lies on the TSO; in case the network cannot support flows corresponding to the amount of transmission rights granted to market participants, the congestion revenues collected by the system operator will be less than his payments to transmission-right holders. On the contrary, in the flowgate approach, market operators are required to purchase the quantity of flowgate rights that they actually use; if for example actual *PTDF*s turn out to be different from those expected ex-ante, market participants will be required by the system operator to settle the difference between the number of flowgate rights they own and the number ex-post necessary to cover their injections/withdrawals. Hence the risk of actual network capacity being different from the ex-ante assessment is placed on market participants.[36].

[34] Chao, Peck, Oren, Wilson (2000) develop an example of iterative power and flowgate negotiation mechanism, cast in a semi-decentralized setting. Even for very small number of nodes and players the number of iterations needed to achieve the equilibrium outcome is not trivial.

[35] Or an equivalent financial right, see section 3.3.1.

[36] This feature of the flowgate approach is not necessary. In fact a flowgate definition of the transmission rights could be implemented without placing risk on market participants. Nevertheless, the use of ex-post PTDFs is part of the original proposal.

3. References

Borenstein S., Bushnell J., Kahn E., and Stoft S. (1996a), "Market Power in California Electricity Markets", *Utilities Policy*, 5(3-4), 1-21.

Bushnell J. and Stoft S. (1996b), "Electric Grid Investment Under a Contract Network Regime", *Journal of Regulatory Economics*, 10, 61-79.

Bushnell J. and Oren S. (1997), "Transmission Pricing in California's Proposed Electricity Market", *Utilities Policy*, 6(3), 237-244.

Bushnell J. and Stoft S. (1996), "Electric Grid Investment Under a Contract Network Regime", *Journal of Regulatory Economics*, 10(1), 61-79.

Chao H. P. and Peck S. (1996), "A Market Mechanism for Electric Power Transmission", *Journal of Regulatory Economics*, 10, 25-29.

Chao H. P. and Peck S. (1997), "An Institutional Design for an Electricity Contract Market with Central Dispatch", *The Energy Journal*, 18, 85-110.

Chao H. P., Peck S., Oren S. and Wilson R. (2000), "Flow Based Transmission Rights and Congestion Management Draft", *The Electricity Journal*, 13(8), 38-58.

Green R. J. (1994), "Do Electricity Transmission Networks need Optimal Spot Prices?" Mimeo, University of Cambridge.

Green R. J. (1997), "Electricity Transmission Pricing - An International Comparison" *Utilities Policy*, 6(3), 177-184.

Green R. J. (1998), "Electricity Transmission Pricing: How much does it cost to get it wrong?", *Power Working Paper PWP-058*, University of California Energy Institute

Harvey S., Hogan W. and Pope S. (1996), "Transmission capacity reservations and transmission congestion contracts", Harvard University.

Hogan W. (1992), "Contract Networks for Electric Power Transmission", *Journal of Regulatory Economics*, 4(3), 211-242.

Hogan W. (1994), "Reshaping the Electricity Industry." presented at *the Federal Energy Bar Conference: Turmoil for the Utilities* (November 17) Washington D.C.

Hogan W. (2000), “Flowgate Rights and Wrongs”, mimeo, Center for Business and Government, JFK School of Government, Harvard University.

Hsu M. (1997), “An Introduction to the Pricing of Electric Power Transmission” *Utilities Policy*, 6(3), 257-270.

Oren S. (1997), "Economic Inefficiency of Passive Transmission Rights in Congested Electricity Systems with Competitive Generation", *The Energy Journal*, 18(1), 63-84.

Stoft S. (1997), “Transmission rights and wrongs”, *The Electricity Journal*, 10(8), 91-95.

Chapter II

Market power in generation

and transmission constraints

1. Introduction

Market power is defined as the ability of a firm to profitably increase market price above the competitive level, by reducing its output or directly raising its price.

In a perfectly competitive market, market power is not present as no single firm, or small group of firms can determine market prices. In such a case, each firm is said to be *price-taker*, meaning that it makes its output decisions taking the price as given, believing that the action it takes cannot affect the market price.
A price taking firm is willing to sell its output as long as the market price is above its marginal costs of production.

The other extreme case is monopoly. A monopolist (or dominant competitor) has the ability to influence the market price through its choice of production level and /or offer price.
By taking such actions, the firm runs the risk to sell less, but it raises the price it will get for all the output it will sell. Notice that in case a firm exercises market power, all firms operating in the market benefit from it. In fact the other firms typically benefit more than the company that is exercising market power since they are able to exploit the price-increase induced by the "active" firm, expanding their production and making more profits (these are known in economics as free-riding profits: the "other" firms free-ride on the price increase induced through the uncompetitive strategy).

There are some features of the market that might limit the firm's ability to profitably raise market price of a certain good in a certain time:

consumers' ability to react to the price increase: in case consumers are able to reduce their consumption of that good, switching it to other times, places or even to other products, then the exercise of market power becomes less profitable.

competitors' reaction to the price increase: in case competitors are able to expand their production in response to an increase in price, then the increase in price will be reduced by their extra production, lowering the profitability of the exercise of market power.

potential competitors' reaction to the price increase: in case other firms, who are not currently on the market, find it profitable to enter the market when they observe the price increase, then they will enter increasing the aggregate production, lowering the profitability of the exercise of market power.

In the present chapter, we briefly recall the reasons why these elements that might mitigate market power are particularly weak in the specific case of the electricity market and we will investigate the specific issue associated with transmission.
First, we will focus on the ability to exercise market power in generation and we will present two alternative methods – physical and economic withholding - that a generator might adopt in order to exercise market power.
Second, we will analyze specific issues related to transmission in electricity markets that are likely to introduce market power, such as the (possible) lines' congestions, the allocation of transmission rights or the type of transmission rights.

2. Market power in electricity generation

As pointed out by many, the electricity industry is particularly prone to the exercise of market power because of the physical peculiarities of electricity and the characteristics of demand and supply of power.
We will now consider each of the three factors limiting market power in the specific case of the electricity market.

1 - Consumers' ability to react: the response of demand to price is often negligible
Let us consider the elements that might explain the responsiveness of demand to price changes.
On the one hand, electricity has a very low level of substitutability with other goods. On the other hand, since electricity is a non storable good, it is not feasible for consumers to buy more electricity when it is cheaper to use it in expensive periods. So consumption of electricity at time *t* cannot be substituted by consumption of electricity at another point in time. (we could label this second element as low intertemporal susbstitutability)
We can conclude that the demand for power is very inelastic, so the ability of consumers to react to an increase in price is very low, facilitating the exercise of market power in this industry.

2 - Competitors' ability to react: supply elasticity
The ability of competitors to react to an increase in price is limited by their capacity constraints, which limit their elasticity of supply.
Given the high variability of demand, there are necessarily times when supply is tight, especially when we move closer to real time. As stated by Borenstein and Bushnell (2000):

"If the grid has only a few percent margin of reserve capacity at that time and there is a producer that is supplying more than a few percent of the total output, then that producer is pivotal in meeting the demand. Put differently, that producer can ask for an extremely high price in order to deliver the power and consumers--more specifically the local utility that represents them in the wholesale power market-- will pay it.”

The other element reducing competitors’ ability to react is related to what we have introduced above: since power supply is inelastic, whenever the other firms are capacity constrained, they are unable to expand their production in response to an increase in price; so the market power of the remaining unconstrained firm is very strong.
Notice that this means that even small suppliers can influence price in these circumstances, avoiding any reaction by bigger competitors. This is somewhat at odds with the Antitrust Authority’s concern with the relationship between big market shares (and concentration) and greater potential for anticompetitive harm. As a consequence the concept of pivotality has been introduced as a market power indicator, integrating (or substituting) the traditional indexes.

3 - Potential competitors’ ability to react: entry

Entry in the electricity market is very difficult both in the short run and in the long run.
In the short run entry from adjacent markets is not very likely because of the little flexibility of plants; in the long run, instead, the entry process is limited by relevant frictions due to investment characteristics, such as lumpiness and risk.

2.1 Methods for market power exercise

In the following we assume the spot market being characterized by a non discriminatory auction in which generators bid their power and consumers make offers to buy. Since the auction is non discriminatory, the benchmark strategy (given by the perfectly competitive outcome) is that generators bid at their variable costs.

There are two main methods of exercising market power in generation: physical withholding of capacity and economic withholding. We now introduce these two methods that, even if they prescribe different strategies, typically lead to the same observable outcome.

2.1.1 Physical withholding

A method to exercise market power in electricity markets is "physical withholding", which refers to the intentional restriction of the power supply placed on the market. The firm can apply this strategy by not bidding part of its capacity either by reporting an outage of a plant or by simply not bidding its entire capacity in the market.

By restricting output, the firm is able to increase the market price. This can be profitable if the loss associated with the missing sale of the quantity withhold is compensated by higher profits on the remaining supplies of energy under its control, which can be sold at a higher price.

There might be different strategies to implement physical withholding in electricity generation. For instance, in the medium run, physical withholding may be the result of a mothballing strategy: generation in a power plant can be stopped and the plant can be kept out of work for a certain period. In the short run, a generator might report a false outage of a plant to the TSO, thus withholding the entire capacity of that plant. Another strategy is to report a longer maintenance schedule than the real one. Finally, a firm might simply not bid all of its capacity. This last strategy is not always feasible, as in many markets it is mandatory to bid the entire capacity, in order to guarantee system security.

The advantage of this strategy with respect to the economic withholding stems from the fact that this is hardly proved by an authority (it is indeed difficult for an authority to show that a plant is not on outage or that it does not need mantainance)

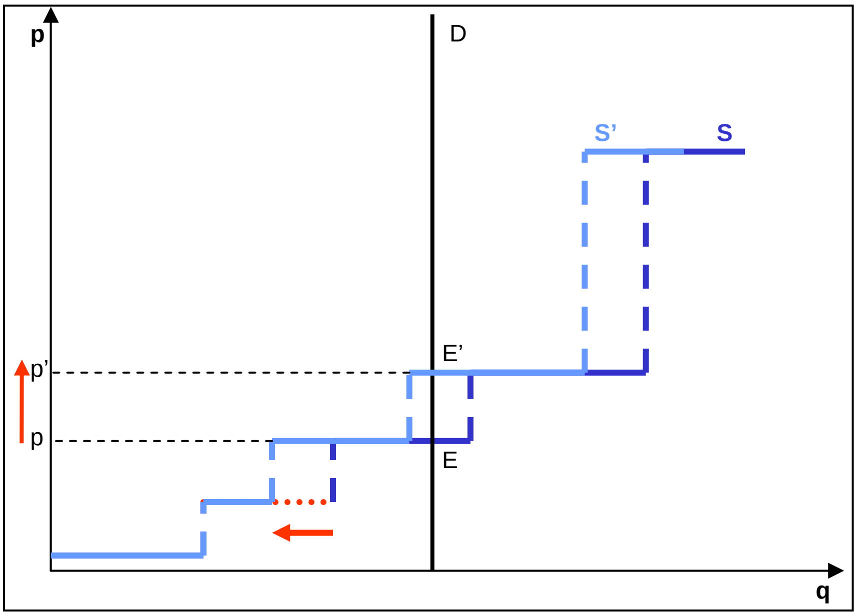

Figure 25 . Physical Withholding

2.1.2 Economic withholding

Another method to exercise market power in electricity markets is called "economic withholding"; it is related to the ability of the generator to enter bids that are substantially higher than the actual cost of production, with the intention of raising the market clearing price.

By setting higher prices, the firm is able to increase the market price, but it is likely that it sells less. As physical withholding, this strategy would be profitable in case the loss in profits associated to the unsold quantity is offset by the positive inframarginal effect on the remaining capacity (that is, the increase in profits due to the fact that the equilibrium price is higher).

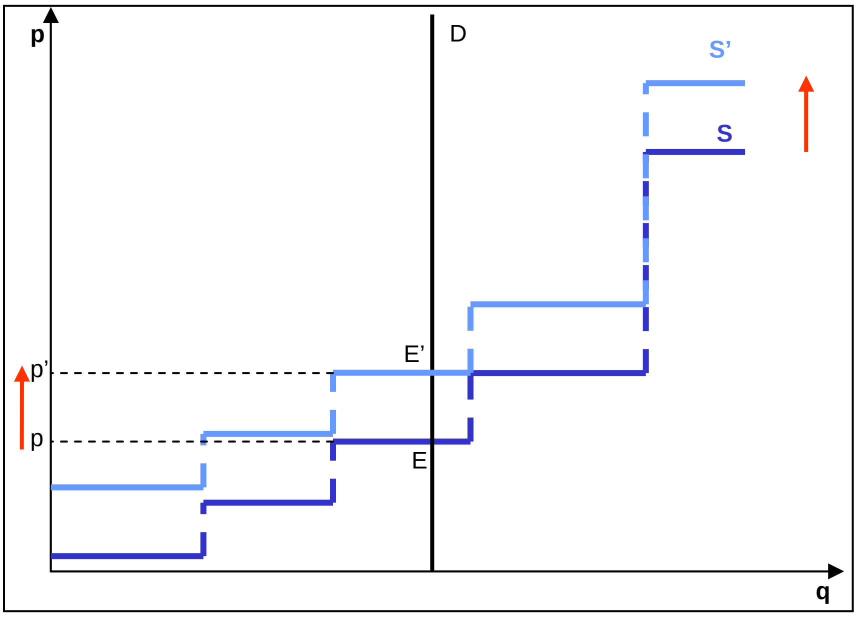

Figure 26 . Economic Withholding

As we introduced before, under perfect competition a non discriminatory auction as this one, leads to bids revealing the true marginal costs of the producers. This result is a lot weaker when we consider pay-as-bid elements in the market design or the option for the power producers to arbitrage among sequential markets.

These elements introduce a distortion with respect to the benchmark in which the bid equals marginal costs and complicates the problem of investigating on the exercise of market power in such a market.

Under this respect, it is impressing the difficulties experienced to investigate whether in the Californian crisis, firms exercised market power.

[Reference to Rajaraman Alvarado 2003 to be added]

3. Power grid and market power

So far we have introduced some characteristics of electricity that make power industry prone to the exercise of market power; in doing so, we have considered only demand and supply characteristics.

In the present paragraph we will make a step further in the analysis: we will introduce the electric grid in the analysis of market power. In the previous chapter we have shown that the presence of a grid allows for an efficient allocation of production: production is allocated where it is cheaper.

The chance of having a connection among markets has a very strong impact on market power: on the one hand, it enlarges the market that can be served by each generator; on the other hand, opening the door to wider competition, it reduces the ability of firms to profitably increase prices.

In order to show the way the network capacity might affect market power, we will introduce very simplified examples, with two nodes. We will first study the equilibria with two separate markets and then we will investigate the effects on the equilibrium quantities and prices of the introduction of a connecting line between the two markets.

The aim of the current paragraph is to underline the competition-enhancing role of transmission due to the substitutability between transmission and generation for demand satisfaction. In the following paragraphs, we will investigate more in detail how competitive strategies depend on the interconnection capacity.

We will first introduce a simplified model with two symmetric nodes characterized by market power (we will analyze the extreme case of two monopolies), solving for the separate equilibrium and then in case of line connection between them. We will then investigate the effects of connecting two nodes when only one is characterized by monopoly while in the other supply is perfectly competitive. Finally we will deal with the effects of the connection of two monopolistic markets characterized by different levels of efficiency.

3.1 Two markets with market power

The first example we provide is very simple: we assume there are two nodes characterized by monopoly, the same production costs and the same demand.

EXAMPLE I.A

Assume we have a two-node network. The two nodes are labeled North and South. Let G_1 be the unique generator located in node North and G_2 be the unique generator located in node South. These generators produce, respectively, q_1 and q_2. Assume for simplicity that generators and consumers in the two nodes are identical[37]. This means that generators have the same production costs and consumers have the same demand function.

In particular, assume that the generators' marginal costs are constant and equal to 4 ($C_i'(q_i) = 4$ where $i = 1, 2$).

We assume generators to compete *à la* Cournot, that is, each firm chooses to produce the level of output that maximizes its profits, given the choice of its competitor(s).

Demand is identical in the two nodes and described by the following linear equation:

$$q_j = 10 - p_j$$

where $j = N, S$.

The following figure represents our stylized network:

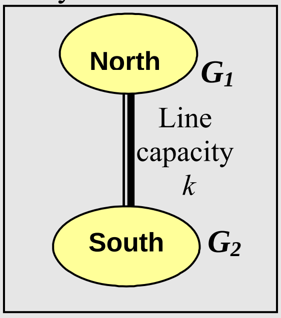

Figure I.A.1 – Two nodes transmission network

[37] We will remove this assumption in the second example, allowing different levels of competition in the two nodes; and in the third example, introducing asymmetry in the efficiency level of the two monopolists.

(i) SEPARATED MARKETS

Let's first characterize the equilibrium in case the two nodes are disjoint, i.e. if $k=0$.

In such a case, there are two identical monopolies in the two nodes. Each monopolist i will choose to provide the quantity that maximizes its profits:

$$\max_{q_i} \left(10 - q_i - 4\right) q_i$$

The first order condition of this maximization problem yields the optimal monopolistic quantity for firm i, $q_i^M = 3$, and the demand function provides the equilibrium price in node j, $p_j^M = 7$.

Profits of each firm are *9*, while consumers' surplus, that is, the area under the demand curve and above the price, is *4.5* in both nodes.

(ii) CONNECTED MARKETS, WITH NO CONGESTION

Let's now investigate the impact of a network on the level of competition.

Since, by assumption, k is so large that there is not going to be any congestion in equilibrium, the two former monopolists behave as two duopolists in a larger market, whose demand is exactly the (horizontal) sum of the two local demands.

The presence of a line between the two nodes allows each firm to sell into the other's node as it sells into its own; hence, the two firms' maximization problem is now depending on the other firm's production decisions:

$$\max_{q_i} \left(10 - \frac{1}{2} q_i - \frac{1}{2} q_z - 4\right) q_i$$

In other terms, the effect of the connecting line between the two monopolies is twofold: on the one hand, this means that each monopolist can sell its product in a wider market; on the other hand each of them has to acknowledge the presence of a competitor when making its production decisions.
The new optimization problem shows that each firm is choosing its output in the wider market in order to maximize its own profits given the choice of its competitor.
The presence of a line between the two nodes has a positive impact on the level of competition, leading to a reduction in prices and an increase in quantity in both nodes.

In equilibrium, each firm will produce a quantity $q_i^D = 4$ and the equilibrium price in both nodes will be $p^D = 6$.
This is associated with single firm's profits equal to *8*, which is a little less than before[38], and a consumer surplus of *8*, which is – as expected –greater than the monopoly case. The overall effect on each single node's total welfare is positive: it grows from *13.5* in the case of separate markets to *16* when they are connected.

In case the two markets are separated each monopolist will choose the quantity that maximizes its profits (monopolistic quantity). The introduction of a line connecting the two nodes as a twofold effect on firms' decisions: on the one hand, it enlarges the demand served by each former monopolist; on the other hand, it introduces a higher level of competition in supply.
This leads to an increase in the quantity supplied by the two former monopolists and a decrease in the equilibrium price in both markets.

[38] The negative effect of having a competitor is partially offset by the fact of having a larger demand to serve.

It is worth noticing that, even if the introduction of the line is effective in fostering competition, **the line is not actually used**.
Since the line introduces the possibility of having competition, each producer facing this new structure of the market will then choose to produce exactly the Cournot duopoly output. A higher level of production would reduce the price and profits; on the other hand, a lower level of production would increase the "domestic" price increasing the firm's profits, but would also induce the competitor to expand its production, vanishing the price-increasing effect of the output contraction.
Thus, although the line is not used, it does provide competition, leading to an output expansion and a consequent lower price (that means a higher consumers' surplus) in both markets.

Although this result that the line is not used depends crucially on the assumptions on firms' and demand symmetry in the two nodes,[39] more in general, it is important to stress that **the effectiveness of a new line is not proportional to the actual flows it exhibits**. Borenstein, Bushnell and Stoft (2000) make exactly this point saying that "there may be no relationship between the effect of a transmission line in spurring competition and the actual electricity that flows on the line in equilibrium". In other words, the role of the line connecting the two nodes is to create a credible threat of competition in each node, but this does not imply that in equilibrium one should observe any flow on that line.

Although the example provided was very simple, it highlighted the importance of the network in fostering competition and mitigating market power. Borenstein, Bushnell and Stoft (2000) assess indeed that "transmission facilities will also provide important competitive links between potentially isolated markets, thus mitigating the potential for market power."

[39] In one of the next examples we will show a modified version of the previous example in which the two monopolists have different marginal costs. In such a case, the equilibrium outcome will imply asymmetric production levels for the two former monopolists, which in turn implies that some power flows on the line in equilibrium.

3.2 One node with market power and a competitive one

We can see a similar effect in a different example in which the line is built between a monopolistic market and a node characterized by perfect competition. Assuming that the congestion will not to occur in equilibrium[40], the result of the connection between the two markets is to have perfect competition in both markets.

EXAMPLE I.B

Let us assume that the producer located in the Southern node is a monopolist as in the previous case, while the Northern node is now characterized by competitive generation.

(i) SEPARATED MARKETS

In case the two nodes are disjoint, i.e. if $k=0$, we have the same monopolistic equilibrium computed in point (a) occurring in the Southern node: $q_2^M = 3$ and $p_2^M = 7$. Profits of firm 2 are *9*, while the consumer surplus is *4.5* in the Southern node.

The Northern node, being perfectly competitive, is going to be characterized by an equilibrium price equal to marginal costs of production; the equilibrium quantity is then determined by demand: $Q_1^C = 6$ and $p_1^C = 4$. Firms have zero profits, while the consumers' surplus is *18* in the Northern node.

(ii) CONNECTED MARKETS, WITH NO CONGESTION

Let's now investigate the impact of a network on the level of competition.
The competitive firms located in the Northern node will be able to expand their production to the Southern node and decrease the price in the South to the marginal cost level.

[40] As anticipated, we will remove this assumption in the next paragraph, analyzing the role of congestions on firms' incentives to exercise market power.

So the new equilibrium is going to be the perfect competition equilibrium outcome in both markets: $Q_j^C = 6$ and $p_j^C = 4$.
All firms have zero profits, while the consumers' surplus is *18* in both nodes.

Notice that the connection between the two markets, opening for (perfect) competition, has a (very) negative impact on the profits of the former monopolist, driving them down to zero.
From a social welfare perspective, this negative effect on profits is offset by the positive effect on consumers' surplus (associated to larger production), so that the overall effect on welfare is positive.
In particular, while total welfare in the competitive node remains unchanged, the Southern node, which had a total welfare of *13.5* (monopolist's profits are *9* and CS is *4.5*) in case of separate nodes, achieves a total welfare of *18* (zero profits and CS equal to *18*) when the two markets are connected.

The connection between the two markets allows the competitive firms located in the Northern node to expand their production in the other node as long as the Southern node price exceeds their marginal costs. This will lead the Southern node to perfect competition.

From a social welfare perspective, there is no change in the Northern node. Regarding the Southern node, we observe that the former monopolist's profits are now null, but this negative effect is offset by the positive impact of the increased competition on consumers' surplus, so that the overall effect on welfare is positive.

3.3 Two asymmetric markets with market power

So far we have only considered the positive effect of the network on the level of competition: through the connection of previously disjoint nodes, consumers can be served by more firms, so competition among suppliers increases, leading to an equilibrium characterized by larger quantities supplied and higher total welfare.

This is not the only effect of the presence of network connecting different production nodes: there is also a potential increase in the **efficiency of production allocation**.
In other words, the presence of an electric grid enables the market to allocate production where it is cheaper.[41]

EXAMPLE I.C

In order to highlight this effect we modify the example I.A introducing a form of asymmetry on the supply side between the two nodes: the monopolist located in the Northern node is more efficient than the other (in particular, we assume it to have marginal costs equal to *2.5*, instead of *4*).

(i) SEPARATED MARKETS

In case the two nodes are disjoint, i.e. if $k=0$, we have the monopolistic equilibrium in both nodes. Given the asymmetry in firm's efficiency, however, the equilibrium price and quantity will differ in the two locations.

The Southern node will be characterized by the equilibrium price and quantity computed the example I.A: $q_2^M = 3$ and $p_2^M = 7$. Monopolist's profits will be equal to *9* and consumers' surplus will be *4.5*.

[41] The assumption on symmetric costs we have done so far, prevented us to see this effect in the previous examples.

In the Northern node, there is a more efficient monopolist who will produce more allowing for a lower equilibrium price. In particular the equilibrium quantity of the efficient monopolist will be $q_1^M = 3.75$ and the equilibrium price will be $p_1^M = 6.25$. So the efficient firm's profits will be higher (namely, *14.0625*) than the profits of the firm located in the Southern node. Consumers' surplus will be higher as well (*7.03125*).

(ii) CONNECTED MARKETS, WITH NO CONGESTION

As we did in the previous examples, we now investigate the impact of a network on the level of competition without considering the congestions. It is worth recalling that the introduction of asymmetric efficiency between firms opens the door to another effect of networks on welfare: production allocation efficiency.

Since we assume that the connection between the two nodes is so wide that it does not represent a constraint for the power transmission between the two nodes, we have that the two former monopolists will now compete *à la* Cournot over aggregate demand.

The inefficient firm will face the same maximization problem as the one in the example I.A:

$$\max_{q_2} \left(10 - \frac{1}{2}q_1 - \frac{1}{2}q_2 - 4\right) q_2$$

On the other side, the efficient firm will have a different target function, given its different marginal costs:

$$\max_{q_1} \left(10 - \frac{1}{2}q_1 - \frac{1}{2}q_2 - 2.5\right) q_1$$

In absence of transmission capacity constraints, the market equilibrium will be an asymmetric Cournot equilibrium in which the efficient firm located in the Northern node, firm 1, produces more than the inefficient one.

Namely, we have that firm 1 will produce $q_1 = 6$, while firm 2 will produce $q_2 = 3$. The resulting equilibrium market price will be $p = 5.5$.
So the profits will be *18* for firm 1 and *4.5* for firm 2, while consumers' surplus in each country will be *10.125*.

Total welfare in case the two markets are separated is *34.59*, while in case they are connected by a line without congestion problems is *42.75*.

The connection of two asymmetric markets has a twofold effect on welfare: on the one hand, it opens up two former monopolies to competition, on the other hand, it allows for a more efficient allocation of production, which has a positive impact on total welfare.
Notice that in this case, we do have an equilibrium flow along the connecting line between the nodes. The equilibrium with the connecting line between the nodes will be characterized by higher quantities, lower prices and an increase in the production of the efficient firm.
All these elements lead to an increase in total welfare.

Summing up, the increase in total welfare associated with the introduction of the network is due to two elements. On the one hand, the fact that the two former monopolists are now operating on the same market leads to an increase in their production; on the other hand, there is an efficient reallocation of production: the increase in production is not symmetric, the most efficient firm will expand its production more than the inefficient one (which in this case will not actually change it).
In other words, the effect of the introduction of the line between the two nodes is not only that existing firms have more competitors than before, but also that production might be reallocated in the most efficient way.

Concluding, the positive contribution of the introduction of a transmission grid on social welfare is twofold: on the one hand, the grid is beneficial as it fosters competition, inducing larger production and lower prices; on the other hand, it allows the market to allocate production where it is cheaper, reducing aggregate production costs and thus enhancing efficiency.

4. Congestions and market power

In the previous paragraph we have underlined the positive effect on competition due to the possibility to connect two separate markets by means of transmission network. Moreover, building a network might have the benefit of leading to a more efficient allocation of production.

However, as we pointed out throughout the previous paragraph, we have made a strong simplification considering networks that were characterized by such large interconnection that no congestion would have occurred in equilibrium.

In fact, networks are characterized by lines that might be congested as consequence of generators bidding output strategies: the scheduled flows on some lines might exceed their capacity (i.e. some lines might be congested). It is worth investigating how this characteristic of power grid affects firms' market power.

The possible restrictions in the network transmission capacity play indeed a relevant role in the analysis of market power. Firms might have the option to expand or reduce their output to induce a congestion on a certain line and reduce competition on a portion of the market.

In other words, if the connection capacity k is not high enough to provide a credible threat of "unlimited" competition, the firm might find it profitable to restrict its output even below the monopolistic level allowing the other firm to enter its market selling k units and then behaving as a monopolist on the residual demand.

With respect to the examples provided in the previous paragraph, the dynamics we are describing now are significantly different: if one of the two firms chooses a strategy that in equilibrium leads to a line congestion, then in equilibrium the two markets are separated and the other firm can behave as a monopolist on the residual demand.

It is easy to see that there might be the incentive to induce transmission congestions in order to isolate some generators, and provide them with a higher market power.

Notice that the fact that the scheduled flows might exceed the lines' capacity and this can cause a separation of some zones, introduces a big difference in the way market power can be exercised in the electricity market with respect to all other markets: while in the conventional setting market power is exercised through output restrictions, in this peculiar market, a generator of electricity might exercise market power by increasing its production in order to induce a congestion and block transmission of a disproportionate amount of competing generation[42].

4.1 Two markets with market power

Borenstein, Bushnell and Stoft (2000) analyze a setup very similar to the one we have introduced in example I.A of the previous paragraph: they assume to have two identical monopolistic markets connected by a line of size k. They investigate the effects of a transmission line on the level of competition by looking at the impact of building a line between two markets that were previously monopolies.

Borenstein, Bushnell and Stoft (2000) derive the best reply functions of the two firms, i.e. each firm's optimal strategy for every strategy of the opponent. In order to define a firm best reply function, three different cases need to be investigated: *(i)* congestion towards the firm's node, *(ii)* no congestion, *(iii)* congestion towards the other firm's node.

EXAMPLE II.A

As in example, I.A we have a two-node network. The two nodes are labeled North and South. G_1 is the monopolist located in node North and G_2 the single generator located in node South.
Generators' marginal costs are constant and equal to 4 ($C_i'(q_i) = 4$ where $i = 1, 2$). Demand is perfectly symmetric in the two nodes and described by the following linear equation:

$$q_j = 10 - p_j$$

where $j = N, S$.

[42] We will show this in two different examples: one with two nodes characterized by asymmetric monopolies; the other in a three node network following the example provided in Cardell, Hitt and Hogan (1997).

(i) SEPARATED MARKETS

The equilibrium in case the two nodes are disjoint coincides with the one we computed in Example I.A: if $k=0$, the equilibrium is characterized in both nodes by monopolistic production, $q_i^M = 3$, and monopolistic price, $p_j^M = 7$. Profits of each firm are *9*, while consumers' surplus is *4.5* in both nodes.

(ii) CONNECTED MARKETS

We now depart from example I.A by considering equilibria potentially characterized by congestions, which we ruled out in that example by assuming a sufficiently high transmission capacity.

The best reply function of each of the two former monopolists depends on the size of the connection, *k*, on the other firm's output and on the fact that a congestion will occur or not (and in which direction).

It is analytically characterized as follows:

$$q_i \equiv BR_i(q_j, k) = \begin{cases} 3 - \dfrac{k}{2} & \text{if} \quad q_i \leq q_j - 2k \\ 6 - \dfrac{q_j}{2} & \text{if} \quad q_j - 2k \leq q_i \leq q_j + 2k \\ 3 + \dfrac{k}{2} & \text{if} \quad q_i \geq q_j + 2k \end{cases}$$

$$= \begin{cases} \textit{optimal passive output} & \textit{if} \quad \textit{congestion towards its own node} \\ \textit{Cournot output} & \textit{if} \quad \textit{no congestion} \\ \textit{optimal aggressive output} & \textit{if} \quad \textit{congestion towards the other node} \end{cases}$$

We now describe this reaction function in grater detail.

(i) OPTIMAL PASSIVE OUTPUT

In case of congestion towards the firm's node, the optimal response for the firm is to act as a monopolist on the residual demand (i.e. the demand minus *k*).
Borenstein, Bushnell and Stoft (2000) label this strategy as ***optimal passive output***.
The firm implementing this strategy finds it more profitable to let the other firm enter its market selling *k* units and then behave as a monopolist on the residual demand. It is intuitive that such a strategy will be implemented when the firm anticipates that the other firm's production is going to be a lot higher than its own (and hence the congestion is going to be towards its own node). Notice that the quantity chosen in such a case is smaller than the monopoly quantity (which is *3*).

(ii) UNCONSTRAINED COURNOT REPLY FUNCTION

In case no congestion occurs, the best reply function coincides with the ***unconstrained Cournot reply function***.
This is the case we have analyzed in the previous paragraph in example I.A: if the transmission capacity is high enough so that no congestion will occur in equilibrium, then the two former monopolists can compete as duopolists for the aggregate market.

(iii) OPTIMAL AGGRESSIVE OUTPUT

In case the congestion is towards the other firm's node, the firm will produce what Borenstein, Bushnell and Stoft (2000) label the ***optimal aggressive output***, which is larger than the monopoly output.
The firm anticipates that the other firm's production is going to be a low; in such a case, it is profitable for the firm to export as much as it can (that is *k*) to the other node, where the price is high due to the low production of the competitor. So the firm will produce the maximum quantity possible, given the transmission capacity, and then it will behave as a monopolist in the domestic node. Notice that the quantity chosen in such a case is larger than the monopoly quantity (which is *3*).

Depending on the size of the connecting line, different equilibria might arise.
Notice that the optimal passive output strategy can be profitable as long as *k* is relatively small: if *k* increases, it increases also the portion of market taken by the (aggressive) competitor; in other terms, residual demand is reduced. So equilibria with aggressive/passive optimal output might occur only if *k* is small.

The condition on *k* is not sufficient for having an equilibrium with aggressive/passive optimal output. In an example as the one we are looking at now, characterized by completely symmetric monopolists, the equilibrium with aggressive/passive optimal output does not occur.
We will show in a subsequent example (Example III.C) that in case of asymmetric efficiency level among firms, there might be an equilibrium in which a firm (the most efficient) is producing the optimal aggressive output and the other is producing the optimal passive output.

For high enough capacity of transmission lines, the Cournot equilibrium will occur. We can now define as k^* the smallest transmission capacity for which we have the Cournot outcome.

It is easy to derive this threshold value, k^*: it is the value of *k* such that playing Cournot yields the same profits as acting as a residual monopolist on domestic demand.
The Cournot profits are *8* and do not depend on *k*, while the profits from letting the other player to export *k* to its node and act as a monopolist on the residual demand are characterized as

$$\left(10 - k - q_i^* - 4\right) q_i^*$$

The optimal q_i^* is

$$q_i^* = \frac{6-k}{2}$$

So the condition for computing k^* is

$$\left(10-k-\frac{6-k}{2}-4\right)\frac{6-k}{2}=8$$

$$\left(\frac{6-k}{2}\right)^2=8$$

$$36 + k^2 - 12\,k = 32$$

$$k^2 - 12\,k + 4 = 0$$

From which we have $k^* = 0.3431$.

So we can conclude that whenever $k > k^*$, the only possible pure strategy equilibrium of this game is the Cournot duopoly equilibrium (by construction).[43]

In case $k < k^*$ instead, the transmission capacity is not high enough to provide a credible threat of competition on each other; so the two firms are not going to play the Cournot reaction functions. However, since the two firms are perfectly symmetric, there is not going to be any equilibrium with aggressive/passive optimal output (they have exactly the same incentives).[44]

We can conclude that a relatively small transmission capacity is needed in order to induce the maximum possible level of competition (which in this case is duopolistic Cournot competition). In order to see this, we can compare k^* – the smallest transmission capacity for which we have the Cournot outcome – with the output expansion that it can yield, or with the size of the market.

[43] See example I.A for the characterization of the Cournot equilibrium in this setting.
[44] In such a case a mixed strategy equilibrium will arise in which firms will play a combination of two pure strategies: monopolistic output and Cournot output.

Following the intuition provided by Borenstein, Bushnell and Stoft (2000), we look at the relative **size** of the connection needed to induce the Cournot duopoly competition with respect to its effects in terms of increased production. We then compare such size of the connection with the market size.

The total increase in output associated to this level of competition is

$$2\ (q^D - q^M) = 2\ (4 - 3) = 2.$$

So the minimum transmission capacity needed for the Cournot equilibrium to occur is only *17%* of the total increase in output in the two markets ($\frac{k^*}{2\left(q^{\mathrm{D}} - q^M\right)} = 0.17$).

Borenstein, Bushnell and Stoft (2000) show that "it is easy to demonstrate that these ratios are the same for any linear demand with constant marginal costs."

Another interesting issue related to this threshold value is that it is fairly low with respect to the size of the market, as measured by the maximum quantity that could be sold in the market with the existing production technology (that is in our case, the intercept of demand, *10*, minus the marginal cost, *4*).

The size of the line needed to lead to the maximum possible level of competition among the two existing firms is less than *6%* of market size.

Although the results obtained in this section are very dependent on the specific assumptions behind the model, we can draw some general results : (i) the effectiveness of the network in terms of increase in competition, output and consumers' surplus; (ii) the relatively small size of the network needed in order to provide the maximum level of competition.

4.2 One node with market power and a competitive one

This section describes an example similar to the one proposed in paragraph 3.2, in which we have analyzed the case of a line built to connect a node characterized by monopolistic generation with a node with perfectly competitive generators (see example I.B); in that example, we assumed the connecting line to have a transmission capacity so high that no congestion occurs in equilibrium: under that scenario, the firms located in the competitive node expand their production to the other node as long as the price exceeds their marginal costs.
So in equilibrium both markets are competitive, prices equal marginal cost, profits are zero and consumers' welfare is maximum.

We now introduce the case in which the connecting line is possibly congested by the scheduled flows: as long as the size of the market[45] in the monopolistic node exceeds the transmission capacity, the competitive firms cannot serve the entire formerly monopolistic market, so the former monopolist has the option of letting the competitive firms export into its node as much as they can (k) and then acting as a monopolist on the residual demand (i.e. on its domestic demand minus the amount k supplied by the competitive firms).
In other words, it has the option of playing the strategy that we have introduced before as *optimal passive output*.

EXAMPLE II.B

Let us assume that the producer located in the Southern node is a monopolist as in the previous case, while the Northern node is now characterized by competitive generation.

[45] As in the previous paragraph, by size of the market we mean the quantity that could be sold in the market with the specific technology in use, if there was perfect competition.
As in the previous case, a market with demand $P = 10 - Q$ and constant marginal cost equal to 4, has size 6.

(i) SEPARATED MARKETS

In case the two nodes are disjoint, i.e. if $k=0$, we have the same monopolistic equilibrium computed in example I.C occurring in the Southern node: $q_2^M = 3$ and $p_2^M = 7$. Profits of firm 2 are *9*, while the consumer surplus is *4.5* in the Southern node.

The Northern node, being perfectly competitive, is going to be characterized by an equilibrium price equal to marginal costs of production; the equilibrium quantity is then determined by demand: $Q_1^C = 6$ and $p_1^C = 4$. Firms have zero profits, while the consumers' surplus is *18* in the Northern node.

(ii) CONNECTED MARKETS

As long as the size of the market in the monopolistic node exceeds the transmission capacity, the firms located in the Northern node cannot serve the entire formerly monopolistic market, so the generator located in the Southern node has the option of letting the competitive firms export into its node as much as they can (k) and then acting as a monopolist on the residual demand. This is the optimal passive output strategy, which in analytical terms, means that the quantity produced by the generator in the Southern node is:

$$q_i\,(k^*) = \frac{6 - k^*}{2}$$

Borenstein, Bushnell and Stoft (2000) investigate the different impact of building a connection on the former monopolist's strategies in two different cases: in case in the other node there is one single competitor (which the case we have analyzed in the previous paragraph), or in case in the other node there are many rivals engaged in a perfect competition (which is the case we are analyzing now).

Their aim is to compare the competitive pressure on the former monopolist induced by the introduction of a single rival with market power with the pressure induced by the connection to a node with perfectly competitive suppliers.

In other words, their question might be interpreted as whether is it more welfare improving building transportation capacity in a setting where it implies introducing competition between two monopolist or between a monopolist and a competitive set of producers.

(iii) COMPARISON OF THE COMPETITIVE EFFECTS

We will now replicate the result by Borenstein, Bushnell and Stoft (2000) in our setting: let's assume, as they do, that the transmission capacity of the connecting line is $k = k^* = 0.34$[46].

CASE 1. CONNECTION WITH ANOTHER MONOPOLISTIC NODE

We have already analyzed the firm's decisions in case two monopolistic nodes are connected: the Cournot equilibrium will occur, with $q_i^D = 4$, $p^D = 6$, single firm's profits 8 and consumers' surplus 8.

CASE 2. CONNECTION WITH A COMPETITIVE NODE

In the opposite case, the former monopolist is connected to a node characterized by perfectly competitive supply. In such a case, facing the competition of the competitive firms, it will act as a residual monopolist on the demand reduced by k^*, that is, it will play the optimal passive output associated to $k = k^*$:

[46] They assume the connection capacity to be exactly the minimum capacity needed to yield the Cournot outcome in the first case, that is k^* as we defined in the previous paragraph.

$$q_i\,(k^*) = \frac{6-k^*}{2} = 2.83$$

COMPARISON

In order to compare the two cases, Borenstein, Bushnell and Stoft (2000) look at the increase in the quantity consumed in the formerly monopolistic node.
In the first case, that is, when the monopolistic node is connected with another monopolistic node, the quantity sold in the node under analysis increases from *3* to *4*. In the other case, when the former monopolist is connected with a perfectly competitive node, the quantity has a much smaller increase: from *3* to *3.17*.

In the comparison between the two cases (connection with a monopolistic node and connection with a competitive node), Borenstein, Bushnell and Stoft (2000) conclude that the increase in quantity induced by the competitive pressure of one single competitor is stronger than the pressure exercised by competitive players.

It is important, however, to underline that this result depends crucially on the assumption on the level of transmission capacity that they do.
Borenstein, Bushnell and Stoft (2000) assume transmission capacity to be exactly equal to the minimum level of transmission capacity that yields Cournot equilibrium in case two monopolies are connected. However, assuming a small connection capacity implies that profits in case an optimal passive output strategy is played are low.
Relaxing this assumption the results change: if $k>2$, the result is exactly the opposite. In such a case, the production in the formerly monopolistic node increases more if the node is connected with a competitive one, than if connected with another monopolistic node[47].

[47] Notice that k does not need to be equal to market size (which is *6* as computed before) in order to make the second case more welfare improving than the first.

4.3 Two asymmetric markets with market power

In the analysis of the equilibria in paragraph 4.1, we underlined that for $k < k^*$ there are only mixed strategy equilibria in the case of symmetric firms, while if $k > k^*$, the Cournot equilibrium occurs.

If firms are asymmetric, however, a pure strategy equilibrium might occur even in case $k < k^*$: for a sufficient level of cost heterogeneity, an equilibrium in which the efficient firm is playing the optimal aggressive output strategy and the inefficient one plays the optimal passive output strategy[48] can occur.

EXAMPLE II.C

As in the example I.C, we assume that the monopolist located in the Northern node has marginal costs equal to *2.5*, while the monopolist located in the Southern node has marginal costs equal to *4*.

(i) SEPARATED MARKETS

The equilibrium prices and quantities in case the two nodes are disjoint ($k=0$) have been computed in the example I.C: in the Southern node, the inefficient monopolist produces $q_2^M = 3$ and the nodal price is $p_2^M = 7$, while in the Northern node, the efficient monopolist produces $q_1^M = 3.75$ and the nodal price is $p_1^M = 6.25$.

(ii) CONNECTED MARKETS

We now depart from example I.C by assuming a connection potentially congested.

[48] In such a case, the line is congested.

The two reaction functions are analogous to that introduced in the previous paragraph: the best reply function of the inefficient firm (the one located in the South) is

$$BR_2(q_1,k)=\begin{cases}3-\dfrac{k}{2} & if \qquad q_2\le q_1-2k\\ 6-\dfrac{q_1}{2} & if \quad q_1-2k\le q_2\le q_1+2k\\ 3+\dfrac{k}{2} & if \qquad q_2\ge q_1+2k\end{cases}$$

$$=\begin{cases}\text{optimal passive output} & \text{if} \quad \text{congestion towards its own node}\\ \text{Cournot output} & \text{if} \quad \text{no congestion}\\ \text{optimal aggressive output} & \text{if} \quad \text{congestion towards the other node}\end{cases}$$

The efficient firm has an analogous best reply function; the differences are due to the different marginal costs.

$$BR_1(q_2,k)=\begin{cases}3.75-\dfrac{k}{2} & if \qquad q_1\le q_2-2k\\ 7.5-\dfrac{q_2}{2} & if \quad q_2-2k\le q_1\le q_2+2k\\ 3.75+\dfrac{k}{2} & if \qquad q_1\ge q_2+2k\end{cases}$$

It is straightforward to derive that congestion towards the Northern node cannot occur; the intuition is quite simple: the monopolist located in that node is more efficient, that is, it is able to produce at lower costs, so the price in its node will be lower than the prices in the other node. This implies that any flow on the connecting line will be in the opposite direction.

If the line has a capacity smaller than *0.75*, then the equilibrium will be characterized by a line congested from the cheap generation (Northern node) to the Southern node.

In such a case, the market equilibrium will depart from the unconstrained Cournot equilibrium and will be characterized by a passive/aggressive optimal output.

- k is small enough that the inefficient firm will find it profitable to let the other firm export k into the Southern node to satisfy part of the domestic demand and will act as a monopolist on the residual demand.
- On the other side, the efficient firm will react to the price increase in the South associated to the output reduction by exporting as much as it can (that is, k) to that node, inducing a congestion towards South.

Assuming, for example, that $k=0.5$. In equilibrium, the efficient firm will produce the optimal aggressive output, $q_1^{M+} = 4$, while the inefficient firm located in the Southern node, will produce the optimal passive output, $q_2^{M-} =2.75$.
Prices will differ in the two nodes[49] will then be $p_N = 6.5$ in the Northern node and $p_S = 6.75$ in the Southern.
The aggregate welfare of the two nodes increases as a consequence of the higher competition and the efficient reallocation of production.

In this case, where the limited connecting capacity is built between two nodes characterized by asymmetric efficiency levels, the equilibrium characterized by a passive/aggressive optimal output strategy might occur.

It is interesting to notice that the efficient producer located in the Northern node anticipates the opportunity to sell k "units" at its domestic price, without an impact on domestic demand.
In other words, the quantity that the efficient producer is going to sell will be sold at the domestic price, but since it is going to be an export, this does not reduce the price that local consumption is willing to pay. The result is that this efficient (former) monopolist

[49] As we have seen in the previous chapter, if the line between two nodes is congested the nodal prices diverge.

will produce as much as it can for the other node (the price effect on the other node's demand is affecting the other generator only); in the domestic market it will restrict production with respect to the monopolistic level, in order to get a higher price for all its production (that is also for the *k* units that are not affecting domestic demand).
This implies that the price in the Northern node increases after the line is built.
Looking at total welfare in the Northern node, the impact of building a line between the nodes is positive anyway: the positive effect on firm's profits is so high that it more than compensate for the loss in consumers' surplus. The efficient generator has more market power than before (it is able to raise prices above the monopoly level) and it can expand production.

***

We can conclude that the presence of a transmission grid, fostering competition, reduces market power in these markets; moreover it enhances welfare through a more efficient production allocation as we have underlined in the previous examples.
However, the presence of possible congestion over the grid introduces a new source of market power: by expanding its production a generator might be able to separate its node from the other and set a price even higher than the monopoly price.

5. Loop flows and market power

The previous paragraphs have emphasized the importance of the transmission grid in determining market power in electricity markets. In the previous analysis, however, we have not considered the presence of loop flow. The simplified examples of a single wire network adopted for highlighting the different features of the transmission grid, could not introduce this element.

In this paragraph we introduce two main examples that show the relevance of the loop flows in affecting not only the presence of market power, but also the way it might be exercised.

We proceed as follows: first, we introduce an example similar to the one provided by Joskow and Tirole (2000), showing that there might be an increase in market power associated to the presence of a congestion; we will also show the interpretation provided by Joskow and Tirole (2000) on the relationships among competitors' output in that particular case.

Second, we will show an example based on Cardell, Hitt and Hogan (1997), in which the congestions induce an increase in market power which can be exercised by an output expansion instead of an output contraction.

5.1 Strategic substitutes or strategic complements?

Joskow and Tirole (2000) analyzed a very simplified three-node model with two generation nodes and one consumption node. They assumed a thermal limit on the line between the two injection nodes, showing that in such a case, the production in one node would be a strategic complement to the other node's production.

This change in the relationship between injection in the two nodes has an impact on market power, as it changes the effectiveness of the output contraction on consumption prices, as it will be clear in the example below.

We will now present a simplified example which replicates their results and is coherent with the analysis we have done so far.

EXAMPLE III.A

Consider the simple power market represented in the following figure:

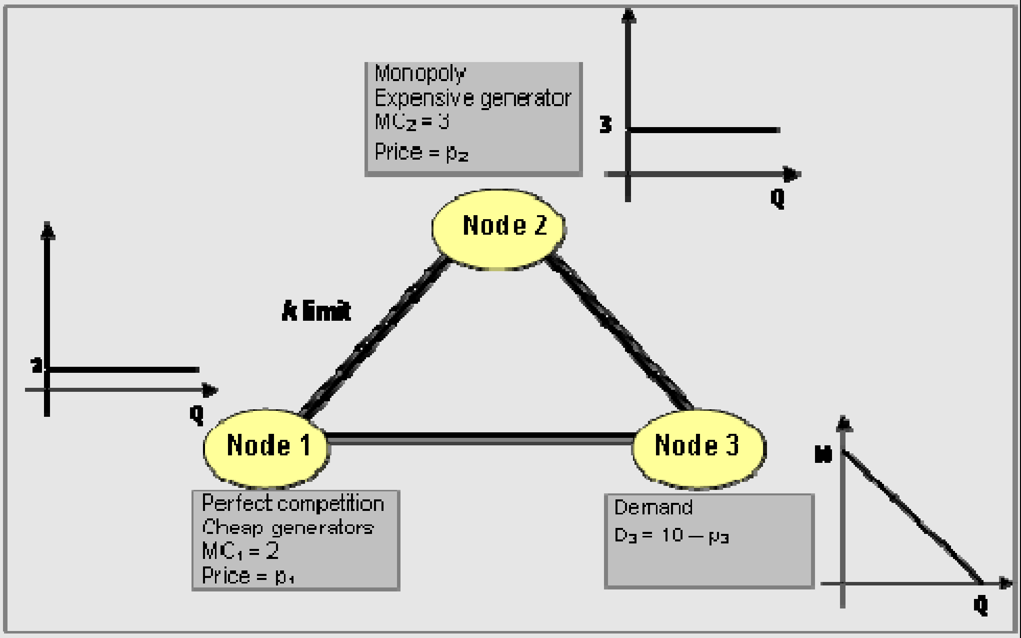

Figure III.A.1 – Three-node transmission network

The market consists of three nodes (1, 2 and 3) and three lines.

For simplicity, we assume each line to have the same electrical characteristics, but for one feature: the transmission capacity of the line connecting node 1 and node 2 is constrained at *k*.

Let k be equal to 2, for simplicity, so that we can compute the model results avoiding this parameter.

Demand is concentrated at node 3, while power is supplied both at node 1 and node 2[50].

Generators in node 1 behave competitively and are more efficient than the single generator in node 2.

In particular we assume that the monopolist located in node 2 has constant marginal costs $c_2 = 3$, while the competitive set of generators in node 1 are characterized by lower constant marginal costs: $c_1 = 2$.

Demand in the consumption node (node 3) is assumed to be linear as described by the following function:

$$D_3 = 10 - p_3$$

***

BENCHMARK CASE: NO CONGESTION

In case the transmission limits are so high, that no congestion occurs in equilibrium, the equilibrium is characterized as follows. The competitive producers located in node 1 bid at their marginal costs and satisfy the demand, while the production supplied by the monopolist located in node 2 is zero.

Summing up, in the benchmark case we have:

[50] Alternatively, one might interpret these as net demand and net supplies.

$q_1 = 8$ $p_1 = 2$

$q_2 = 0$ $p_2 = 3$

$D_3 = 8$ $p_3 = 2$

***

SOLUTION WITH MARKET POWER

We now relax the assumption of no congestion adopted to compute the benchmark and we compute the solution with market power[51] and possible equilibrium congestion.

The equilibrium must satisfy some constraints: on the one hand demand must be met by supply, i.e. in analytical terms,

$$D_3 = q_1 + q_2; \tag{1}$$

on the other hand, the transmission capacity constraint on line 1-2 must not be violated, i.e. in analytical terms,

$$| q_1 - q_2 | \leq 3k. \tag{2}$$

Using the fact that there is a lower marginal cost and hence a larger production in node 1 than in node 2 and since congestion will occur in equilibrium, we can rewrite the second constraint (eq. (2)) as:

$$q_1 - q_2 = 3k = 6 \tag{2'}$$

The price in the consumption node is:

$$p_3 = ½ (p_1 + p_2). \tag{3}$$

Given the assumption on perfectly competitive behavior in node 1, we can assume that power is supplied at marginal cost in that node, i.e. $p_1 = c_1 = 2$.

[51] Assuming perfectly competitive supply located in node 1, monopolistic generation located in node 2 and demand located in node 3.

On the other hand, this condition does not hold for the single generator located in node 2, who is able to sell at a price different from its marginal cost.

The generator in node 2 is going to maximize its profits subject to constraints (1), (2') and (3). From equation (3), we can derive the value of p_2 as a function of the other prices:

$$p_2 = 2p_3 - p_1$$

Since node 1 is perfectly competitive,

$$p_2 = 2p_3 - 2$$

From the constraint on quantities (eq. (2')), we know that

$$q_1 = 6 + q_2$$

From the demand function and eq. (1), we know that

$$p_3 = 10 - q_1 - q_2$$

$$= 10 - (6 + q\,) - q_2$$

$$= 4 - 2\,q_2$$

Substituting this equation for p_3 in the condition on prices, we find the relationship between the quantity set by the monopolist and the price it gets:

$$p_2 = 2p_3 - 2$$

$$= 8 - 4\,q_2 - 2$$

$$= 6 - 4\,q_2$$

So the maximization problem of the monopolist located in node 2 is:

$$\max_{q_2}(p_2(q_2) - c_2)q_2$$
$$\max_{q_2}(6 - 4q_2 - 3)q_2$$

The FOC for this problem is:

$$3 - 8q_2 = 0$$

From which we have the equilibrium quantities and price in the monopolistic node:

$$q_2 = \frac{3}{8} = 0.375$$

$$p_2 = 6 - 4\,q_2$$

$$= 6 - 1.5 = 4.5$$

Producers in the competitive node can produce only

$$q_1 = 6 + q_2$$

$$= 6.375$$

And the consumption in node 3 will be

$$D_3 = 6.375 + 0.375 =$$

$$= 6.75$$

Summing up, in presence of market power, the equilibrium prices and quantities are:

$$q'_1 = 6.375 < q_1 \qquad p'_1 = 2 = p_1$$

$$q'_2 = 0.375 > q_2 \qquad p'_2 = 4.5 > p_2$$

$$D'_3 = 6.75 < D_3 \qquad p'_3 = 3.25 > p_3$$

It is important to notice how the effectiveness of output contraction is affected by the presence of a transmission constraint: while in a standard monopoly the impact of an output contraction on prices is 1:1, in presence of a transmission constraint as the one described above, the impact on consumption prices is doubled (this is straightforward looking at the way we have rewritten the demand function, when including the transmission constraint: $p_3 = 4 - \underline{\mathbf{2}}\, q_2$).

This is due to the fact that contraction in output in node 2 induces an equal reduction of output in node 1, because of the presence of the transmission constraint on the line connecting the two producing nodes.

In this simplified model, we have shown that the presence of transmission constraints on the power grid has dramatic effects not only on the degree of market power, but also on the strategic relationship among output in different places. The latter argument is necessary to understand the way market power is enhanced.
As Joskow and Tirole (2000) point out, in a setting in which a binding transmission constraint is set between two generating nodes, the output in node 1 and output in node 2 instead of being strategic substitutes, become two *local complements*, where "local" refers to the fact that the result holds for relatively low values of k. A reduction in the output in one node forces an equal reduction in output in the other node.

This implies that the effectiveness of output contraction on consumption prices is doubled: by reducing its output of one unit, the monopolist is able to force a reduction by one unit from the competitors because of the transmission constraint. This means that consumption is reduced of two units, so the monopolist can get the price increase associated with the withholding of two units by reducing output of one unit only.

This result cannot be generalized to any constraint on the transmission line: in order for the production of two nodes to be local complements, the constraint must be placed on the interconnection between them or, more in general on a line over which the power flows injected by the two regions have opposite directions.

5.2 Output contraction or expansion?

Loop flows might affect the way market power can be exercised in another way; under some conditions, a firm might be able to exercise market power through an output expansion (instead of the usual means of contracting output).

Cardell, Hitt and Hogan (1997) analyze a three-node model with generation in all nodes and one single consumption node. Generation is assumed to be competitive in one node, while a generator with market power owns a plant in each of the two other nodes (one of which is the one where demand is located). They assumed a binding transmission constraint on the line between the nodes where the generator with market power has its plants. They show that in such a case, the monopolist is going to exercise market power expanding its production with respect to the competitive levels.

EXAMPLE III.B

Consider the simple power market represented in the following figure:

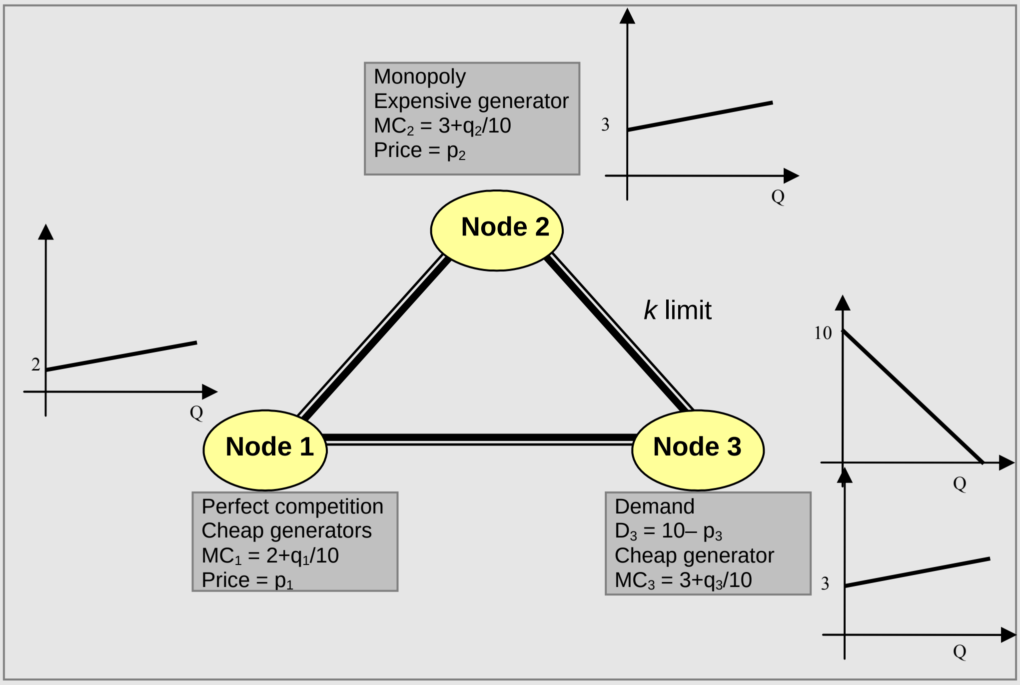

Figure III.B.1 – Three-node transmission network

The market consists of three nodes (1, 2 and 3) and three lines.
For simplicity we assume each line to have the same electrical characteristics, but for one feature: the transmission capacity of the line connecting node 2 and node 3 is constrained at *k*. Let *k* be equal to 1, for simplicity, so that we can compute the model results avoiding this parameter.
Demand is concentrated at node 3, while power is supplied in all nodes.
Generators in node 1 behave competitively and have linear marginal costs described by the following function:

$$c_1 = 2 + \frac{q_1}{10}$$

The generation in nodes 2 and 3 is controlled by a single monopolist, which has linear marginal costs in both nodes:

$$c_2 = 3 + \frac{q_2}{10}$$

$$c_3 = 3 + \frac{q_3}{10}$$

Demand in the consumption node (node 3) is assumed to be linear as described by the following function:

$$D_3 = 10 - p_3$$

***

COMPETITIVE BENCHMARK

In the competitive benchmark, all the firms will set their bids at their marginal cost.

The quantity accepted in node 1 will be the maximum feasible, given the constraint on line 2-3:

$$q_1 = 3k = 3$$

This way the line 2-3 will be congested, so the plant in node 2 (which is less efficient and so will bid its capacity at a higher price) will not inject anything in the grid.

The plant located in node 3, however, is not affected by the line congestion as its injection does not need to pass through the grid. That plant will bid at its marginal costs and sell:

$$q_3 = 40/11 \approx 3.64$$

The aggregate quantity will be

$$D_3 = 73/11 \approx 6.64$$

And the benchmark equilibrium price in node 3 (which coincides with the marginal cost of firm 3 to produce q_3) will be

$$p_3 = 37/11 \approx 3.36$$

Summing up, in the benchmark case we have:

$q_1 = 3$	$p_1 = 2.3$
$q_2 = 0$	
$q_3 = 3.64$	
$D_3 = 6.64$	$p_3 = 3.36$

***

SOLUTION WITH MARKET POWER

We now relax the assumption of competitive behavior to solve our model with perfectly competitive supply in node 1, monopolistic ownership of generation in node 2 and in node 3 and demand in node 3.

In this case we cannot apply the same solution method, used in the previous example: in example III.A, the monopolist was sure that its bid would have been accepted, even if at a high price, as any quantity injected in node 2 would have relaxed the constraint and allowed the cheap injection in node 1.
As we pointed out discussing that example, this is due to the fact that the constraint was exactly on the line connecting the two producing nodes: this makes the injection in one node a local complement to the injection in the other node (instead of having the standard relationship of substitution).

In the current case, however, the constraint is on a line along which the flows injected in node 1 and 2 have the same direction, so the relationship of substitutability is even stronger than in the standard case.
It is important to notice this peculiar feature of this kind of network.

The monopolist knows that its production in node 2 has a double impact on congesting the line with respect to the production that the competitive fringe can inject in node 1: given the Kirkoff's laws that we quoted in the previous chapter, we know that the flow on the line 2-3 will be $\frac{2}{3}q_2 + \frac{1}{3}q_1$, so the constraint given by the thermal limit on line 2-3 can be represented as:

$$\frac{2}{3}q_2 + \frac{1}{3}q_1 \leq 3$$

This enables the monopolist to reduce the aggregate injection in the grid (q_1 + q_2) through an output expansion.
Such a reduction in the aggregate injections means a lower import level in the consumption node and hence a higher residual demand for the plant located in that node.

So the monopolist, owning a plant both in node 2 and in node 3, might decide to bid in node 2 undercutting the bids in node 1 (that means also bidding under its costs); this way, it could reduce the others' production and the aggregate injection in the grid. This would lead to a higher residual demand in node 3 for the monopolist itself.
In the limit, the monopolist will prevent any injection from the others and congest the grid by undercutting the bids from node 1. Applying this strategy, the monopolist will produce exactly half the quantity injected in the grid in the benchmark case. This means that the residual demand in node 3 will be $P_3 = 8.5 - q_3$ instead of $P_3 = 7 - q_3$, which was the residual demand in the benchmark case.

So the monopolist has to decide whether (*i*) to let the competitive generators use (and congest) the grid, injecting a large quantity, 3k, in the grid and produce in node 3 only; or (*ii*) to undercut partially the bid of the competitive producers, bidding under costs, thus reducing aggregate grid injections and increasing the residual demand in node 3; or in the limit (*iii*) to undercut the bid of the competitors located in node 1 producing 2/3 k, thus congesting the grid and enjoy the increased residual demand in node 3.

In case the monopolist does not undercut the bids by producers in 1 (i.e. case (*i*)), then they will act as in the benchmark case, setting $q_1 = 3$ and $p_1 = 2.3$. The monopolist is not selling anything in node 2 and maximizes its profits on the residual demand in node 3 setting $q_3 = \frac{20}{11} \approx 1.82$.

Summing up, in this first case, in which the monopolist does not produce in node 2, we have:

$q_1 = 3$ $\quad\quad\quad$ $p_1 = 2.3$

$q_2 = 0$

$q_3 = 1.82$

$D_3 = 4.82$ $\quad\quad\quad$ $p_3 = 5.18$

And the profits of the monopolist are $\Pi_M = \dfrac{40}{11} \approx 3.64$

Anyway, the monopolist can set a better strategy, by undercutting the bids of the firms located in node 1 and increasing the residual demand it faces in node 3. This strategy has a cost as it prescribes a bid smaller than the costs in node 2. However it turns out that monopolist's profits increase as q_2 increases.

In the extreme case in which the monopolist decides to bid its maximum production (3/2 k) at a price slightly smaller than the price made in node 1, the following happens.

The producers in node 1 won't be able to sell their production on the market, while the production in node 2 will be $q_2 = 3/2 = 1.5$ at price $p_1 = 2$.

The monopolist will then maximize its profits on the residual demand in node 3:

$$\max_{q_3} \quad \left(p_3(q_1, q_2, q_3) - c_3(q_3)\right) q_3$$

$$\max_{q_3} \quad \left(10 - 1.5 - q_3 - 3 - \frac{1}{10} q_3\right) q_3$$

$$\max_{q_3} \quad \left(5.5 - \frac{11}{10} q_3\right) q_3$$

The FOC for this problem is:

$$5.5 - \frac{11}{5} q_3 = 0 .$$

From which we have that the optimal quantity for the monopolist, assuming that it is also congesting the grid by producing 1.5 in node 2, is to produce

$$q_3 = \frac{5}{2} = 2.5$$

The profits associated with such a strategy will be:

$$\begin{aligned}
\Pi_M &= \left(p_2 - c_2(\bar{q}_2)\right)\bar{q}_2 + \left(p_3(q_1, \bar{q}_2, q_3) - c_3(q_3)\right) q_3 = \\
&= \left(2 - 3 - \frac{1.5}{10}\right) 1.5 + \left(10 - 1.5 - 2.5 - 3 - \frac{2.5}{10}\right) 2.5 = \\
&= \underbrace{-1.725}_{\substack{\text{Losses from} \\ \text{the sales under} \\ \text{costs in node 2}}} + \underbrace{6.875}_{\substack{\text{Profits from} \\ \text{the sales} \\ \text{in node 3}}} = 5.15
\end{aligned}$$

So we can conclude that the best strategy for the monopolist is to produce the amount needed to congest the grid and bid it at a price lower than the competitive bid of the producers located in node 1. This way the monopolist is able to prevent the others from injecting power into the grid; this, in turn, allows it to have the minimum import possible in the consumption node.
So the monopolist is bidding under cost in node 2 in order to raise the price it can make in node 3.

This example was built in order to replicate the results introduced by Cardell, Hitt and Hogan (1997): under some hypothesis[52], the generators can exercise market power by expanding output rather than contracting it.
This result is based on the peculiar relationship between the production in node 1 and in node 2, as described by the constraint on line 2-3:

$$\frac{2}{3}q_2 + \frac{1}{3}q_1 \leq 3$$

The constraint on line 2-3 introduces an exclusionary relationship between the injection in nodes 1 and 2, that – unlike the example by Joskow and Tirole (2000) – leaves the substitution relationship between them and, more interestingly, leads to a situation in which the expansion of production by the monopolist in 1 induces a "double contraction" of the output in node 2.
In a setup as the one proposed by Cardell, Hitt and Hogan (1997), the monopolist will incur in losses selling part of its production under cost in order to increase the price it can get in the consumption node.
Notice that the price increase is not benefited by the plant that is actually expanding output, but by the other plants belonging to the same producer and located in the consumption node.

[52]In particular, there are two basic assumptions that are necessary for the result to hold: (i) the congestion must be on a line along which the flows of all the injections have the same directions and (ii) the monopolist must hold a plant in the consumption node, in order to benefit from the price increase induced by the strategy described above.

In other words, the monopolist can exploit the fact that the line might be congested for exercising market power in a more effective way.

In the equilibrium we have found, the monopolist expands its production in node 2, selling power under its production costs; by doing so, the monopolist is able to prevent any injection from the competitive firms located in node 1 and to reduce the total quantity "imported" by node 3.[53]

This has a positive effect on the price in node 3; the monopolist will enjoy this (costly) reduction in "imported" power, in terms of higher prices for the sales in node 3.

[53] This is due to the particular constraint described above: the maximum quantity that the competitive firms located in node 1 could inject is 9, while the monopolist can induce a congestion on line 2-3 by injecting only half of that quantity.

6. Transmission rights and market power

Another important feature in the analysis of market power in electricity markets is the presence of transmission rights. Joskow and Tirole (2000) analyze the impact of different transmission rights' allocations on the level of market power.

Firstly, in line with the result we presented in the previous chapter, Joskow and Tirole show that, assuming the existence of "use-it-or-lose-it" conditions for electricity generators that own transmission rights, there is an **equivalence** between financial transmission contracts and physical transmission contracts. For this reason, in the present paragraph we are going to treat the problem of the way market power is affected by the allocation of transmission rights, without considering their type.

6.1 Two asymmetric markets with market power

In this section, we analyse the impact on market power of different allocations of transmission rights. This section describes the basic setup, while the next sections analyse the effects of allocating the transmission rights to the importing or to the exporting generator.

We will show that market power is enhanced if transmission rights are allocated to importing generators, while in the opposite case, that is, if transmission rights are held by exporting generators, we have a mitigation of market power.

EXAMPLE IV.A (SETUP)

In order to show the impact on market power of alternative allocations of transmission rights, we adopt the asymmetric setting introduced in paragraph 3.3: we assume a two-node network, where the two nodes are labeled North and South.

Let G_1 be the single efficient generator located in node North and G_2 be the single inefficient generator located in node South. These generators produce, respectively, q_1 and q_2. Assume that G_1 has constant marginal costs equal to 2.5, while G_2 has constant marginal costs equal to 4 ($C_1'(q_1) = 2.5$ where $C_2'(q_2) = 4$).

Demand is perfectly symmetric in the two nodes and described by the following linear equation:

$$q_j = 10 - p_j$$

where $j = N, S$.

The two nodes are connected by a line with transmission capacity k.

Following the approach of the previous sections, we can distinguish what happens in case the nodes are disjoint, in case they are connected with a large capacity line, so that no congestion occurs in equilibrium and in case they are connected by a possibly congested line.

(i) SEPARATED MARKETS

As we have seen in example I.C, if the two nodes are separated (i.e. $k = 0$), the two monopolists choose to produce the monopolistic output, by equalizing marginal revenues and marginal costs.

In the Northern node, the former monopolist G_1 produces 0.75 and the market price is 3.25, while in the Southern node, G_2 produces 3 and the price is 7.

(ii) CONNECTED MARKETS, WITH NO CONGESTION

As derived in example I.C, if the two markets are connected by a large capacity transmission line, the Cournot equilibrium with asymmetric costs will occur.

Under the unconstrained Cournot equilibrium, the efficient producer located in the Northern node G_1 produces *6* and the inefficient generator G_2 produces *3* and the single market price is *5.5*.

(iii) CONNECTED MARKETS, WITH CONGESTION

In both these cases, since congestion does not occur in equilibrium, the transmission rights are not going to pay back any congestion rent, so they do not impact on firms' decisions, regardless of the way they are allocated.

The relevant case for analyzing the role of transmission rights allocation is when the two markets are connected by a relatively small transmission line[54]. In such a case, the equilibrium will be characterized by aggressive/passive output and congestion will occur in equilibrium.

As we have seen in example II.C, the efficient generator, G_1, is now enabled to sell into the Southern node at the equilibrium price of the domestic node[55]. Since the domestic price is not affected by this exporting quantity, there are two effects of the connection on the monopolist decisions:
(i) G_1 is willing to reduce domestic output with respect to the monopolistic one in order to increase the price it will earn both on domestic and exported quantities;
(ii) it is willing to expand exports as much as it can, as these quantities do not affect domestic market price (so G_1 is willing to use all the connection capacity between the nodes). But exports are constrained not to exceed *k*.

[54] In analytical terms, this means that $k < k^* = 0.75$, we look at this case as this is the parameter region where congestion occurs in equilibrium and thus the transmission rights play a role.
[55] The opposite case, that is, G_2 exporting to the Northern zone is implicitly ruled out by the fact that the Northern market price is always smaller than the Southern one.

As cited in example II.C, Borenstein, Bushnell and Stoft (2000) label *optimal aggressive output* the optimal strategy in this case: the efficient generator, G_1, will export as much as it can (k) to the other node and will restrict output in the domestic node.
The optimal behavior for the inefficient monopolist located in the South, is to choose *optimal passive output*, that is, to act as a monopolist on the residual demand.

So the equilibrium in case the two nodes are connected is characterized by G_1 that produces a smaller quantity in the domestic market, inducing a price increase, and exports k to the Southern market and by G_2 that reacts to the import of k units by reducing its domestic production.

Analytically, G_1 produces $0.75 + k/2$, the domestic sales in the Northern node are $0.75 - k/2$ (less than the monopolistic output, which is 0.75), and the price will be $3.25 + k/2$; in the Southern node, G_2 produces the optimal passive output, $3 - k/2$ (less than the monopolistic output, which is 3), and the equilibrium price is $7 - k/2$.

The key question is now whether and how the allocation of transmission rights to any of the two agents affects their market power.

6.1.1 Transmission rights allocated to the exporting generator

Let's first consider how the market power of the exporting generator is affected by the allocation of transmission rights to such an agent.
We first analyze how the transmission rights change the incentives looking at their impact on the profit function and then we discuss the intuition behind the result.

EXAMPLE IV.A (SECTION A)

We now consider the case where the transmission rights are allocated to the generator located in the Northern exporting node.

Assume that G_1 holds a generic share α of these rights; in such a case its objective function becomes:

$$\max_{q_1} \left(p_N(q_1,k) - 2.5\right)q_1 + \alpha\, k\left(p_S - p_N(q_1,k)\right)$$

$$\max_{q_1} \underbrace{(10 - q_1 + k - 2.5)q_1}_{\substack{\text{Profits from producing } q_1 \\ \text{sold at the domestic price,} \\ 10\text{-}(q_1\text{- k}), \text{ and produced} \\ \text{with marginal cost 2.5.}}} + \underbrace{\alpha\, k\left(p_S - 10 + q_1 - k\right)}_{\substack{\text{Congestion rents of} \\ \text{a share } \alpha \text{ of} \\ \text{transmission rights.}}}$$

The first term in the objective function is the firm's profits from the sale of q_1, that is, the objective function it would have in case no transmission rights are allocated to the exporting generator.

Notice that the price is determined solely by the quantity consumed in the Northern node, $q_1 - k$, but it also applies to the quantity exported to the Southern node (the net price is multiplied by q_1).

The second component of the objective function represents the revenues associated with the transmission rights held by the former monopolist. A transmission right over a certain line entitles its owner to receive the corresponding share of the associated congestion rents, that is the share α of the transmission capacity, k, times the difference in the two nodes' price, $p^S - p^N$.

G_1's choice of the total quantity to produce affects both terms, in other words, on top of the usual effect of quantity choice on production profits, there is an effect on the congestion rents.

It is straightforward to see that while the profits from the sale of q_1 increase in the price of the Northern node, the second component is actually decreasing in that price.

This means that the allocation of transmission rights to the exporting generator is **mitigating** G_1's **market power**, by providing an incentive to reduce its domestic price.

The effect of transmission rights is to allow the exporting generator to sell part of its capacity at the other node's price. This mitigates the incentives to increase the domestic price.
As α increases, this market power mitigation effect gets stronger: the fraction of output sold at the Northern market price decreases and this in turn lowers the incentive to increase the price in the North over the monopolistic price. In the extreme case that the monopolist located in the North is assigned all the transmission rights over the line (i.e. $\alpha = 1$), then it can sell all the export capacity at the Southern node's price (and not at the Northern price); it is like it can perfectly price discriminates between the two nodes. In such a case, the domestic price and quantity will be back to the monopolistic ones.

Analytically, this is straightforward: if $\alpha = 1$, the objective function becomes:

$$\max_{q_1} \left(p_N(q_1-k)-2.5\right)(q_1-k)+\left(p_N(q_1-k)-2.5\right)k+k\left(p_S-p_N(q_1-k)\right)$$
$$\max_{q_1} \left(p_N(q_1-k)-2.5\right)(q_1-k)+\left(p_S-2.5\right)k$$

The exported units, k, are sold at the Southern price, while the maximization related to the quantity sold in the domestic node, $q_1 - k$, is the standard monopolistic problem.

So we can conclude that assigning transmission rights to a generator with market power located in an exporting node has procompetitive effects: since the generator is earning a congestion rent inversely proportional to the price it sets in the domestic market, it has an incentive to keep this price low.

In other words, as the generator raises the domestic price, it loses money on the transmission rights; this means that

Thus, the allocation of transmission rights to the generator located in the exporting node, reduces its incentives to increase the domestic price, lowering the level of market power.

6.1.2 Transmission rights allocated to the importing generator

In case the transmission rights are held by the generator located in the Southern importing node, the effect on market power is exactly the opposite.

EXAMPLE IV.A (SECTION B)

Let the importing generator, G_2, hold a generic share α of the transmission rights; G_2's objective function becomes:

$$\max_{q_2} \left(p_S(q_2,k)-4\right)q_2 + \alpha\, k\left(p_S(q_2,k)-p_N\right)$$

$$\max_{q_2} \underbrace{\left(10-q_2-k-4\right)q_2}_{\substack{\text{Profits from producing } q_2 \\ \text{sold at the domestic price,} \\ 10-q_2-k\text{, and produced} \\ \text{with marginal cost 4.}}} + \underbrace{\alpha\, k\left(10-q_2-k-p_N\right)}_{\substack{\text{Congestion rents} \\ \text{of a share } \alpha \text{ of transmission rights.}}}$$

The first component of G_2's objective function is the firm's profits from the sale of q_2 considering the residual demand. This is the objective function of the importing generators, absent transmission rights.
The second element of the objective function is the revenues of the transmission rights held by the importing generator.

As in the previous example, the quantity produced by the right holder, G_2, affects both terms of the profit function.

In this case, however, the congestion rents component is proportional to the domestic price, so the incentives to keep a high domestic price are reinforced by the allocation of transmission rights.

This effect is increasing in α, the share of transmission rights held by the importing generator: as α increases, the total volume of energy it sells at the Southern node price increases.
In the extreme case that the monopolist located in the South is assigned all the transmission rights over the line (i.e. $\alpha = 1$), then it is earning the Southern price both on the quantity it is producing and on the quantity imported from the Northern node. It is like it can sell the imports as well.
In such a case, the domestic price will be back to the monopolistic one.

***

We can conclude from the analysis of these simple examples that if a generator with market power at an importing node is allocated transmission rights over the congested line, this will be equivalent to an increase the total volume of energy it sells at the Southern node price and increases its incentive to withhold domestic output to increase the domestic price.
Market power is therefore enhanced if importing generators hold transmission rights.

On the contrary, if an exporting generator with market power holds transmission rights, it is as if the fraction sold at the domestic market price decreases and hence the incentive to raise the domestic price decrease as well (Green, 1999).

Transmission rights held by exporting generators mitigate market power.

In the present exposition we have adapted the examples from Joskow and Tirole (2000) to our setting in order to show their main results; another interesting result they have is that considering a three-node network, they show that, in such a case, **transmission rights held by exporting generators can enhance market power.**

7. References

Borenstein S. (2000), "Understanding Competitive Pricing and Market Power in Wholesale Electricity Markets", *The Electricity Journal,* 13 (6), pp. 49-57.

Borenstein S., Bushnell J., Kahn E. and Stoft S. (1996), "Market power in California 's Electricity Market", *Utilities PolicyI,* 5, 1-12.

Borenstein S., Bushnell J., Kahn E. and Stoft S. (2000), "The competitive effects of transmission capacity in a deregulated electricity industry", *RAND Journal of Economics*, 31 (2), 294-325.

Byggar D. (2005), "Competition Issues in the Electricity Sector", *OECD Journal of Competition Law and Policy*, Vol. 6, No. 4.

Cardell J., Hitt C. and Hogan W. (1997), "Market Power and Strategic Interaction in Electricity Networks," Resource and Energy Economics 19:1, pp. 109 - 137.

Gilbert R., Neuhoff K. and Newbery D. (2004), "Allocating Transmission to Mitigate Market Power in Electricity Networks", *RAND Journal of Economics*, 35 (4), 691-709.

Hogan W. (1997), "A Market Power Model with Strategic Interaction in Electricity Networks", *The Energy Journal*, 18 (4), 107-141.

Joskow P. and Tirole J. (2000), "Transmission Rights and Market Power on Electric Power Networks", *RAND Journal of Economics*, Vol 31 (3), 450-487.

Oren S. (1997), "Economic inefficiency of passive Transmission Rights in Congested Electricity Systems with competitive Generation", *The Energy Journal*, 18 (1), 63-83.

Pierce R. J. (1996), "Antitrust Policy in the new electricity industry", *Energy Law Journal*, 17, 29-58.

Chapter III

A primer on market coupling

1 Introduction

In this chapter we discuss how the theoretical notions analyzed in Chapter 1 have been transposed in the policy debate on Market Coupling mechanisms (MC).

Currently very different arrangements are implemented to allow electricity transactions across different European borders. In most European regions capacity is currently allocated through explicit auctions, while in some others implicit auctions are in place.

In spring 2006 ERGEG launched the Electricity Regional Initiatives (ERI) in order to foster integration of European electricity markets, with the final goal of achieving a single pan European market. The ERI approach is based on integrating groups of Countries, or Regions, in a first phase. In a second phase inter-regional coordination should be achieved[56].

We discuss in detail the ERI progresses in the following Chapter. Here we focus on the main issues featuring in the debate on Market coupling in Europe.

In section 2 we discuss methods to assess the transmission capacity available for cross-border transactions under different allocation methods. In section 3 we analyze alternative mechanisms of spot market coupling. In section 4 we discuss the nature of long term transmission rights in a market coupling environment. Finally in section 5 we consider the much discussed issue of criteria to split among the involved TSO's the congestion rents resulting from implicit-auction based transmission capacity allocation mechanisms.

[56] Seven ERI regions have been identified: Central-South, Central-East, South-West, Baltic, France-UK-Ireland, Northern and Central-West.

2 Transmission capacity for cross-border transactions: from NTC-based to flow-based market coupling

Alternative notions of interconnection capacity are relevant to different allocation methods. The traditional "explicit" method allocates to market participants a "commercial" notion of capacity, which gives the owner the right to balance withdrawals in a Country (the importing Country) with corresponding injections in another Country (the exporting Country). This is the notion market participants are ultimately interested in, since it is directly related to cross-border transactions (buying electricity in a country and selling it in another country). In this sense this notion is referred to as "commercial".

The implicit auction method does not allocate to market participants any right on the use of the interconnections; a notion of interconnection capacity equivalent to the "commercial capacity" is defined and allocated in the form of financial transmission rights (see section 4)

In either case the network representation assumed for the purpose of clearing the energy markets may approximate to a different extent the real network. In the following sections we investigate the implications of assuming network approximations when clearing the market in terms of sub-efficient use of the available resources.

To do so we consider two extreme approaches: the so called "interface approach" – implemented in the traditional "explicit" allocation methods – assuming the "commercial" interconnection capacity for the purpose of market clearing – and the nodal flow-based approach, considering the real physical network when clearing the energy markets. We also consider a somehow intermediate flow-based model, implemented assuming a simplified representation of the network.

2.1 Interface Approach

In the interface the TSOs allocates the right to balance withdrawals in a Country (the importing Country) with corresponding injections in another Country (the exporting Country).

Allocating commercial capacity guarantees the efficient use of the network for very simple network topologies, like the one represented in figure 1.

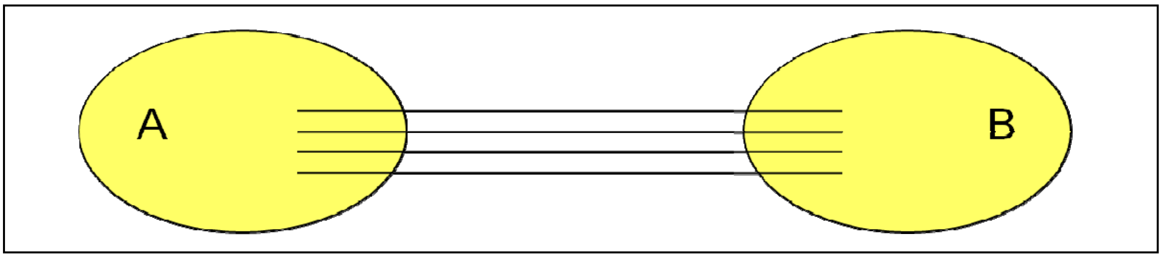

Figure 1 – A two-nodes network. Physical representation.

In this figure we have assumed that the two nodes are connected by a transmission grid composed of four parallel wires, each with a capacity of 1000 MW. The interface approach identifies the overall capacity of the interconnection between the two countries, net of the security margin, the so called Net Transfer Capacity (NTC), as shown in figure 2[57].

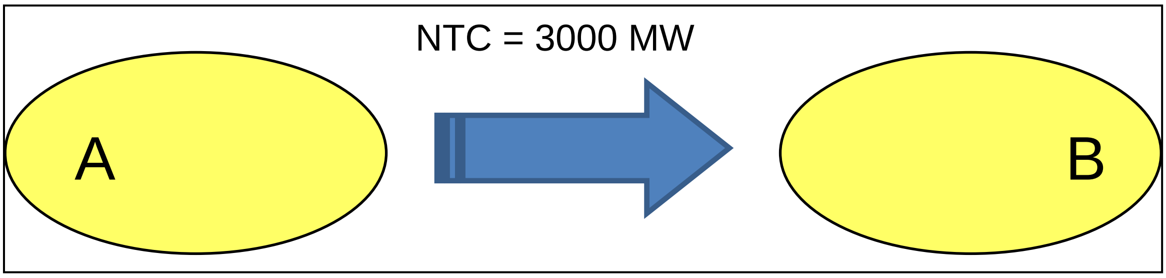

Figure 2 – A two-nodes network. NTC representation.

[57] In this example we assume that the security margin is determined according to the *N-1 security margin* principle. That is, the available capacity is determined as if one of the lines is out of service.

This approach becomes inadequate when applied to more complex grids. As shown in Chapter I, in general a number of different combinations of net injections and withdrawals are compatible with the same physical infrastructure.

EXAMPLE

In this example we assume the presence of four countries (called A, B, C, and D) and four interconnection lines. We further assume that there is no congestion within each country. This allows us to refer to A, B, C and D equivalently as to countries or nodes.

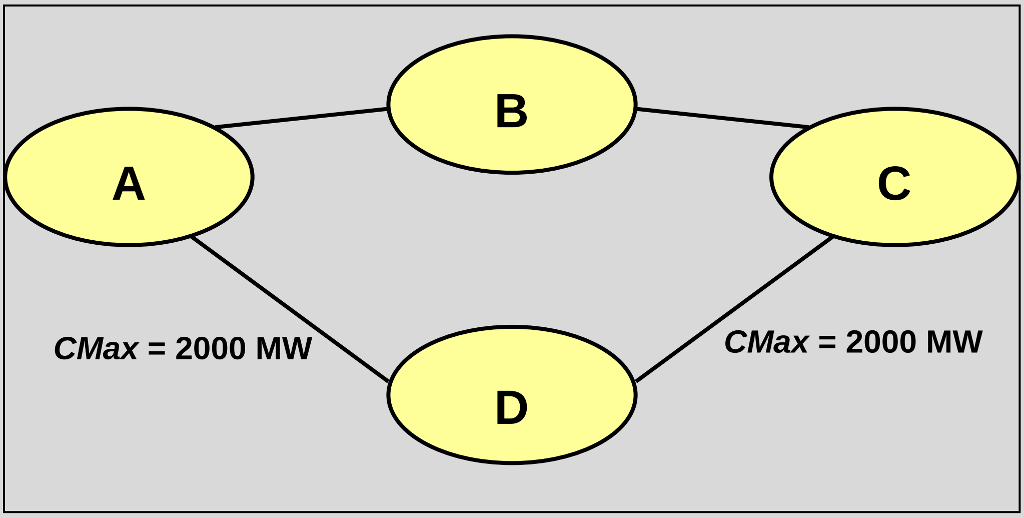

Figure 3 – The four-nodes, four-line network. Physical structure.

Let country D be the *swing node* and assume all the lines are identical. These assumptions yield to the following PTDFs matrix[58].

[58] See Chapter I.

Country \ Line	A - D	A – B	B - C	C – D
A	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
B	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
C	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$	$\frac{3}{4}$

Table 1 – PTDFs matrix in the base case

Let us finally assume that the interconnection lines connecting A and C to D have a thermal limit (or *Maximum Capacity, CMax*) of 2000 MW.

Transmission constraints may then be expressed as:

$$A - D : \frac{3}{4}a + \frac{1}{2}b + \frac{1}{4}c \leq 2000$$
$$C - D : \frac{1}{4}a + \frac{1}{2}b + \frac{3}{4}c \leq 2000$$

Where,
- *a*: net injection in country A;
- *b*: net injection in country B;
- *c*: net injection in country C.

The meaning of the two inequalities is that the overall power flow on each line cannot exceed the capacity limit that we have assumed.

We define *feasible transaction* as a vector made of balanced net injections which do not violate any of the transmission constraints. We can now characterize for our simple network the *set of feasible transactions* that can be implemented in that network.

To simplify the graphical representation in what follows we will assume that a, the net injections in A, equals 1000 MW.

This enables us to draw the constraints in terms of b and c only and thus to represent the set of feasible transactions in a 2D plot. By assuming a=1000 the transmission limits become:

$$A - D : 750 + \frac{1}{2}b + \frac{1}{4}c \leq 2000$$
$$C - D : 250 + \frac{1}{2}b + \frac{3}{4}c \leq 2000$$

These new constraints can be represented as lines in the parameter space (b,c), as shown in figure 4.

On the horizontal axis we have shown power injections in country B, while injections in C are represented on the vertical axis.

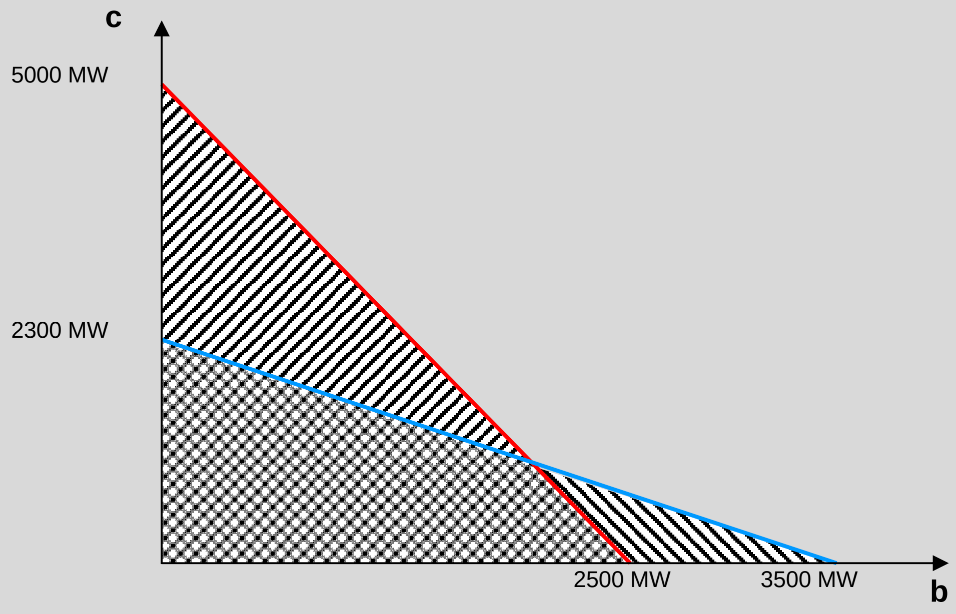

Figure 4 – Feasible transactions in the (b,c) space

All the combinations of injections which are below the red line are compatible with the A-D line constraint, while those below the blue line are compatible with the C-D line constraint.
However, only the combinations which are below both lines are compatible with both the constraints. We call the area below both the constraints *region of feasible transactions*, since it contains all the possible vectors of injections which can be implemented, without violating any capacity constraint.

The main drawback of the interface approach results from the fact that TSOs have to fix ex-ante the NTC vector to allocate to the market. In doing so the TSO must guess the feasible transaction vector that is *most valuable* to market participants. This would require information that is not available ex-ante: the (equilibrium) value of power at each node.

If the NTC vector allocated to the market is not the most valuable one, then some gains from trade will not be exploited.

Assume for example that the TSO allocates to the market the vector NTC_1, as shown in figure 5[59].
This point is located on the frontier of the region of feasible transactions, meaning that the usage of the available capacity is maximized. If the as-bid costs at the different nodes turns out to be such that transactions between B and D are much more valuable than transactions between C and D, a different vector of transactions, NTC_2, could generate a higher total surplus and the initial selection of NTC_1 by the TSO would be inefficient.

[59] We still assume 1000MW injections in A.

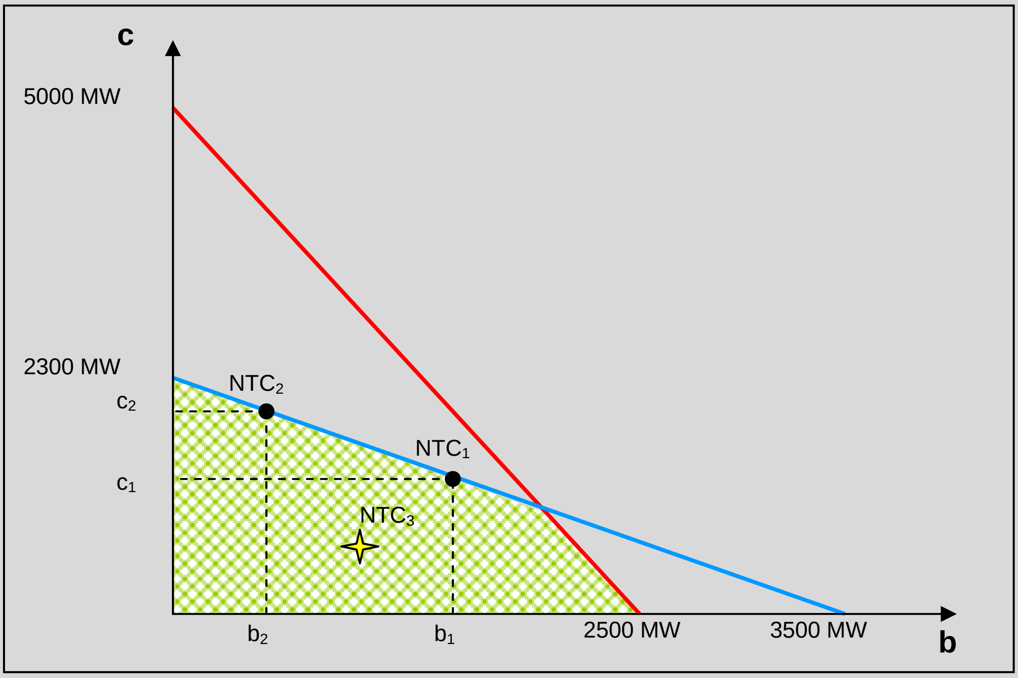

Figure 5 – The NTC in the region of feasible transactions.

Moreover, the assumption that injections at all the nodes of the same country cause the same flows on interconnection lines does not hold in practice. This means that NTC reflects also a guess by the TSO on the intra-country injection/withdrawal patterns. That creates a further source of uncertainty, which may bring TSOs to assess lower values for NTC, like NTC_3 in the figure, which lies *inside* the region of feasible transactions, rather than on the frontier.

***

ETSO describes the NTC calculation process in the following terms. The starting point in computing NTC consists in considering the Total Transfer Capacity (TTC). It represents the maximum exchange program between two areas compatible with operational security standards applicable at each system if future network conditions, generation and load patterns were known in advance.

TTC is computed by TSOs assuming a given power system scenario, which is determined looking at past flows on the considered line in certain time periods and assuming invariance of load flows on the lines not explicitly considered. Effectively TTC equals the maximum exchange program between the two considered areas only if the generation and load pattern in these areas and in others strictly connected exactly correspond to the assumptions made in the base case.

ETSO recognizes that forecasts become more precise when approaching the program execution, so that the TTC value may vary while approaching real time. There fore NTC values are recomputed (and possibly allocated to the market) in several stages. ETSO refers to two moments in time in which materially different information on the flows are available: the *planning phase* and the *capacity allocation*.

Planning NTC calculations are based on typical system conditions based mainly on past observations of few network snapshots. Capacity allocation NTC calculations take into account the exchange programs already nominated and actual system conditions.

The Net Transfer Capacity is given by the difference between the TTC and the Transmission Reliability Margin (TRM), a margin needed to deal with uncertainty about actual flows.

The NTC assessments are performed by the TSOs through load-flow calculations which require a set of data, including a network model and input data, describing load and generation patterns forecast, and network topology at the studied timeframe. In order to guarantee a certain degree of coherence among the calculations implemented by the different TSOs, the capacity assessments are based on commonly agreed datasets, also called base cases.

Starting from the same base case, several ways exist on how to simulate the power exchanges between two areas. The choice about the methodology to apply is left the TSOs and has to be made with the aim of reaching the most realistic result.

Once NTC is calculated, TSOs have to decide whether to allocate all or only part of it to market participants. This is an important decision since, due to variations from the base scenario, the effective available capacity could be greater or smaller than expected. In general, TSOs adopt a conservative approach with this respect, allocating less than the entire available capacity at the beginning, eventually increasing the remaining part closer to real time[60].

However, some TSOs – like the Italian one, for example – immediately allocate all the computed NTC to market participants, eventually curtailing part of the rights to use it in case they exceed the effective capacity of the network in real time.

ETSO publishes twice a year indicative and non-binding NTC-values on its website in order to provide a general European overview. TSOs are instead responsible for calculating and publishing net and available cross-border transmission capacities more frequently, as often as requested by national rules, market actors or allocation mechanisms.

[60] On the Swiss-German border, for example, capacity is allocated very close to real time, in order to account for the power flows generated by the wind power generators located in Northern Europe, whose size is not predictable in advance.

2.2 Flow-based approach

Contrary to the NTC-based approach, the flow-based approach directly internalizes the information about the value of electricity at different locations when choosing the optimal use of available transmission resources. On the basis of energy offers and bids at different locations, the welfare-maximizing level of net injections in each node is determined.

Therefore the actual set of feasible transactions is chosen as the one maximizing the value of transmission capacity. In this sense the "commercial capacity", defined as the possibility of producing a net-surplus in a country to compensate for a net-deficit in a neighboring country, is not given ex-ante, but it is endogenously defined as part of the market outcome in an efficient way. Further, what we called "commercial" rights are entirely allocated to the TSO, or market operator, that buys electricity in the low-price countries and sells it in the high-price country, thus appropriating the economic value of capacity. However, as shown in Chapter I, those rights can be transferred to market operators in the form of a financial rights.

In terms of figure 5 the point on the frontier corresponding to the net-injections determined in the flow-based modal is determined by the relative prices of electricity in the involved countries. If, for example, the price in B is lower than in C then the combination of transactions which will be chosen will be shifted toward the B axis[61] so that most of the injections will occur in this country.

As long as market clearing in the flow-based (as well as in the interface model) is carried out with reference to the real network representation, net-injections allowed by the TSO will lay on the frontier of the feasible transaction space. We now turn to the case in which a simplified network representation is assumed for market clearing purposes, in which nodes are aggregated into zones within which it is assumed no congestion. Examples of the application of such a "simplified" FB model are: Italy, for within country congestion management; CWE for cross-country congestion management; Texas and California (in the past) in the US.

[61] Note that we are still assuming that the injections in A remain fixed at the level of 1000 MW.

We can discuss this approach within the context of our example:

Let us assume that real network shown in figure 1 is simplified, for flow-based market clearing purposes, in the one shown in the following figure 6. The model implemented in the energy market clearing stage will not consider B and C as two separate nodes, but will aggregate them into a single “zone”, called E.
That amounts to assuming that injections in either node B or C have identical impact on the AD and CD lines and that the (market equilibrium) value of power in B and C is the same.

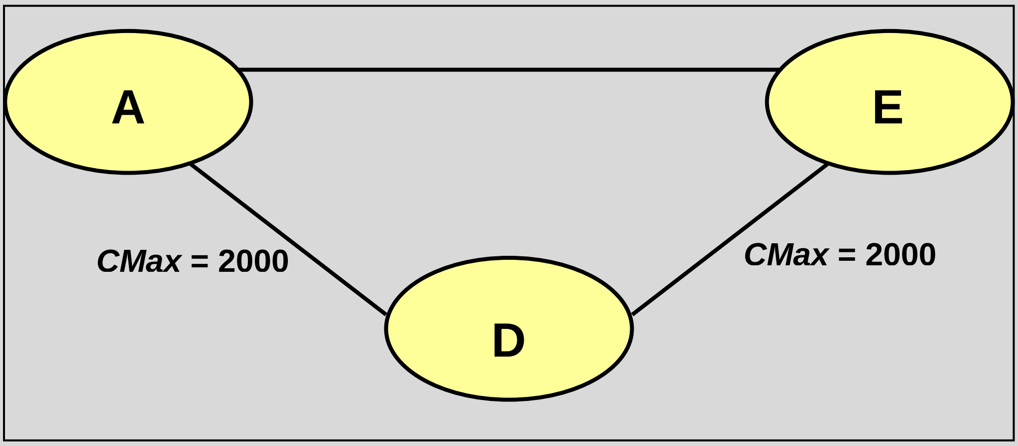

Figure 6 – The structure of the network in the “zonal” FB approach.

Notice that, in our example, despite zonal aggregation, the degree of simplification of the network representation is still lower than in the NTC model because the physical structure of the system is to some extent explicitly taken into consideration when clearing the market. However, lines and nodes used for market clearing purposes are now only indirectly related to the physical infrastructure.

In this new setting the network is still described by a PTDF matrix and Thermal Limits for each line. PTDFs have different values since they are intended to represent the impact on line loading of a unit injection in any of the nodes included in the zone. We will call these new PTDFs *zone-to-line* as opposed to the base ones, called *node-to-line*[62].

The problems resulting from approximating the network with a zonal representation are of the same nature as those described before when discussing the interface approach. The zonal representation of the network implicitly relies on an assumption about the market outcome.
In case this assumption is not ex-post correct, underutilization of the grid or need for within-zone redispatching might occur.

In order to show that, we need to investigate how *zone-to-line* PTDFs can be calculated based on the "real" node-to-line PTDFs.
Assume, for example, the highly intuitive method of calculationg zone-to-line PTDFs, between E and a generic line *I,* as a weighted average of the node-to-line PTDFs, in which weights are equal to the share of production in every node within country E, that is:

$$PTDF_{E,i} = x \cdot PTDF_{B,i} + (1-x) \cdot PTDF_{C,i}$$

where x is the share of net injections in node B, according to the forecasts of the TSO.

[62] We take this jargon from Smeers (2008).

The PTDF matrix for the simplified network is then the following:

Line / Country	A – D	A - E	E – D
A	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
E	$\frac{1}{4}(1+x)$	$-\frac{1}{4}(1+x)$	$\frac{1}{4}(3-x)$

Table 2 – Zone-to-line PTDF matrix

For the simplified network representation, then, the transmission constraints can be expressed as follows:

$$A-D: \frac{3}{4}a+\frac{1}{4}(1+x)e \leq 2000$$
$$E-D: \frac{1}{4}a+\frac{1}{4}(3-x)e \leq 2000$$

In the two inequalities, x is exogenously determined by the TSO and e is the variable representing injections from country E.

Notice that by implementing the simplified representation for market clearing purposes, the TSO (or market operator) can constrain, by the choice of a suitable "zonal" electricity price, only the sum of injections in nodes B and C, i.e. total injections in zone E.

Recalling our assumptions that the level of net injections in A is equal to 1000 MW the constraints in country E, becomes:

$$A-D: e \leq \frac{8000-3000}{1+x}$$
$$E-D: e \leq \frac{8000-1000}{3-x}$$

Let us now assess what set of injections in node B and C is compatible with the above constraints. We can easily do that by replacing $b+c$ for e in the constraints.

We obtain:

$$A-D: c \leq \frac{8000-3000}{1+x} - b$$

$$E-D: c \leq \frac{8000-1000}{3-x} - b$$

Assume for example that the share of zonal production in node B, forecasted by the TSO is ¼. Then, substituting x = ¼ in the previous conditions, we obtain the feasible transactions region.

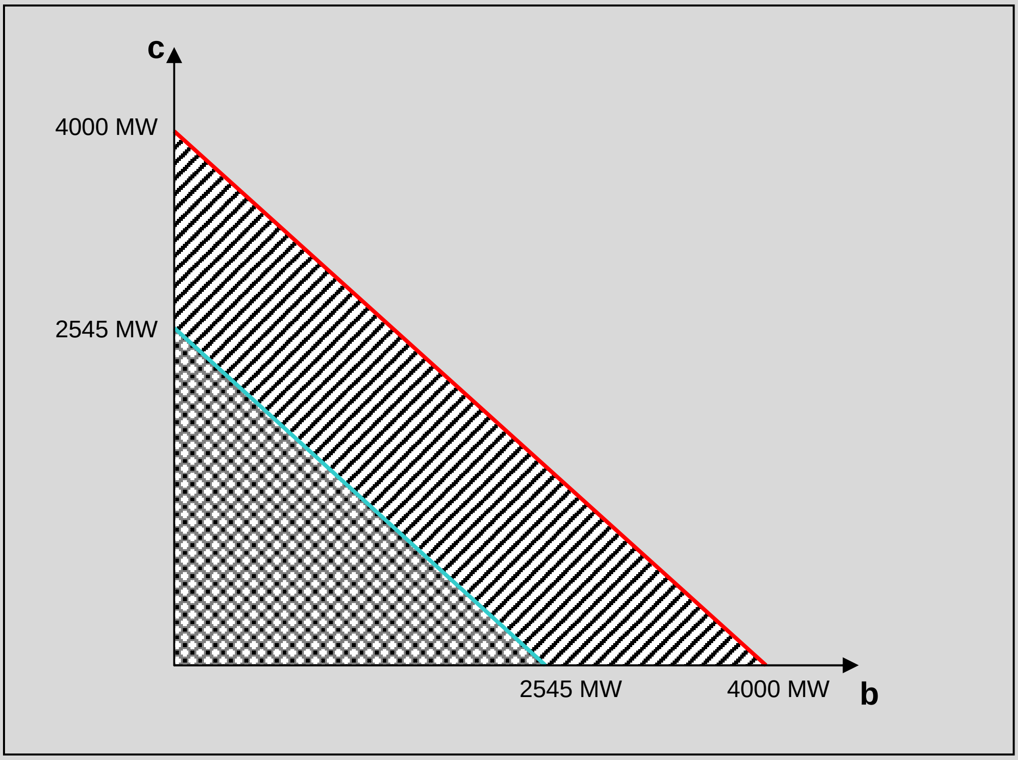

Figure 7 – Transactions space in the "zonal" FB approach[63].

The feasible transaction region is the surface below both the constraints.

[63] Note that we are still assuming that net injections in A are fixed at the level of 1000 MW.

Figure 8 shows a comparison between the region of feasible transactions obtained when considering the real network representation and the one resulting from zonal aggregation.

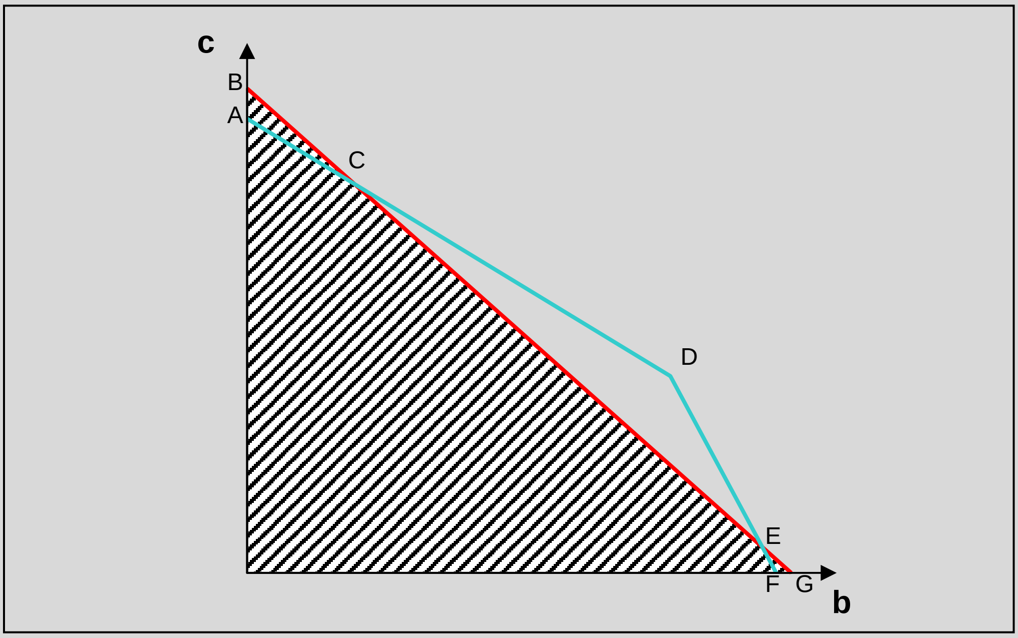

Figure 8 – Regions of feasible transactions in case of "pure" and zonal FB approach.

Areas of redispatching and *area of under-utilization* of the grid result from assuming the simplified network. Consider the areas identified by the triangles ABC and EFG. These areas are included in the new region, but are above the frontier of region obtained. This means that all the transactions included into these areas violate the real transmission limits of the network, despite being considered ad "feasible" under the simplified network assumption. In case the net injections corresponding to the zonal flow-based equilibrium fell in that regions some redispatching by TSOs will be necessary to guarantee system security. Triangles ABC and EFG represent, therefore, the *areas of redispatching.*

The transactions included in the CDE area excluded from the feasible under the simplified network representation, although they do not violate any real transmission constraint.

Triangle CDE therefore identifies an *area of under-utilization* of the grid.

As already noticed in correspondence to the NTC example, since real-time redispatching is costly, TSOs might want to assume a conservative position about the region of feasible transactions, in order to avoid real-time intervention[64].

For example the TSO might set the reliability margins so that the constraint implemented at the market clearing stage becomes the green line connecting A and F.

Such constraint ensures that any combination of net injections (b, c) that may result in the market equilibrium does not violate the real network constraints.

That choice by the TSO comes to the expense of enlarging the under-utilization area.

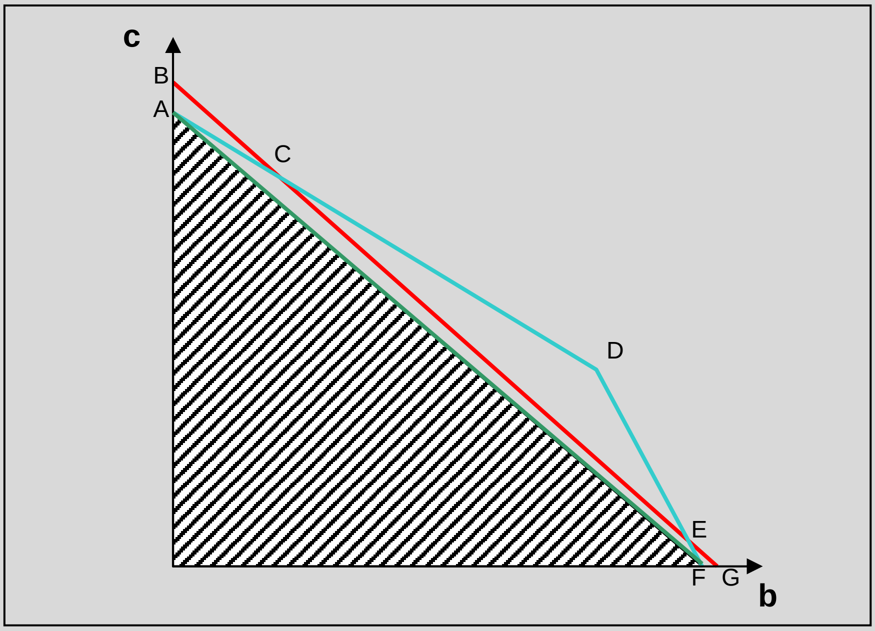

Figure 9 – The conservative approach to the region of feasible transactions.

[64] In the example, this means that the TSO wants to avoid transactions falling within the ABC and EFG areas.

The analysis previously presented shows that the simplified version of the FB approach turns out to have a drawback similar to the NTC approach.

Transmission capacity does not depend only on the physical characteristics of the network but depends also on assumptions on the allocation of injections among the internal nodes, i.e. on assumptions on the value of power at different locations.

The size of the inefficiencies resulting from the zonal approximation is a matter for empirical assessment.

3. Spot markets coupling mechanisms

Market integration can be implemented in different ways. In particular, the policy debate has focused mainly on two notions: *Market Coupling* (MC) and *Market Splitting* (MS). Since both of the approaches aim at the efficient result, that is inducing the market transactions which maximize social welfare, the difference among them is in the way they reach this outcome.

MC evokes the existence of separate markets, linked together by the existence of a coupling mechanism. MS, instead, is based on the assumption that the markets of all the involved regions are merged into a single one.

The differences between the two designs can be analyzed under three intertwined points of view: institutional, algorithmic and economic. Next we discuss the features of alternative market coupling models from those perspectives.

Under MS a unique PX operates across all the countries which are part of the system. Under MC the markets of the regions involved are to some extent cleared separately by each national (or regional) PX. A central entity exists whose role is to couple the outcomes of the different markets in order to determine the efficient set of cross-border transactions. Different MC designs assign different roles to the coupling entity and to the national PXs.

In this respect a common taxonomy calls *Price-Based market coupling* the setting where the coupling stage determines the complete market outcome, including the cross-border transactions and the equilibrium prices in each national market.

Volume-Based market coupling is the setting where the coupling stage determines cross-border flows (and the corresponding payments), but domestic prices are determined locally by each PX in a separate stage. In order to appreciate the notion of Volume-based market coupling it can be useful to represent a coupling mechanism as composed of three stages. In the first stage, each national PX sends to the coupling entity a Net Exportation Curve (NEC), based on domestic bids and offers. The NEC[65] represents the relationship

[65] A description of how the NEC is computed is presented in chapter 1, §3.2.

between the domestic price and the quantity exported (or imported) in each area. In the second stage a central institution calculates the cross-border equilibrium transactions based on the NECs, thus determining cross-border flows and the equilibrium prices at each side of the borders. In the third stage each national PX computes its internal market equilibrium, taking as given the import/export flow resulting from the second stage

The more the NEC and/or the coupling algorithm embed the constraints and matching rules featuring in the domestic markets (e.g. the rules applied for domestic congestion management, or the block bid constraints), the closer the equilibrium prices resulting in the coupling stage will be to the domestic equilibrium prices resulting from the last stage. In the extreme case in which the NEC and the coupling algorithms reflect all the nation specific constraints the prices resulting in the two stages will be identical.

Volume-based market coupling is said *Tight* or *Loose* according to the extent to which the coupling algorithm incorporates the set of information and matching rules characterizing the domestic market outcomes. In a Tight model, most of the features of each national markets are reflected in the coupling algorithm, so that the prices resulting in the coupling phase will be very close to those determined by the national clearing algorithms. In a Loose model, the prices resulting in the coupling phase may turn out to be materially different from those determined by the national clearing algorithms.

Market Splitting and Price-coupling determine the same (efficient) utilization of cross border transmission capacity, given the as bid supply and demand information. Achieving this outcome, though, may be costly in terms of coordination of national market designs and exchange of information.

Volume-based MC and, in particular, loose volume-based MC require a lower degree of information exchange and coordination between the separate PXs and the central coupler, thus mitigating the need for total harmonization. However, that model may result in inefficiencies. Consider, for example, the case of two countries, A and B. At the end of the (stage 2) coupling process, the central coupler determines an equilibrium flow of 1000 MW from A to B, and corresponding prices, say, in country A 54 €/MWh and in country B 58 €/MWh. Since the coupling algorithm does not match exactly the domestic clearing rules, it can happen that domestic prices obtained by assuming 1000 MW cross border flow will not be consistent with the direction of that flow. For example, price is A could be confirmed at 54 €/MWh while price in B could be 53 €/MWh, because of block bid constraints that were not accounted for in the coupling stage. In that case, the flows determined in stage 2 are inconsistent with the values of power in the involved countries. Market participants in country B would find it convenient to sell power in country A, rather than importing it. Note that this kind of inefficiency is more likely to determine when the values of electricity in the coupled markets are closer.

3.1 Coupling Intra-day Markets

Intra-day markets (IDM) are becoming an increasingly important element of wholesale power market design for at least two reasons. First, as the generation industry evolves towards a less concentrated structure, the availability for small market players of balancing becomes crucial to enable them to compete against large portfolio players. Second, the increasing role of renewables as wind and sunlight in production capacity makes efficient IDM even more necessary, since predictions about their production patterns are much more accurate close to delivery. Since liquidity is a central issue for IDMs, the possibility to bring together the liquidity of different national intra-day markets is crucial, especially for relatively small markets that could benefit from the liquidity of larger neighboring markets.

Some-how surprisingly the IDM market design around which consensus is growing is different from the one prevailing for day-ahead market coupling. In fact, while the implicit auction system appears to be “the model” for day-ahead market coupling, continuous trading with first-come first-served transmission capacity allocation is often referred to as the most appropriate way to integrate IDMs.

One of the arguments against MC in intraday markets[66] is that, due to the very short timeframe, the market conditions would change immediately after the closing of an auction, making the market result inefficient. Guaranteeing an efficient result with MC would therefore require an excessively high number of auctions. In a continuous trading system bids and offers are continuously matched. Bid/offer couples corresponding to

[66] Sustained, for instance, by traders in the ELBAS market.

injections and withdrawals in different countries are matched only if transmission capacity between the two countries is available.

ETSO (April 2008) states that, provided IDMs result in low trading volumes, a continuous trading system with no charge for capacity could be implemented.

It is not clear why the reduced size of the IDMs justifies implementing a design so different from the one pursued in the day-ahead; there is not even evidence of any transaction cost savings, since the implicit auction algorithms and coordination procedures should be put in place for day-ahead market coupling. Further, we already mentioned trends that are likely to increase the importance and the scope of IDMs.

Second, the first-come first served method does not guarantee that the welfare maximizing set of transactions will be carried out, or, in other terms, that available transmission capacity will be allocated to those party that value it the most.

Finally, a mixed model in which day-ahead capacity allocations are done at a charge (although implicit) and IDM capacity allocations are done on a free (first come first served) basis, may create distortions in the bidding strategies of market players endowed with market power. In particular, it could turn out to be profit maximizing for a market participant with market power, to shift part of the transactions to the IDM market. Incidentally, we notice that such strategies would (artificially) contribute to increasing the size of the IDM, which would: 1) run against the assumption on which support for a continuous trading system appears to be based; 2) could make dispatching more complicated or increase the need for reserve capacity.

In general, the value of implementing intraday markets depends on the structure of a country's production capacity and has to be compared to the costs of a more dynamic management of the transmission network. In a country like Germany, for instance, implementing intraday markets would allow allocating potentially huge amounts of low cost electricity, produced by wind power plants, close to real time. In this case, the value of intraday markets would be very high, thus outweighing the costs induced by the added complexity to network management.

The result would not be as clear in a country like Italy, where the contribution of renewables to generation capacity is limited. In this case, the value of intraday markets for the system as a whole would be likely not to compensate for the higher costs induced by the management of injection/withdrawal programs close to real time.

4. Long Term Rights

As pointed out by the analysis presented in Chapter I[67], transmission rights – either physical or financial - allow implementing long term cross-border transactions at a fixed price.

Physical Transmission Rights (PTRs) assign their holders the right to inject a electricity in one country, accompanied by the right to withdraw an equal amount in another country. It allows the physical utilization of a portion of the interconnection capacity among the involved countries. *Financial Transmission Rights* (FTRs) are a financial hedge against the variability of spot price-differences in the importing and exporting countries. A FTR is a purely financial instrument entitling its holder to receive a payoff at a maturity date.

Fixing in advance the price for the utilization of interconnection capacity, PTRs and FTRs provide market participants with an instrument to hedge against the risk induced by price volatility. They are explicitly auctioned by TSOs on a yearly, quarterly and monthly basis and the acquirer of these transmission rights enters into contractual relationship with the TSOs.

Both PTRs and FTRs can be implemented either as *options* or *obligations*. In the former case, the holder acquires the right to use a certain amount of a congested interconnector – in case of PTRs – or to receive an amount equal to the price-spread between the involved countries[68] – in case of FTRs. In order to exercise this right, the holder is required to *nominate* the energy transfer – in case of PTRs – or nominate his right to receive the payment – in case of FTRs – at the D-1 timeframe.

In case of obligations, instead, the market party who acquired a PTR is compelled to carry out the corresponding energy transfer, while the holder of a FTR obligation has to receive the price difference at the maturity date, which could also be negative. In case it is negative, the FTR holder has to pay the price difference.

For both PTRs and FTRs, the amount of rights which can be auctioned cannot exceed the available transmission capacity. Physical Transmission Rights allocate the

[67] See Paragraph 3, Chapter I.
[68] Given two countries, A and B, the price spread is the difference between the price in country B and the price in country A.

right to use interconnection capacity and clearly a TSO cannot allocate an amount of capacity which is greater than the available capacity.

In the case of Financial Transmission Rights, TSOs finance the payment to the right holders through congestion rents. Auctioning more transmission rights than the available capacity a TSO would risk having to face an amount of payments greater than the congestion revenues. To be risk balanced for the TSOs, the overall quantity of FTRs sold at the purchase date must be simultaneously feasible under "physical" conditions at the maturity date.

In general, *ceteris paribus*, FTRs are superior to PTRs because they do not constrain the spot market outcome, while allowing perfect hedge for transmission costs. Since FTRs do not allocate any physical capacity in advance, the spot market is free to determine the equilibrium transactions which are most suitable given the spot prices – i.e. the efficient outcome – even if it involves transactions which are completely different from those resulting from the FTRs allocation. This result is efficient for the TSOs as well, since the congestion rents resulting from the market equilibrium transactions will always be high enough to pay the revenues to the FRTs holders.

The allocation of physical transmission capacity through PTRs, instead, does not allow the spot market to reach the efficient outcome. This induces an inefficient allocation of generation[69] and damages also the TSOs that will receive less congestion rents.

Since with PTRs interconnection capacity between neighboring countries is explicitly allocated at a price fixed in advance, their implementation requires a low degree of coordination among the involved markets, even in absence of liquid spot markets. PTRs actually do not even require that spot prices are determined in each country.

Their simplicity of implementation makes PTRs extremely compatible with a context of relatively immature electricity markets, where they would reduce the risk of cross-border transactions for market players, inducing new entrants to join the market increasing its liquidity. The small requirements in terms of national market design coordination explain the widespread utilization of PTRs across European interconnections.

[69] See Chapter 1, paragraph 3.3.1, for more a more detailed analysis.

Since the revenue guaranteed by FTRs depends on the price-spread between the involved markets, their implementation requires the presence of liquid of spot markets which guarantee the determination of spot prices. Moreover, FTRs require a high degree of coordination among the involved markets. An implicit auction based market coupling design, guaranteeing coordination among the coupled markets, provides the ideal framework for the utilization of Financial Transmission Rights.

Incidentally, we notice that the distinction between physical and financial rights becomes blurred, the easier physical rights may be transferred between market participants. If the right holder finds it no more convenient to execute the underlying economic transaction, he can sell the right to use the interconnection capacity to other market participants willing to sell power in the other country. This reduces somehow the connection between the PTR and the use of the interconnection capacity, making it more similar to a pure hedging system like a FTR.

However, Physical Transmission Rights are also compatible with the market coupling design and can perfectly coexist with Financial Transmission Rights. This result is obtained allowing the PTRs holders to exert their right matching the scheduled injections and withdrawals in the involved countries or selling the right to the MC institutions, which will implicitly allocate the underlying capacity in the spot market.

This can be done on a voluntary basis, allowing to the right holder who does not exercise it at the nomination stage to transfer it to the MC institution, in exchange for the congestion rent that the capacity will generate in the spot market, according to the *use-it-or-sell-it* (UIOSI) principle.

Another implementation scheme is also possible in which, if the right holder does not exercise his right, it expires after the nomination stage and the freed capacity is implicitly allocated to market participants in the spot market, according to the *use-it-or-lose-it* (UIOLI) principle.

***

An important item of the policy debate on the features of long term transmission rights is the *firmness* of those rights in case capacity turns out to be such that cannot be supported.

In presence of Physical Transmission Rights this means that, due to unexpected events – for instance the breakdown of physical infrastructures or errors in the definition of the available capacity at the issuing date (under the NTC approach) –, scheduled injections and withdrawals are not compatible with the security of the system. Whenever these events occur, TSOs have to implement some (costly) redispatch in order to guarantee the physical implementation of the transactions.

In presence of Financial Transmission Rights, the unexpected reduction of available capacity may cause the payments to the right holders to be greater than the congestion rents collected in the spot market by the TSOs, exposing them to losses.

In absence of firm transmission capacity rights, whenever an unexpected event occurs causing a reduction of the available capacity, the allocated capacity is curtailed. Currently, at the European level, right owners receive a compensation for curtailments at 100% or 110% of the price paid for capacity. This does not assure "firmness" in the meaning discussed above, because curtailed market participants are not compensated with the economic value of the capacity they lose.

The firmness issue, from an economic perspective, is a mere risk allocation issue. With non-firm transmission right, right holders bear the risk associated to the content of the right they purchased. In presence of firm transmission rights, such a risk is borne by the TSO and – to an extent that depends on the regulatory arrangements – socialized among all network users.

The optimal risk allocation depends on:

- The agents having more opportunities to manage that risk;
- The incentive implications of each possible allocation arrangement.

For what concerns the first issue, it appears that TSOs are the agents in the best position, since they have access to information about the status of the transmission network and the status of generation capacity that other market participants do not have.

For what concerns the second issue, the regulatory incentive structure placed on TSOs has to be considered. If, in particular, TSOs are called to bear part of the cost of ensuring firmness of rights, they have an incentive to understate the available transmission capacity. On the other hand, if all the risks are borne by market participants, TSOs might not have the right incentive to arbitrage between the real costs of curtailment and other counter-measures, such as coordinated redispatching, to guarantee the availability of the allocated capacity.

In the policy debate on firmness, two positions might be identified.

On the one side, regulators and traders state that if all the risk is borne by market parties, it imposes an extra cost to the system. In bidding for capacity, market players will consider a risk premium and their estimated value for capacity will be lower than it would be otherwise. The TSOs incentive not to consider alternative solutions to redispatching, would worsen the situation further. The solution supported by regulators and market players would be to introduce a market based compensation scheme, according to which any curtailment of capacity rights should be compensated at the market price spread (i.e. the price difference between the source and destination countries), except in the case of *Force Majeure*[70].

On the other side, TSOs agree that curtailments should be avoided, thus relieving market parties from having to bear any risk, only *after* the D-1 nomination stage. If an unexpected event occurs *before* nomination, transmission rights should be curtailed and compensation should be provided.

In ETSO's view, compensation at full market spread would merely transfer risk from the right holders toward end users, which would risk paying twice in case of curtailment: once through the increase in the market price and once through the compensation by the TSOs[71]. According to ETSO, maximizing the volumes of transmission rights traded increases the risk of curtailments.

The combination of compensation at full market spread and the need to maximize the volumes of transmission rights allocated in forward auctions, leads to maximize the risk transferred to end consumers. In order to avoid that, it is reasonable to assume that TSOs would be tempted to auction less cross-border capacity in order to reduce their financial risk. ETSO therefore believes that the key point on compensation is the establishment of an appropriate balance risk between market participants, TSOs and end consumers when, defining the compensation rules in case of curtailment.

[70] *Force Makeure situations* refer to those cases in which TSOs are facing conditions that could lead to large-scale system blackouts, requiring estreme measures outside of normal market operations. Since these circumstances are outside the direct control of TSOs, the risk of their occurrence should not be borne by TSOs and, as all risks beyond their control, should be considered as "market risk".
[71] This could be transferred to end consumers through increased tariffs.

5. Inter-TSO Compensation Mechanism

As introduced in the first chapter (reference in footnote), power flows on every parallel path from source to sink. This is the base of the inter-TSO Compensation Mechanism: cross-border transactions generate flows on interconnected third countries. In order to deal with the economic implications of that, shortly after the implementation of the first Electricity Market Directive 96/92/EC, European TSOs introduced cross-border transmission tariffs for wholesale contracts crossing national borders. These tariffs proved to be an obstacle to cross-border trade because, in some cases, eroded margins: for contracts crossing several borders, "pancaked" tariffs charged upon export, transit (sometimes more than once) and import of nominated power, severely reduced trade incentives.

Since 2002, agreement has been reached upon the idea that transmission tariffs should be substituted by a "non transaction-based" compensation system, i.e. the Inter-TSO Compensation (ITC) mechanism. A voluntary "transitional" ITC mechanism has been in operation since March 2002. Initially intended to last only until the end of 2002, it has repeatedly been extended and modified. Regulation 1228/2003 formalized the ITC arrangements as a right on the part of those TSOs "hosting" transits to receive payments from those "causing" the related flows. According to Article 3(1)-(2) of the Regulation:

"Transmission System Operators shall receive compensation for costs incurred as a result of hosting cross-border flows of electricity on their networks. The compensation shall be paid by the operators of national transmission systems for which cross-border flows originate and the system where those flows end."

Despite the Regulation has been into force since 2003 an active debate on the features of the ITC design is still in place and, due to the complexities underlying the ITC mechanism, a final agreement about its designs has still not been reached.

The key technical issues on which the current debate is focused refer to the identification of the Horizontal Network, the determination of the network costs and the allocation of the cost responsibilities to the TSOs. The aim of this paragraph is to shed light on these issues, discussing the proposals emerged in the debate.

5.1 The Horizontal Network

The Horizontal Network (HN) represents the portion of the total network which might be *significantly* affected by cross-border transactions, therefore the cost of the HN should be somehow shared among different countries. Part of the analytical effort up to date has been devoted to developing procedures to identify which part of the grid in the various TSO areas is relevant for international flows.

An ITC mechanism can define compensation on the basis of cross-border flows or transits[72]. In the former case, involved TSOs are compensated for all the power flows crossing their borders, including import/export flows. In the latter, compensation is provided only for those flows which use the network of a TSO, but which both originate and end in other networks.

Currently, the HN is identified through a standardized procedure based on transits. A simulation of the flows induced on an empty network by the injection of 100MWh in a line and the load of the same amount from another is run for every pair of tie-lines. Grid elements in the overall transmission network are included in the HN if they transmit at least 1 MWh in at least one of the series of DC load flow calculations performed.

5.2 Network Costs

For the purpose of the ITC mechanism, according to Article 3(6) of Regulation 1228/2003, network costs should be defined using standardized costing methodologies, taking new infrastructure – as well as existing ones – into account. Moreover, the amount of compensation should be reduced considering the benefits induced by cross-border transactions. This paragraph discusses the main issues connected to these requirements.

[72] A detailed description of the debate on the utilization of cross-border flows or transits is provided in paragraph 5.2.

Standard Costing Methodologies

The current solution to the definition of standardized costs, decided by ETSO in 2003, grounds on "regulated costs" – i.e. the costs accepted by the international regulators for the domestic transmission access tariffs – for the estimation of the forward looking average incremental costs. The practice of using regulated costs has been accepted by the European Commission, as stated in the consultation document on ITC, discussed in December 2008[73].

On the one hand, this solution solves the complex problem of finding an adequate standard methodology to calculate LRAIC. On the other hand, it introduces some problems. The adoption of regulated costs as proxies for the network costs implies that more efficient countries, having lower network costs, receive less compensation than less efficient countries, facing higher network costs. In other words, efficient countries pay a sort of "subsidy" to the less efficient ones, thus raising a *fairness issue*.

In theory, considering national costs[74] could also generate an incentive for TSOs and/or national regulators to artificially increase network costs in order to receive more compensation. Each TSO compensation is, in fact, based the proportion of flows in its area – defined as cross-border flows under the ITC mechanism – and the cost of the HN (elements) in that area. Artificially raising the network costs a TSO could increase its entitlement to compensation, without affecting the amount of compensation it is required to pay for using the HN in other TSO areas. The net effect will be to increase the net ITC entitlement, or reduce the ITC payments.

However, the incentive for TSOs to report artificially high network costs relies on the possibility to appropriate the resulting increased compensation received. This possibility seems, however, very unlikely in reality: national regulators would normally impose to use the higher compensation from ITC to reduce national transmission tariffs.

National regulators themselves do not seem to have real incentives to report artificially high network costs, since this would determine an increase of the fees charged on consumers. The only possibility for regulators to report higher network costs without affecting consumers would be to shift "common" costs – i.e. costs generated by network

[73] European Commission. DG TREN. *Consultation Document on the Inter-TSO Compensation Mechanism and on Harmonization of Transmission Tarification*. DG TREN/C2. (December 2008).
[74] We consider, here, national costs and regulated costs as synonyms.

elements not included in the HN – on the HN. This would allow increasing the compensation received, without affecting the fees charged on consumers.

The passage to a standard costing methodology to determine network costs, required by the previously recalled Art 3(6) of the Regulation, raises a series of further issues. There exist, in fact, some country specific characteristics – topography and quality/security standard, for instance – that can have significant effects on the cost of an electricity network. In order to take this into account, the discussion has evolved towards the opportunity to implement a standardized cost measure that reflects those national characteristics that can increase (or reduce) costs, relative to the European average.

New Infrastructures

Regulation 1228/2003 requires considering the costs of new infrastructure, as well as the costs of existing infrastructure. Generally, the decision to invest and on the size of the investment are taken by local TSOs, typically with the involvement of their regulator. Other TSOs have little to say on such topics and therefore on the costs of the new infrastructure.

These decisions are taken on the basis of "internal" needs. However, if the TSO considering whether to invest is concerned only with flows inside its boundaries, it might have lower incentive to undertake the investment, than in case it considers the whole effect on the system.

The current structure of the compensation mechanisms is "use based", meaning that the amount of compensation provided for every network element depends on its degree of utilization. Such an ITC scheme raises an *incentive issue*, possibly detrimental to the development of the EU transmission network. If a line is effectively heavily used for cross-border transactions, the compensation received might provide sufficient additional incentive to build the asset[75]. If this is not the case, the TSO might not be able to recover the costs of the investment and might decide not to invest.

[75] Note that this does not mean that the full cost recovery would come from foreign TSOs, since there would generally be domestic usage of the asset as well, even if its primary purpose was to relieve congestion abroad or to facilitate external utilization of the TSO's network.

Benefits

Typically, in the debate on the definition of the network costs, it is recognized that cross-border flows or transits can free transmission capacity for within-country transmission capacity, thus enabling market players to do more transactions or reducing the need for redispatching – and therefore its costs – by TSOs.

The following example can clarify this point.

EXAMPLE

Consider the three-country system represented in figure 12.

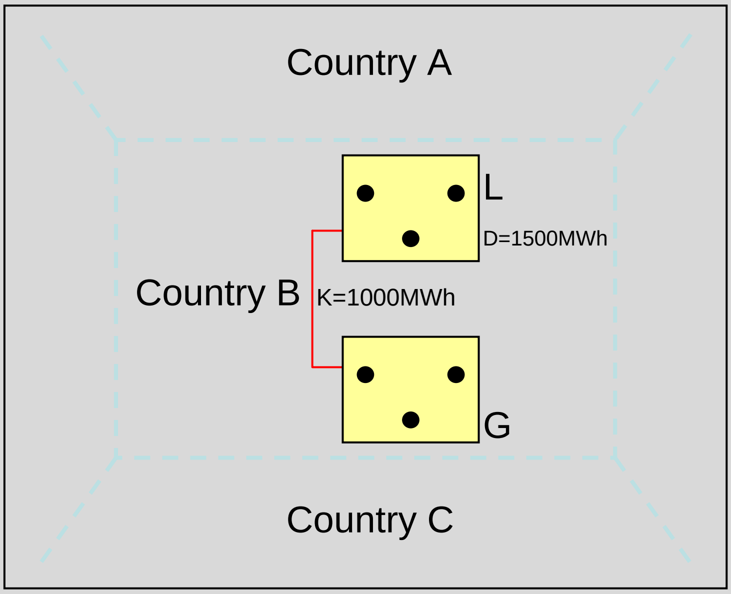

Figure 12 – The three country system in absence of transits

Node L (load) is located in proximity of boundary with country A.

Assume that this node is composed by both consumption and generation units, but the marginal cost of the generation units located in node L is higher than the marginal cost of the units located in node G (generation). Because of that reason, L is net buyer of electricity and G is net supplier[76].

Assume, further, that the demand in L is equal to 1500 MWh.

[76] Assume, for simplicity, that no consumption occurs in node G.

Since, by assumption, all the generation plants in L are less efficient than those located in G, then L will import all its demand from G.

If G-L line has a capacity limit equal to 1000 MWh, then the TSO of the B area will have to implement costly real time redispatching, in order to guarantee the physical stability of the internal network.

Assume now that a 1500 MWh transit from A to C, flows through country B. If the TSO in the B area uses the inflow to satisfy the demand in L and uses the generation plants to export 1500 MWh to C, as shown in figure 13, then it does not need to do real-time redispatching anymore. In this extreme example no power flows on the L-G line anymore.

Thanks to the cross-border transits 1000 MWh of within-country transmission capacity are freed-up thus enabling, in theory, consumers located in zone L to import 1000 MWh more from zone G[77].

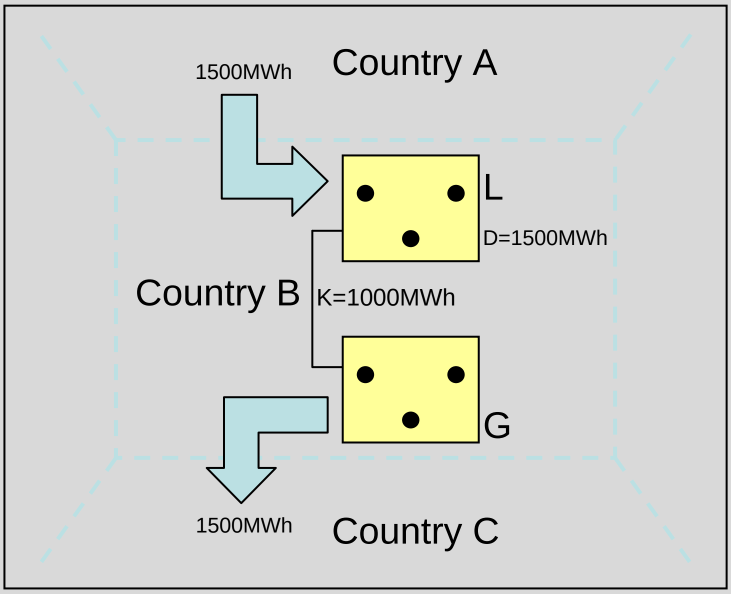

Figure 13 – The three country example in presence of transits from A to C

[77] In this example we do not consider, for simplicity, the presence of production limits in zone G.

Consider, further, satisfying demand in L with the 1500 MWh from country A reduces the transmission losses induced by the power flow that would otherwise occur using the L-G transmission line.
This is especially true in presence of a "long" country, where the distance between the L and G zones is greater than the distance between the L zone and the border with country A.

Notice that our example shows only some of the possible benefits generated by cross-border transits. These unexpected transits might in fact, free capacity on the internal network that the TSO can use in order to implement cross-border transactions that it would not have been possible to implement otherwise. In the much more complex context of a real network, it is therefore extremely difficult to exactly quantify the benefits generated by cross-border transits to the different TSOs.

***

Among the benefits induced by cross-border flows also congestion rents should be mentioned. Congestion rent is a byproduct of the allocation mechanism of a scarce resource and provides an indication about the value of the incremental investment in the congested asset. However, it is by no means an indicator of the contribution of the congested asset to the value created by the entire network.

Congestion rents are provided to TSOs only in presence of congested interconnections. It does not mean, however, that the value of non-congested lines is equal to zero. In general, the value of a high-capacity interconnection line not fully utilized can be very high, despite the fact no congestion rent is provided for its utilization. The following example can clarify this point.

EXAMPLE

Consider a network composed by three countries, A, B and C. Assume that A is the only production country and C is the consumption country. Power only flows through B[78]. Assume, further, that line A-B has infinite capacity, while the B-C line has a small capacity, equal to 1000 MWh.

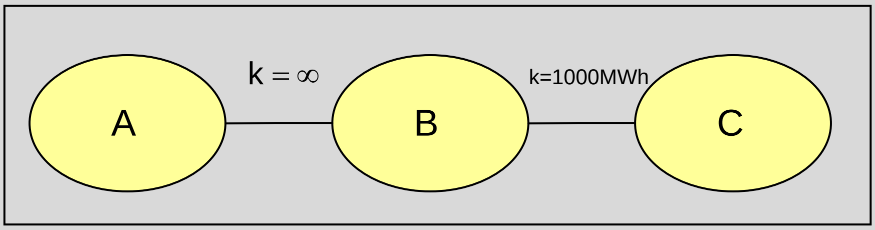

Figure 11 – The three-node example

In presence of a 1400 MWh power flow which goes from A to C, line B-C is congested and country B receives a congestion rent. Since the A-B line is not congested, country A does not receive anything.

This is like assuming that, since the A-B line is not fully used, its value is equal to zero. However, the value of the line is potentially very high, since it allows a large amount of power to flow from A to B.

In our extreme example, if country A "closes" the A-B line (or decides not to build it), no more cross-border transactions can be realized.

Another – less extreme – situation could be imagined, in which the presence of a high-capacity line with a low power flow on it helps satisfying the N-1 security margin of the network, thus allowing more transactions to be implemented in other parts of it.

In general, assessing the contribution of each network element to creating those benefits is problematic from theoretical point of view and impossible in practice.

[78] Note that this are only simplifying assumption, which are not strictly necessary to discuss our point.

Regarding the use of congestion rents, it has been argued that they should be allocated to investments in incremental infrastructures. However, in the current debate it is not entirely clear if this idea is linked to the provision of incentives to implement new investments. Such an incentive could be provided in two alternative ways:

a) Using congestion rents in order to provide a higher return on investments than the one provided by the market. Such a solution would enhance competition among investors in the construction of new infrastructures;
a) Allocating congestion rents to the TSOs investing in new infrastructure. Such a solution would enhance competition among TSOs, favoring those located in more "incentive friendly" countries.

In either case, it is important to point out that the way incentives are provided is not linked to the way incentives are financed. The link between congestion rents and incentives is therefore all but obvious.

5.3 Horizontal Network Cost Responsibility Allocation

Once the showed resources have been identified and the costs have been assessed, the following step is assessing the amount of resources used for cross-border transactions or transits (according to the approach) covered via the ITC.

In the recent debate, a number of alternative ways to allocate network cost responsibility to the different ITC parties have been proposed. This section reviews the main approaches: the Average Participation (AP), the Marginal Participation (MP), the With-and-Without (WW) and the current ETSO mechanism (ETSO2005).

Before describing the alternative approaches we analyze the choice of using the ITC mechanism to alternatively provide compensation either for all cross-border transactions or only for transits.

Cross-border Flows versus Transits

Another element of the discussion on the role of ITC is whether compensation should be provided for the costs incurred from hosting cross-border flows, or it is meant to provide compensation to the costs generated by transit flows. In this context:

- *Transit Flows* over a TSO's grid are defined as power flows which affect its grid but which both originate and end in the grid(s) of other TSO(s);
- *Cross-border Flows* over a TSO's grid are defined as power flows which cross at least one border.

The notion of cross-border flows is therefore wider, including not only transit flows, but also import and export flows (i.e. cross-border flows which either originate or end in the TSO's grid).

The choice between an ITC mechanism addressing cross-border flows or transits is linked to the organizational and regulatory framework.

Historically, cross-border transactions have been regulated by bilateral agreements between neighboring countries. This requires couples of countries to agree on:

- Available cross-border capacity levels;
- Allocation mechanism;
- Splitting of the congestion rents (in case allocation took place via market based systems);
- Transmission charges to be specifically applied to cross-border transactions.

In that approach a mechanism was needed to deal with unaccounted flows that the transactions between each couple of countries caused on the networks of third countries. The following example highlights the relation, in the current regime, between the ITC mechanism and the definition of Available Transmission Capacity (ATC) in a meshed grid.

Consider the simple system composed of three countries, represented in figure 14.

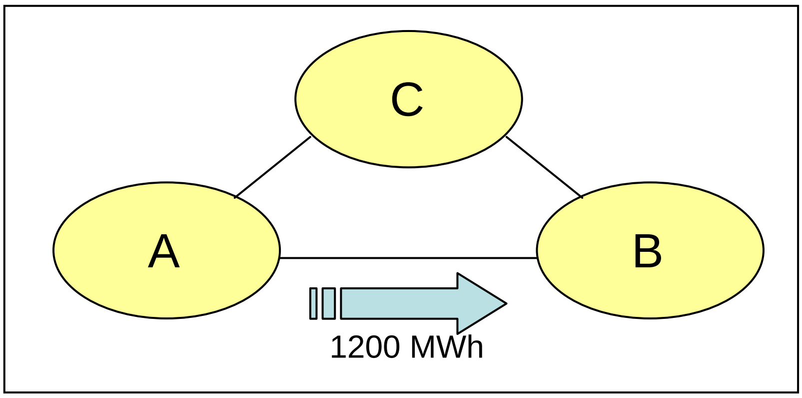

Figure 14 – The current definition of commercial capacity

Assume that country A and B agree that the commercial capacity between to the two countries is 1200 MWh. For simplicity assume also that the capacity on A-C and B-C lines is infinite, that consumption occurs only in country B and that no production occurs in country C. Due to network effects[79], only 800 out of the 1200 MWh flow directly on the A-B line, while the remaining 400 MWh flow through the A-C and C-B lines. According to our definitions, the 400 MWh flow through C is a “transit” for country C, while 1200 MWh is a cross-border flow for countries A and B.

Generators and customers in country C are not responsible at all for that 400 MWh flow through country C, but the TSO (and, consequently, final consumers) bears its costs. Moreover, this flow might require costly real-time redispatching by the TSO in C. According to the approach in place since 2002, compensation of country C is performed via the ITC mechanism applied to *transits*.

There is, however, a conventional element implicit in these trading arrangement. An alternative arrangement is shown in figure 15, where the maximum ATC between A and B is set to 800 MWh and the maximum ATC on the A-C and B-C lines is set to 400 MWh.

[79] We assume for simplicity that the resistance is the same over all the lines.

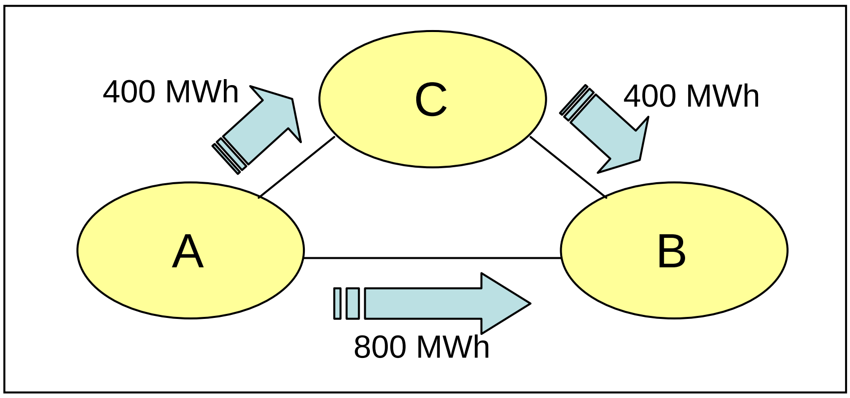

Figure 15 – Cross-border capacity taking into account transits

In this new setting, the power flow through country C is no more considered as a transit, but is now decomposed into two cross-border flows, for which the TSO in C receives compensation through congestion rents.

In the system currently implemented different representations may significantly affect the total compensation received by each country since, typically:

- Congestion rents are split according to ATC;
- ITC is forfeited and results from a multilateral agreement among the TSOs.

In our example, country C is likely to receive more compensation – in the second setting – for hosting the 400 MWh, since the amount received as a congestion rent is probably higher than the compensation received through ITC.

The choice about the representation of the system therefore depends on relevant conventional elements, but it is likely to have a relevant economic impact on the involved parties. The favor of parties in the more recent debate has shifted towards mechanisms designed on cross-border flows, rather than transits.

Applying a compensation scheme based on cross-border flows implies that all the income sources – congestion rents and transmission tariffs – are determined at the same time and then allocated to the different TSOs, in order to cover the costs generated by both the import/export flows and the transit flows. In this case, the ITC mechanism loses its nature of "compensation" scheme, becoming a cost-sharing scheme.

5.3.1 Review of the Main Approaches to Cost Responsibility Allocation on ITC

This section provides an overview of the main approaches to the allocation of cost responsibility developed in the recent debate on the ITC mechanism. The description of every method is accompanied by an example, in order to clarify its underlying logic.

In general, all these approaches assume that the share of cost responsibility to be assigned to any group of agents can be retrieved from their degree of utilization of the network elements. The AP and MP approaches identify the responsibility of every (injection or withdrawal) node for the utilization of the various HN elements. The WW and the ETSO2005 approaches, instead, are based on an assessment of the overall external utilization of the network elements.

The AP and the WW methods can be applied to cross-border flows as well as to transits. When describing the two approaches, however, we will only focus on their implementation in case of cross-border flows. The remaining MP and ETSO2005 methods can exclusively be applied to cross-border flows and transits, respectively.

In what follows only a description of the responsibility allocation rules under the different approaches is provided (the HN is identified as described in section 5.1[80] and network costs are assumed to be calculated on the basis of regulated costs, as described in section 5.2).

The Present ETSO ITC Model

As briefly anticipated in the introduction, the present model for Inter-TSO Compensation was introduced in 2002 and has been modified step-by-step from year to year. We refer to this model as the ETSO2005 model, since 2005 is the year in which the last modifications have been introduced. The ETSO2005 compensation is limited to transits, in a logic of continuity with the pre-2002 regime, in which compensation was provided through transmission tariffs for cross-border transits.

[80] Depending on the approach, the HN is defined on the basis of transits or cross-border flows.

The compensation entitlement for each ITC Party is defined according to a cost claim, computed as the product between the network costs on the HN of each TSO and a “transit key”; this transit key is the ratio between transits and transits plus consumption. Within this model, transits are computed (at the aggregate level, for each country) as the hourly minimum of export and import from an ITC Party, where export and import are the sum of physical flows on all exporting lines and all importing lines, respectively.

For what concerns the financing of the compensation fund, different rules are applied for the ITC Parties – i.e., the countries involved in the ITC mechanism - and the so-called Edge Countries – i.e. the countries which are part of the ITC mechanism, but have at least one electric border in common with one or more non-ITC Parties.

The Average Participation Method

The Average Participation (AP) method assumes that it is possible to trace the electricity flows over the various elements of the HN (lines, transformers) from the nodes where electricity is produced and injected into the network to the nodes where it is withdrawn and consumed, and *vice versa*.

The first step for tracing power flows over the HN consists in identifying the injection and withdrawal nodes. This is done netting the total generation and consumption at every node. In the second step, the amount of power flowing on every line is identified, by the observation of the actual metered flows in every network component. The network utilization responsibility is then assigned to the injection/withdrawal nodes on the basis of the so-called *“proportionality branching rule”*.

According to this simplifying assumption, in each node of the network the *actual* inflows are allocated proportionally to the *actual* outflows, and *vice versa*. In other words, the proportion of the inflow in a given node which is consumed in the node is computed as the ratio of the consumption (in MWh) and the sum of consumption and the outflow from the node (in MWh). The proportion of the inflow which “passes through” the node is determined as the ratio of the outflow from the node and the sum between consumption and outflow.

The responsibility of the agents located in each node for the power flow over a specific network element is determined as the share of the flow over the element that can be traced backward or forward to that node. The responsibility of each TSO in the utilization of a specific HN element is then determined as the aggregate responsibility of all injection/withdrawal nodes located on the grid of that TSO for all the flow on the HN element in question. This responsibility is then expressed in monetary terms by allocating the cost of the HN element to every TSO in proportion to the TSO's responsibility in utilization.

The monetary compensation that one TSO is required to pay for the use of another TSO's grid is thus the sum of the compensations that the former must pay for the use of all the HN elements of the latter's grid.

EXAMPLE – AP

In order to see how the cost responsibility of every node is determined according to the AP approach, consider the following 5-nodes 5-lines example. In every node, g indicates generation and d load, according to the notation in Chapter I. The flow is expressed in MWh.

Node 1: g=100, d=40
Node 2: g=20, d=150
Node 3: g=40, d=80
Node 4: g=50, d=120
Node 5: g=200, d=20

First, at each node the net amount of generation or demand is computed as the net outcome of generation minus load[81]. If there is more generation than load (as it is the case, for instance, in node 1) the node injects power into the network and is marked as an injection node.

[81] The method can be implemented in different ways. An alternative is to treat generation and load at each node separately, without netting them out.

Instead, if load is larger than generation (as in node 2) the node withdraws power from the network and is marked as a withdrawal node.

Then, the net amount of power injected or withdrawn at every node is traced through the grid until it reaches the loads or generators where it is consumed or produced. The results of these two steps are reported in figure 16.

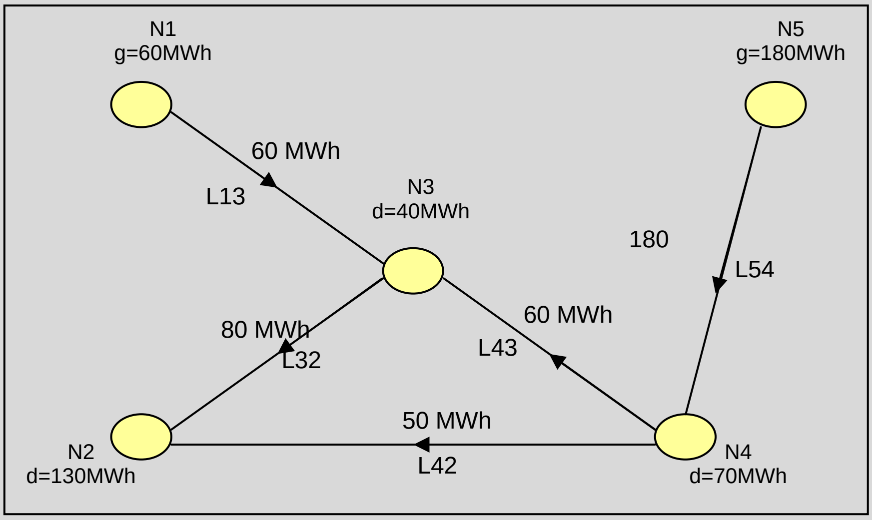

Figure 16 – The five-node five-line system

In the definition of cost responsibilities, injection and withdrawal nodes are dealt with separately. In this way, if desired, there is the option of assigning different global weights to injections (generators) and withdrawals (loads) when computing the final coefficients of allocation of flows.

Consider the allocation to the injection nodes first.

These are nodes 1 and 5 in the figure. Let us start with node 1. When the net injection of 100-40=60 MWh in node 1 reaches node 3, it is split in proportion to the net withdrawal (80-40 = 40 MWh) in node 3 and to the outgoing flow (80) leading to node 2. Therefore, the flow of 60 MWh entering node 3 is split into 60x40/(40+80)=20 MWh that are consumed in node 3 and 60x80/(40+80)=40 MWh that flow into line L32.

This way we have determined that the injection in node 1 is responsible for 60 MWh of flow in line L13 (out of 60) and for 40 MWh of flow in line L32 (out of 80). Node 1 is not responsible for the flows in lines L24, L34 and L45.

The case of the net injection of 200-20=180 MWh from node 5 is a bit more complex. The whole amount (180) flows via line L45, but when it arrives to node 4 it splits in 3 directions:

a) Net withdrawal of 120-50=70 MWh in node 4;
b) The outgoing flow of 50 MWh towards node 2, where it dies;
c) The outgoing flow of 60 MWh towards node 3, where it continues further. This last flow of 60 MWh is split between the net withdrawal of 40 MWh in node 3 and the outgoing flow of 80 MWh towards node 2, where it dies.

The split in node 3 is computed in accordance to the proportionality rule, that is, 60x80/(40+80)=40 MWh goes into line L32 and 60x40/(40+80)=20 MWh remains in node 3.

The same procedure has been applied to the withdrawals, with the only difference that the pattern of the flows is analyzed backward, starting from the withdrawal node.
Consider, for example, node 2. The whole amount of 130 MWh flows to this node through lines L32 and L42. According to the proportionality assumption, an amount equal to 130x80/(80+50) flows on the line L32 and an amount equal to 130x50/(80+50) flows on the line L42. The net withdrawal in node 2 induces an outflow from node 3 equal to 80 MWh. This outflow is responsible for a flow of 80x60/(60+60) on both the L13 and L43 lines[82].

[82] The equality between the flows on the lines L13 and L43 is due to the assumption that the inflows from these two lines into node 3 are equal. This is not a general result. If the two inflows were different, the proportionality rule would have also determined different flows.

The withdrawal in node 2 finally induces an outflow from node 4 equal to 50 MWh (on the L42 line) and equal to 40[83] (on the L43 line). Power flows into node 4 uniquely from the L54 line. The overall utilization for this line is therefore determined as (40x180/180)+(50+180/180). Table 3 summarizes the calculations for all the injection and withdrawal nodes.

	Line 13	**Line 32**	**Line 42**	**Line 43**	**Line 54**
Injections					
Node 1 (60)	60	$60 \cdot \frac{80}{80+40}$	0	0	0
Node 5 (180)	0	$60 \cdot \frac{80}{80+40}$	$180 \cdot \frac{50}{60+50+70}$	$180 \cdot \frac{60}{60+50+70}$	180
Total	**60**	**80**	**50**	**60**	**180**
Withdrawals					
Node 2 (130)	$80 \cdot \frac{60}{60+60}$	$130 \cdot \frac{80}{80+50}$	$130 \cdot \frac{50}{80+50}$	$80 \cdot \frac{60}{60+60}$	50+40
Node 3 (40)	$40 \cdot \frac{60}{60+60}$	0	0	$40 \cdot \frac{60}{60+60}$	20
Node 4 (70)	0	0	0	0	70
Total	**60**	**80**	**50**	**60**	**180**

Table 3 – Allocation of flows in the lines to the injection nodes and to the withdrawal nodes

Table 4 shows these same responsibilities, but expressed in per unit terms. The per unit results are obtained simply dividing the values in each column of table 3 by the total flow in the line corresponding to each column.

	Line 13	**Line 32**	**Line 42**	**Line 43**	**Line 54**
Node 1	1	0.5	0	0	0
Node 2	2/3	1	1	2/3	0.5
Node 3	1/3	0	0	1/3	1/9
Node 4	0	0	0	0	7/18
Node 5	0	0.5	1	1	1

Table 4 – Final allocation (in per unit) of line flows to the injection and withdrawal nodes

[83] 80x60/(60+60).

Before obtaining the final table, we have to decide which weight to assign globally to generators and loads in order to obtain the final allocation factors of lines to nodes. If it is decided that the weights should be the same, we just need to multiply all the numbers in table 4 by 0.5.

However, if the desired global weight is, for instance, 30% to generators and 70% to consumers, then all the per unit values in table XX corresponding to injection nodes have to be multiplied by 0.3 and the per unit values corresponding to withdrawal nodes by 0.7. The results in table 5 correspond to the 50/50 assignment.

	Line 13	Line 32	Line 42	Line 43	Line 54
Node 1	0.5	0.25	0	0	0
Node 2	1/3	0.5	0.5	1/3	0.25
Node 3	1/6	0	0	1/6	1/18
Node 4	0	0	0	0	7/36
Node 5	0	0.25	0.5	0.5	0.5
Total	1	1	1	1	1

Table 5 – Final allocation (in per unit) of lines flows to nodes.

The Marginal Participation Method

The Marginal Participation (MP) method is based on a load flow model, aiming at assessing the extent to which a unit increase in the power injected into or withdrawn from the grid at each node affects the various HN elements. This assessment produces the “marginal participation” of each node in the power flow of over each HN element. The corresponding “total participation” of the node is then obtained by multiplying its marginal participation by the net power injection/withdrawal at the node. The responsibility for the utilization of an HN element – and the corresponding cost – is apportioned to the agents located at the various nodes proportionally to their total participation. The MP method, like the AP method, therefore computes individual responsibilities for market agents. Compensations at the TSO level are computed by aggregating these responsibilities for all the agents located in the TSO area.

Since injections and withdrawals from the grid should always be kept balanced, this method requires matching every increase in the power injection/withdrawal at a node of the HN with a corresponding increase/decrease in the power withdrawal/injection at one or more other nodes. In order to implement the load flow model underlying the MP approach therefore requires the identification of one or more “slack” node(s), that is, node(s) which will offset the marginal increase in the in the power injection/withdrawal simulated at each node. The slack node(s) has to be the same for all injections and withdrawals from the network and its choice is crucial, since the compensation responsibilities and entitlements vary a lot with different slack node(s).

EXAMPLE – MP

Consider the following example, based on the 5-node 5-line previously introduced model. In order to clearly identify how the choice of the slack node affects the result, two alternative situations are analyzed.

In the former we choose node 1 as the slack node, while in the latter node 5 is selected.

Situation 1: Node 1 is chosen as the slack node

To calculate the participation of the agent located in node 2, we study the impact on the network of a marginal increase in load in such node (offset by an equal increase in the generation at the slack node)[84].

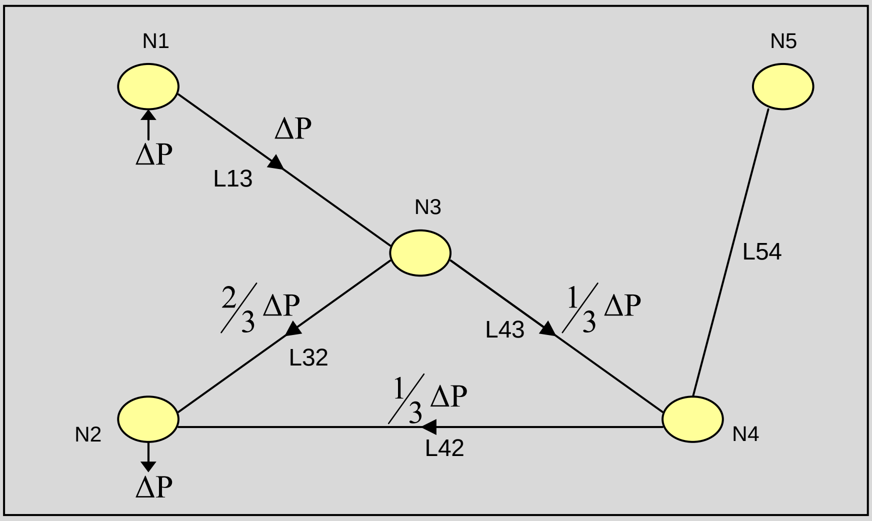

Figure 17 – Marginal participation of node 2 when node 1 is chosen as the slack node

The participation of the load in node 2 is given by the product of the amount of power consumed by the load and the portion of ΔP flowing through each line in the incremental transaction simulated, taking into account the flow direction. Results are shown in the table below, where P_{L13}^{N2} indicates the participation of the agent in node 2 in the use of the line L13[85].

P_{L13}^{N2}	P_{L32}^{N2}	P_{L43}^{N2}	P_{L42}^{N2}	P_{L54}^{N2}
$130 \cdot \frac{\Delta P}{\Delta P} = 130$	$130 \cdot \frac{2/3 \Delta P}{\Delta P} = 87$	$130 \cdot \frac{-1/3 \Delta P}{\Delta P} = -43$	$130 \cdot \frac{1/3 \Delta P}{\Delta P} = 43$	$130 \cdot \frac{0}{\Delta P} = 0$

Table 6 –Participation of node 2 when node 1 is chose as the slack node.

[84] Note that in this example we assume that the resistance is the same on L23, L34 and L24.
[85] The participation is computed multiplying the results of the load flow model (i.e. the marginal participation), represented in figure 17, by the net withdrawal in node 2.

To calculate the participation of the agent located in node 5, an incremental injection (since node 5 is an injection node) of ΔP located in node 5 has to be simulated, together with an increment in load in the slack node 1.

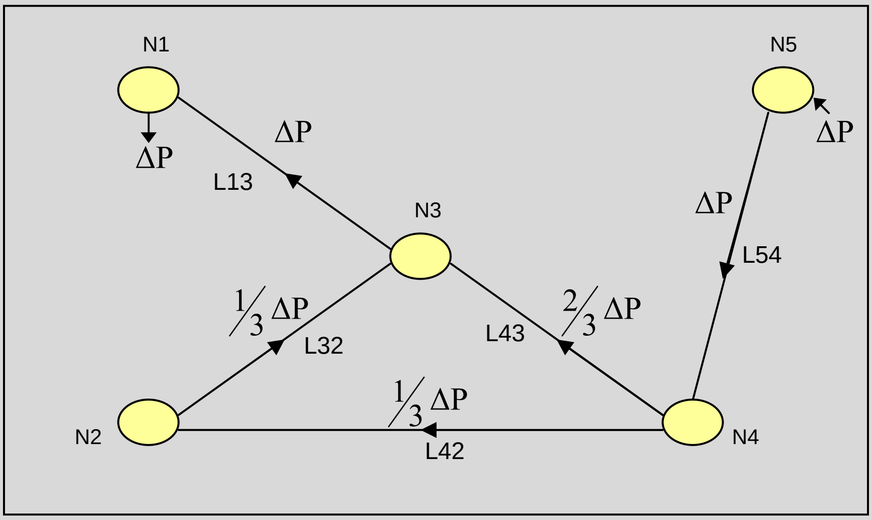

Figure 18 – The marginal participation of node 5, when node 1 is chosen as the slack node.

In this case, the results are:

P_{L13}^{N5}	P_{L323}^{N5}	P_{L43}^{N5}	P_{L42}^{N5}	P_{L54}^{N5}
$80 \cdot \frac{-\Delta P}{\Delta P} = -180$	$180 \cdot \frac{-1/3 \Delta P}{\Delta P} = -60$	$130 \cdot \frac{2/3 \Delta P}{\Delta P} = 120$	$180 \cdot \frac{1/3 \Delta P}{\Delta P} = 60$	$130 \cdot \frac{\Delta P}{\Delta P} = 180$

Table 7 –Participation of node 5 when node 1 is chose as the slack node.

Calculating the participation of the agent located in node 3 of the network grid, we obtain:

P_{L13}^{N3}	P_{L32}^{N3}	P_{L43}^{N3}	P_{L42}^{N3}	P_{L54}^{N3}
$40 \cdot \frac{\Delta P}{\Delta P} = 40$	$40 \cdot \frac{0}{\Delta P} = 0$	$40 \cdot \frac{0}{\Delta P} = 0$	$40 \cdot \frac{0}{\Delta P} = 0$	$40 \cdot \frac{0}{\Delta P} = 0$

Table 8 –Participation of node 3 when node 1 is chose as the slack node.

And, finally, calculating the participation of the agent located in node 4, we obtain:

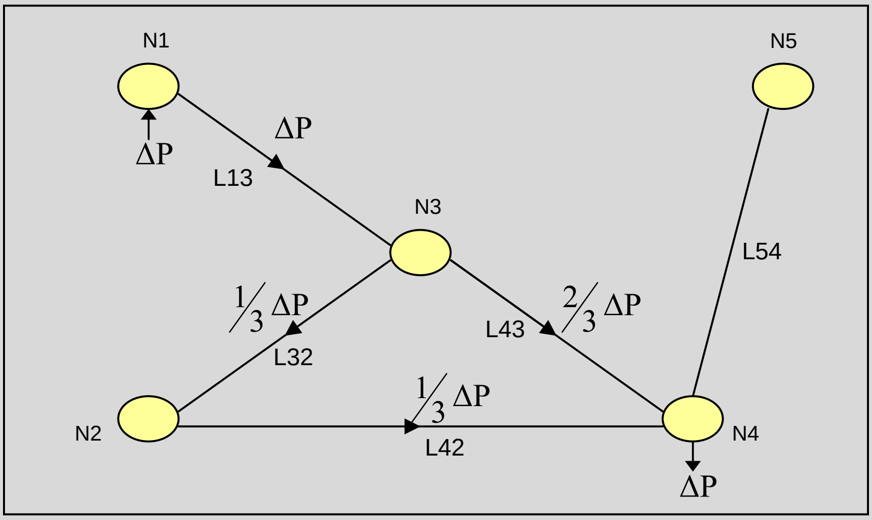

Figure 19 - The marginal participation of node 4, when node 1 is chosen as the slack node

P_{L13}^{N4}	P_{L32}^{N4}	P_{L43}^{N4}	P_{L42}^{N4}	P_{L54}^{N4}
$70 \cdot \frac{\Delta P}{\Delta P} = 70$	$70 \cdot \frac{1/3 \Delta P}{\Delta P} = 23$	$70 \cdot \frac{-2/3 \Delta P}{\Delta P} = -47$	$70 \cdot \frac{-1/3 \Delta P}{\Delta P} = -23$	$70 \cdot \frac{0}{\Delta P} = 0$

Table 9 –Participation of node 4 when node 1 is chose as the slack node

Situation 2: Node 5 is chosen as the slack node

In case node 5 is selected as the slack node, we obtain a totally different result for the calculation of the responsibility of single agents in the network use, as shown in the following calculation.

The participation of node 1 in the network use is:

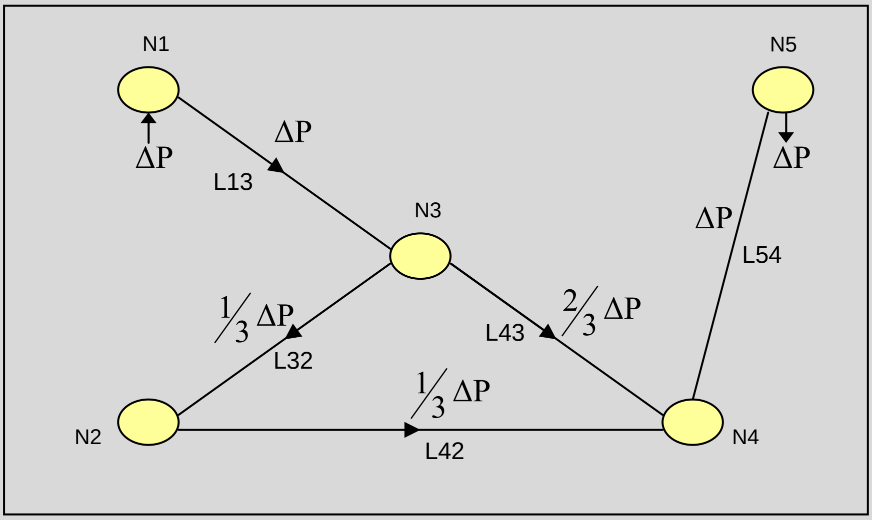

Figure 20 - The marginal participation of node 4, when node 5 is chosen as the slack node

P_{L13}^{N1}	P_{L32}^{N1}	P_{L43}^{N1}	P_{L42}^{N1}	P_{L54}^{N1}
$60 \cdot \frac{\Delta P}{\Delta P} = 60$	$60 \cdot \frac{1/3\,\Delta P}{\Delta P} = 20$	$60 \cdot \frac{-2/3\,\Delta P}{\Delta P} = -40$	$60 \cdot \frac{-1/3\,\Delta P}{\Delta P} = -20$	$60 \cdot \frac{-\Delta P}{\Delta P} = -60$

Table 10 –Participation of node 1 when node 5 is chose as the slack node

In case of node 2:

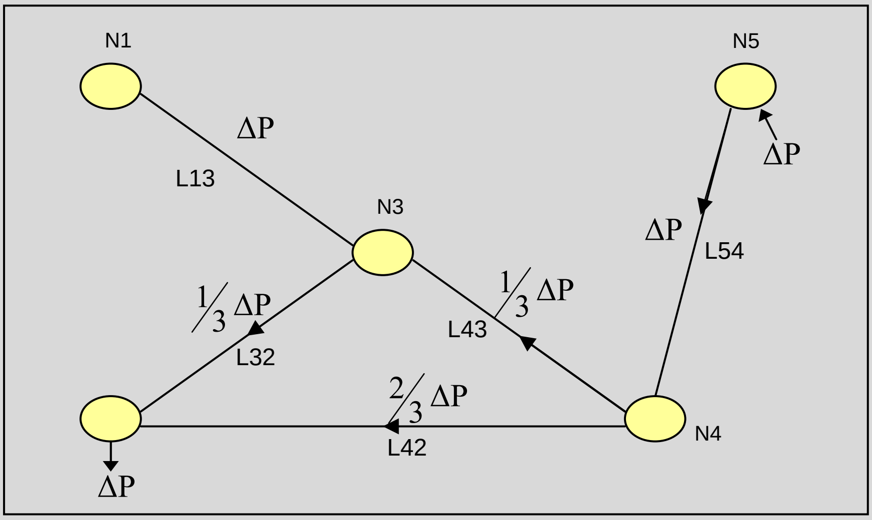

Figure 21 - The marginal participation of node 2, when node 5 is chosen as the slack node

The case of node 3 is very similar and we obtain:

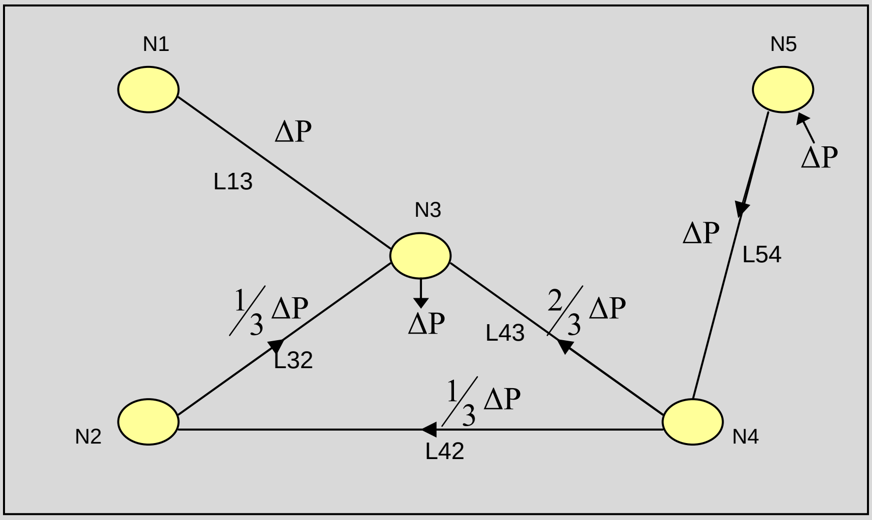

Figure 22 - The marginal participation of node 3, when node 5 is chosen as the slack node

P_{L13}^{N3}	P_{L32}^{N3}	P_{L43}^{N3}	P_{L42}^{N3}	P_{L54}^{N3}
$40 \cdot \frac{0}{\Delta P} = 0$	$40 \cdot \frac{-1/3\Delta P}{\Delta P} = -13$	$40 \cdot \frac{2/3\Delta P}{\Delta P} = 27$	$40 \cdot \frac{1/3\Delta P}{\Delta P} = 13$	$40 \cdot \frac{\Delta P}{\Delta P} = 40$

Table 12 –Participation of node 3 when node 5 is chose as the slack node

In case of node 4:

P_{L13}^{N4}	P_{L32}^{N4}	P_{L43}^{N4}	P_{L42}^{N4}	P_{L54}^{N4}
$70 \cdot \frac{0}{\Delta P} = 0$	$70 \cdot \frac{0}{\Delta P} = 0$	$70 \cdot \frac{0}{\Delta P} = 0$	$70 \cdot \frac{0}{\Delta P} = 0$	$70 \cdot \frac{\Delta P}{\Delta P} = 70$

Table 13 –Participation of node 4 when node 5 is chose as the slack node

Table 14 reports the results of the MP model obtained in the alternative assumptions about the choice of the slack node.

	Line 13		**Line 32**		**Line 43**		**Line 42**		**Line 54**	
	Slack node N1	Slack node N5	Slack node N1	Slack node N5	Slack node N1	Slack node N5	Slack node N1	Slack node N5	Slack node N1	Slack node N5
Node 1	0	60	0	20	0	-40	0	-20	0	-60
Node 2	130	0	87	43	-43	43	43	87	0	130
Node 3	40	0	0	-13	0	27	0	13	0	40
Node 4	70	0	23	0	-47	0	-23	0	0	70
Node 5	-180	0	-60	0	120	0	60	0	180	0

Table 14 – Participation of the five nodes in the utilization of the various lines in the two situations

The results in table 14 clearly show that the choice of the slack node is likely to affect the results of the MP method. In particular, participations in the use of the lines may be much larger than the actual flow on the line (see for example the case of the line L13 when node 1 is chose as the slack node). As a final remark, it is interesting to notice that these results do not take into account the capacity limits of the lines composing the network, thus generating the risk of inconsistencies between compensations and the network capacity.

The With-and-Without Methods

The With-and-Without (WW) method and the With-and-Without method applied to transits (WWT) are based on a comparison between the flows on the HN of each TSO area under two different scenarios. The first scenario corresponds to the actual system operation and the second results from excluding either cross-border flows (in the WW method), or transit flows (in the WWT method).

Transits can be defined and allocated to the different interconnection lines in various ways. The approach considered in this example is the same adopted by the current ETSO compensation mechanism previously described[86]. Transits can then be allocated to the different interconnection lines carrying flows in import and export directions in proportion to the measured cross-border flows on these lines in each direction. The flows over each of the HN elements in the two scenarios are computed using a load flow model. The resulting differential flow (between the two scenarios) over each HN element is attributed to cross-border flows/transits and therefore to external agents.

The compensation a TSO is entitled to receive is expressed in aggregate terms, without distinguishing between the utilization (and cost) responsibility of the other TSOs involved, and can be computed in several ways. The model could define the total amount of compensation entitlement for every single element of the network[87] – in this case the total compensation is obtained summing the amount of compensation of all the network elements – or it can be defined directly considering the entire HN[88].

By summing all compensations entitlements of all TSOs, the monetary value of the total compensation fund is obtained. Responsibility for contributing to this fund is then allocated to the different TSOs on the basis of their responsibility in generating cross-border flows/transits.

[86] Hourly transits are defined as the minimum between the total hourly in-flow and the total hourly outflow.
[87] According to this solution, compensation is defined as the network element cost multiplied by the ratio of the flows attributable to external utilization and the total flows on that element.
[88] According to that solution, total compensation is defined as a proportion of the total cost of the HN, where the factor of proportionality is the ratio between utilization of the HN by external flows and the total (actual) utilization of the HN.

EXAMPLE – WWT

The functioning of the WWT method is explained with the help of a simplified topology of 6 nodes connected by 6 lines.

In particular:

- Country A includes nodes 1,2,3 and the lines L12, L23 and L31;
- Country B includes node 4 and line L14;
- Country C includes node 6 and line L62;
- Country D includes node 5 and line L35.

The WWT method requires comparing the flows on the HN of each TSO area for two different scenarios.
The first scenario corresponds to the actual system operation (scenario 1). The second one results from excluding transit flows (scenario 2).

Using the standard definition, transits are calculated as the minimum between the total import flows and the total export flows. In our example the transit through country A is 3000 MWh, i.e. the minimum between 5000 MWh (import) and 3000 MWh (export).

As previously explained, there are several ways to compute the compensation that a TSO is entitled to receive.
In this example, we compute the compensation the TSO is entitled to receive as a proportion of utilization of the entire HN by external flows. Also the usage of the network of a country could be calculated in many alternative ways. In this example network usage is defined as the sum of the length of the lines multiplied by the flows over these lines.

Thus, the network usage of country A corresponding to the actual operation of the system would be:

$$U = l_{12} \cdot I_{12} + l_{23} \cdot I_{23} + l_{13} \cdot I_{13} = l_{12} \cdot 2000 + l_{23} \cdot 3500 + l_{13} \cdot 1500$$

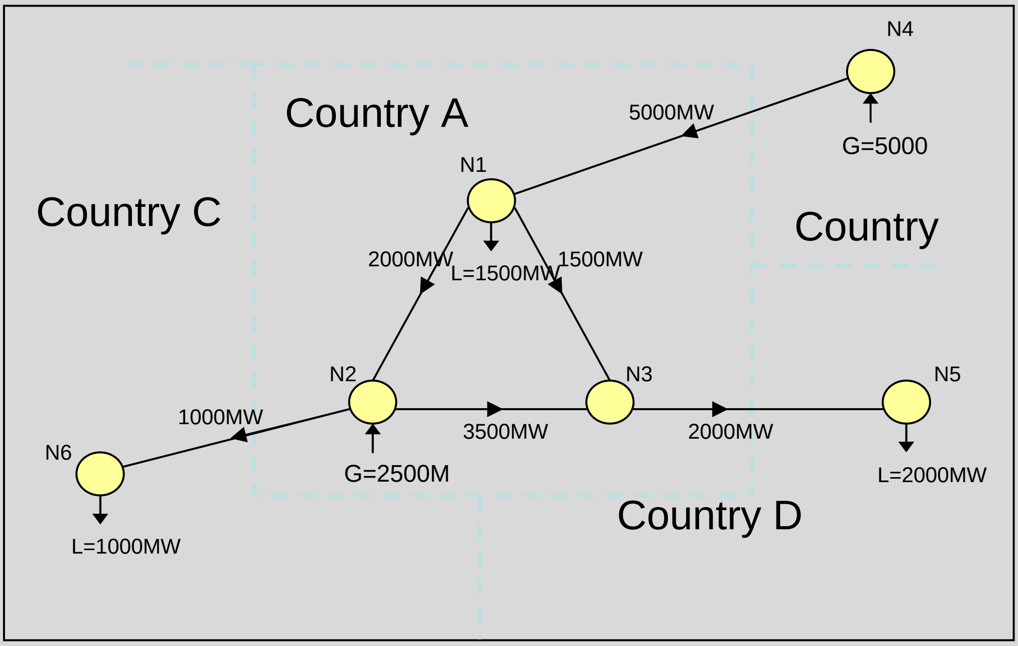

Figure 23 – The three-country system (scenario 1)

In order to calculate the compensation due to country A, a fictional scenario where transit flows have been removed has to be created. Once the transit has been removed, export flows are zero and import flows amount to 2000 MWh.

The following step consists in determining the flow on every cross-border line in this new setting. This is done, in this example, distributing the transit among cross-border lines in proportion to the size of the original flow on each of these lines[89]. Thus, flows on the cross-border lines in the without transit scenario are 0 for lines L26 and L35 and 2000 MWh for line L41.

[89] Different transits allocation rules could be implemented.

Lastly, cross-border flows are changed to these new values and a new load flow is run to determine the flows on the internal lines of the country in this fictitious scenario. This new situation is represented in figure 24.

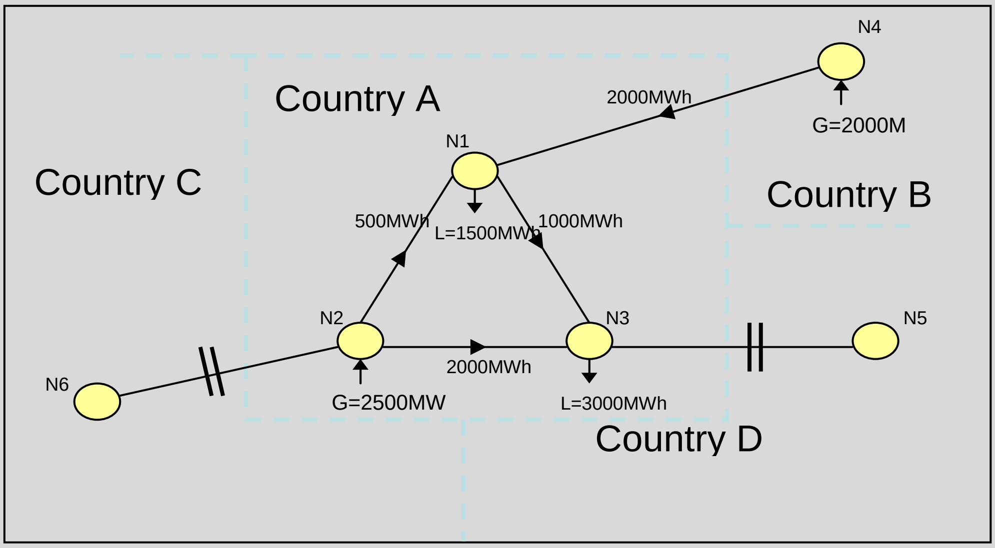

Figure 24 – The three country system in the "without" scenario (scenario 2)

The usage of country A network in the scenario without transits can only be attributed to country A and is computed as:

$$U = l_{12} \cdot I_{12}^{wt} + l_{23} \cdot I_{23}^{wt} + l_{13} \cdot I_{13}^{wt} = l_{12} \cdot 500 + l_{23} \cdot 2000 + l_{13} \cdot 1000$$

Compensation due to country A is obtained as the product of the fraction of total network usage for country A that is attributable to transits and the cost of the horizontal network of country A, that is:

$$Comp_A = Cost(HN_A) \cdot \frac{U - U^{wt}}{U}.$$

Responsibility for contributing to this fund is then allocated to the different TSOs on the basis of their responsibility in generating transit flows.

This responsibility can be determined in various ways. One example could be to consider the net import/export (net(I/E)) from the grid of each TSO as a proxy to the transit flows generated by the injections into and withdrawals from that grid.

In this way, the contribution of each of the countries in the system to the compensation due to country A (assuming there are no other countries) would be:

$$Payment_{AA} = Comp_A \cdot \frac{net(I/E)_A}{net(I/E)_A + net(I/E)_B + net(I/E)_C + net(I/E)_D} =$$

$$= Comp_A \cdot \frac{2000}{2000 + 5000 + 1000 + 2000}$$

$$Payment_{BA} = Comp_A \cdot \frac{5000}{2000 + 5000 + 1000 + 2000}$$

$$Payment_{CA} = Comp_A \cdot \frac{1000}{2000 + 5000 + 1000 + 2000}$$

$$Payment_{DA} = Comp_A \cdot \frac{2000}{2000 + 5000 + 1000 + 2000}$$

5.3.2 Issues on the Network Cost Responsibility Methods

The process of network cost responsibility allocation is extremely complex and all the previously described solutions have strengths and weaknesses, mainly due to their underlying assumptions.

The analysis of these weaknesses shows that, in general, the approaches providing an aggregate measure of compensation – WW and ETSO2005 – prove to have a poorer performance than the approaches identifying the cost-responsibility for the single nodes – AP and MP. When the aggregate measure of external responsibility for the usage of a TSO network is obtained, in fact, ad hoc criteria need to be used to allocate it to the other TSOs involved.

The ETSO2005 method bases responsibility allocation on "transit keys", while in our example for the WW method we used net import/export from the grid of each TSO as a proxy to the transits he caused. Different criteria can be considered more or less suitable, depending on the situation, but the key issue emerging from this discussion is the arbitrariness of the choice, potentially affecting the results.

With respect to the ETSO2005 approach, the reliability WW method is also affected by the arbitrariness in the definition of the "without" scenario, which reduces the transparency of the solution. The definition of the without scenario requires specifying a series of conventions, the modification of which would crucially affect the final result[90].

The AP and MP approaches, assigning responsibility directly to the individual nodes, do not require the implementation of arbitrary allocation rules which could reduce the transparency of the results. Also the responsibility allocation in these two approaches relies, however, on a set of assumptions mining its validity.

As explained in the previous section, the AP approach allocates the cost responsibility to the agents according to a proportionality rule, on the basis of the observed power flows on the network. The idea of allocating the responsibility of every node in the utilization of the various network elements proportionally to the net injections/withdrawals of the node is absolutely arbitrary and leads to results which do not strictly reflect the laws of physics governing the distribution of electricity flows on the grid.

The solution provided by the MP approach is indeed consistent with the physical laws of physics, since it is the result of a load flow model. This solution, however, crucially depends on the arbitrary choice of the slack node(s).

Moreover, the MP approach is by definition focused on the incremental effects power injections/withdrawals at the different nodes. This may lead to power flows – and therefore ITCs – inconsistent with the physical reality of the network. This may occur, for example, in the case of two TSO areas which are only weakly linked. The MP method may assign significant responsibility for the utilization of the network in one TSO area to generators and loads in the other TSO area, if this would be the case when incremental injections and withdrawals are considered. This result, when applied to total injections ad withdrawals, may be at odds with the physical capability of the networks, as the limited

[90] Consider, for instance, the assumption about the distribution of transits on the network that we had to make in our example.

capacity of the interconnector may restrict the extent to which generators and loads in one TSO area can use the network in the other area.

5.4 Concluding Remarks

The current ITC mechanism is the result of market arrangements developed independently in different areas and in different times. It is widely based on negotiations among (couples of) European States and is difficult to root in any theoretical framework.

Our analysis of the issues surrounding the design of the ITC system has highlighted the arbitrary nature of any method to assess the impact of cross-border transactions on the utilization (or relief) of each transmission network element. As a consequence, there is no obvious way to allocate the cost of each network component to a transaction or to a Country’s net surplus/deficit.

The same ambiguity features in the assessment of the contribution of each network element to making cross-border transactions possible. The incremental value of each network element, that can be assessed based on the value of electricity at each location, is by no obvious base to allocate congestion rents.

Coordinating the development of the transmission networks via mechanistic rules for international cost-sharing based on some measure of utilization of the infrastructures for the purpose of cross-border transactions in likely to cause inefficiencies.

A practical solution to the problems related to cross-border transmission capacity cost-allocation, fully consistent with the market integration objective, could be based on some form of pooling of the standard costs of the European States’ transmission grids. In this approach, the total standard cost of the Horizontal Network would be computed and shared among all the European customers. Congestion rents resulting from market based allocation of transmission rights would be used to cover the shared part of the total Horizontal Network standard cost. The remaining part would be split among the different countries according to, for example, total electricity consumptions. National regulators would then be free to design national transmission tariffs allowing collection of the revenue target (as well as of the revenues related to non-standard portion of their network’s costs).

In that way, the total cost of the European Countries, to a large extent in the same way the cost of each Country's transmission system is currently spread among each State's consumers. The reference to standard costs limits the scope for cross-subsidization, while preserving a certain degree of independence in each Country's procurement or network design strategy.

Pooling of the (standard) costs of existing network resources would lead the way to coordination in network development decisions.

6. References

AEEG. "Consultation Document on the Inter_TSO Compensation Mechanism and on Harmonization of Transmission Tarification". *Response to Consultation, Autorità per l'Energia Elettrica e il Gas*.

CW MC Project (August 2008), "Implementation Study. Technical Report", Available from: http://www.apxgroup.com/index.php?id=186.

Brattle Group (March 2008). "Models of Inter-regional Transmission Charging". *Report, The Brattle Group*.

Camacho L. O., Pérez-Arriaga I. (2007). "Comparison of Several Inter-TSO Compensation Methods in the Context of the Internal Electricity Market of European Union". *Energy Policy, 35, 2379-2389*.

Dupuy M. (2008), "Electricity markets: Balancing mechanisms and congestion management", *Master's thesis, Royal Institute of Technology, Stockholm, Sweden*.

ERGEG (2008), "ERI coherence and convergence report", *Conclusion paper, ERGEG*.

ERGEG (2008a), "ERGEG Regional Initiatives Annual Report", *Technical report, ERGEG*.

ETSO (January 2009). "Firmness of cross-border capacities and compensation schemes in case of curtailments". *Position paper, ETSO*.

ETSO (April 2008). "Reference Model for Cross-border Intraday Markets". *Technical Paper, ETSO*.

ETSO (June 2007). "Firmness of cross-border trade and transmission rights in Europe". *Position paper, ETSO*.

ETSO (March 2007). "Regional Flow-Based allocations. State-of-play". *Final paper, ETSO*.

ETSO (April 2006). "Transmission Risk Hedging Products". *Background paper, ETSO*.

ETSO (2005). "The Current ETSO ITC Model and Possible Development". *Technical Report, ETSO*.

ETSO (October 2001). "Capacity Publications and Procedures". *Final Report, ETSO*.

ETSO (October 2001). "Procedures for Cross-border Transmission Capacity Assessments". *Technical Report, ETSO*.

ETSO (April 2001). "Co-ordinate Auctioning". *Final Report, ETSO*.

ETSO (April 2001). "Definitions of Transfer Capacities in Liberalized Electricity Markets". *Final Report, ETSO*.

ETSO (March 2000). "Net Transfer Capacities (NTC) and Available Transfer Capacities (ATC) in the Internal Market of Electricity in Europe (IEM)". *Information for User, ETSO*.

European Commission, DG TREN. "Consultation Document on the Inter-TSO Compensation Mechanism and on Harmonization of Transmission Tarification". *Consultation document, DG TREN, European Commission*.

EuroPEX and ETSO (April 2008). "Development and implementation of a coordinated model for regional and inter-regional congestion management". *EuroPEX and ETSO Interim report*.

EuroPEX (September 2007). "Importance of Cross-border Intraday Markets and the Role of Power Exchanges". *Position paper, EuroPEX.*

EuroPEX and ETSO (2004), "Flow-based market coupling", *EuroPEX and ETSO Technical report*.

EuroPEX (July 2003). "Using Implicit Auctions to Manage Cross-Border Congestion: Decentralized Market Coupling". *Conclusion Paper, EuroPEX.*

Florence School of Regulation (2005). "A study on the Inter-TSO Compensation Mechanism". *Final report, Florence School of Regulation*.

Frontier Economics and Consentec. "Study on the further issues relating to the Inter-TSO Compensation Mechanism". *Final report, Frontier Economics and Consentec*.

Lyons K., Fraser H., Parmesano H. (2000). "An introduction to financial transmission rights". *The Electricity Journal, 13(10), 31-37*.

Meeus L. (2006), "Power Exchange - Auction trading platform design", *PhD thesis*, *Katholieke Universiteit Leuven, Belgium*.

Norwegian Industry and Consumer Association (March 2007). "The ETSO sponsored and EU recognized Inter-TSO Compensation (ITC) Scheme: Jeopardizing Incentives for Network Investment and Efficient Network Operation". *Position Paper. Norwegian Industry and Consumer Association*.

Universidad Pontificia Comillas (November 2002). "Cost Components of Cross-border Exchanges of Electricity". *Study for the European Commission, DG TREN, Final Report*.

Chapter IV

Electricity Regional Initiatives

1. Introduction

Directive 96/92/EC and Directive 2003/54/EC laid down the general conditions for the creation of a single Internal Electricity Market in Europe. They refrained, however, from indicating a common design for electricity markets. As of today, the Internal Electricity Market is still divided into sub-markets, which often coincide with national borders and present a low level of harmonization.

Most of the electricity is traded bilaterally in forward and over-the-counter (OTC) type of markets. Daily and infra-daily trading occur in Power Exchanges (PX) and Transmission System Operators (TSOs) procure ancillary services on balancing markets. Apart from a few exceptions, there is no link between national PXs or balancing markets. Furthermore, even though Regulation 1228/2003 introduced harmonized principles on cross-border congestion management, different kinds of arrangements still characterize the allocation of available capacity on interconnectors between national transmission systems.

A higher extent of market integration would offer potential benefits in terms of economic efficiency and security of supply.

At present, the adoption of common, harmonized rules at a regional level is considered as a necessary, intermediate step in the development the Internal Electricity Market. With the precise objective of facilitating the development of regional markets, in 2006 the European Regulators' Group for Electricity and Gas (ERGEG), in cooperation with the European Commission, launched the European Regional Initiatives (ERIs).

In this part of the report we analyze, within the framework of the ERIs, the conditions for a gradual integration of the Italian market with neighboring countries. In particular, we address the following topics:

- the ERGEG Regional Initiative framework;
- the market coupling experiences in Europe;
- the basic conditions for implementing a day-ahead (and intra-day) market coupling in the Central-South region;
- the current framework for capacity allocation on the Italian borders.

Relying on the institutional and technical reports, as well as on the available literature, we provide insights on the key economic, technical and institutional issues of the proposed integration models.

2. Electricity Regional Initiatives

The European Regulators Group for Electricity and Gas (ERGEG) in spring 2006 launched the Electricity Regional Initiatives (ERIs), with the overall objective of facilitating the development of regional energy markets. The final goal of the ERIs, to be achieved working in co-operation with the stakeholders, is to remove barriers to trade and competition and to create a single electricity market in Europe.

The ERIs consist of seven Regional Energy Markets (REMs) projects. Each REM brings together regulators, TSOs and PXs, Member States, the European Commission and other stakeholders (traders, power producers) to focus on developing and implementing solutions to remove barriers which may be hindering the growth of competitive electricity markets and to make real progress in delivering a competitive liberalized single European market.

The seven REMs are:

- *Central-West*: Belgium, France, Germany, Luxembourg, The Netherlands;
- *Northern*: Denmark, Finland, Germany, Norway, Poland, Sweden;
- *France-UK-Ireland*: France, Ireland, United Kingdom;
- *Central-South* : Austria, France, Germany, Greece, Italy, Slovenia;
- *South-West*: France, Portugal, Spain;
- *Central-East*: Austria, Czech Republic, Germany, Hungary; Poland, Slovakia, Slovenia;
- *Baltic*: Estonia, Latvia, Lithuania.[91]

REMs have similar objectives - integrating fragmented national electricity markets into regional markets - but their priorities reflect their different regional concerns. Since 2006, each Regional Co-ordination Committee (RCC) has created Implementation Groups (IGs), that include TSOs and PXs, and Stakeholder Groups (SGs), which have a wider participation and include also market operators. In addition, RCCs interact with (reports to) the EU

[91] The Regional Initiatives need to be flexible as the number of EU member states grows. In 2007 Bulgaria e Romania joined the EU, creating the need for an 8th region, South-East Europe. As for now the two countries are given the role of observers in the Central-East and Central-South regions. They have also signed the Energy Community Treaty for the application of common congestion management methods.

Commission, Member States, and the Florence Forum. Most of the regions have held workshops and meetings, and launched consultations to tackle particular issues.

Consistency and convergence are important issues for the ERIs. At the time ERIs were launched, there was a number of other initiatives focused on market integration and not under the direct leadership of European regulators.[92] Regulators play an important role in ensuring that different regional developments remain coherent, in the long run, with the vision of a single market, avoiding contrasting priorities and solutions, and accelerating the roll-out of effective solutions developed in one region, to others (ERGEG, 2007).

Section 2.1 provides a summary of priorities and main achievements for each region. Section 2.2 addresses the main issues concerning coherence and convergence in regional development.

2.1 Priorities and developments to date

Progresses have been made in the past two and a half years in all the regions. ERGEG has indicated that most regions have chosen the same topics to work on and to address problems in a similar way (ERGEG, 2008a). Before entering in the details of regional priorities and achievements to date, it is important to recall the preferred methods for congestion management indicated by ERGEG:

- a common transmission model with regionally-calculated PTDF matrix, especially for regions with highly meshed networks;
- a single auction platform with harmonized auction rules, IT interface, and products for long and medium term allocation;
- a market coupling model for the day-ahead time frame;
- an intra-day mechanism, possibly based on continuous trading.

[92] They included the Pentalateral Energy Forum for electricity (initiative of the governments in the Central-West Region), the Mibel Project (for the establishment of an Iberian market), the planned coupling on the German-Danish interconnectors as well as the Nor-Ned cable (Norway-Netherlands) and the day-ahead market coupling on the Dutch-Norwegian interconnectors.

2.1.1 Central-West

A TriLateral Market Coupling (TLC) for the day-ahead market is in place since November 2006 between France, Belgium and The Netherlands (see also Section 3.2).

Market integration in this region is strongly supported by the work developed in the framework of the Pentalateral Energy Forum. A Memorandum of Understanding (MoU) was signed in June 2007, between governments, regulators, power exchanges, TSOs and the electricity associations of Belgium, Luxembourg, the Netherlands, Germany and France.

The signatories of the MoU agreed to design and implement a flow-based market coupling between the five countries with January 2009 as a target date. The latest development is the publication of an Implementation Study on the design of a market coupling solution for the Central-West European region by the CW MC Project (2008), that brings together all TSOs and PXs of the region. The first target is the launch of an Available Transfer Capacity (ATC) based market coupling for the CWE region, expected for the end of 2009. After a period of parallel running in flow-based coupling, there is going to be a final switch to flow-based coupling. Note that such regional market would be larger than PJM (currently the largest market in the world) - with a 191 GW peak load and 1246 TWh annual consumption - although less fine tuned and harmonized in terms of design (Meeus and Belmans, 2008).

The agreement called for strict collaboration between TSOs, and PXs. They had to examine the opportunity in setting up a joint-TSO and/or PX company. TSOs have started working on the harmonization of the explicit auction rules and, together with regulators, have defined a common position. Moreover, the European Commission has approved, in August 2008, under the EU merger regulation, the creation of a joint venture, Capacity Allocation Service Company (CASC), for the allocation of cross-border transmission capacity in the region.[93] There are three other noteworthy initiatives: (i) the creation of a Holding de Gestionnaires de Réseaux de Transport (HGRT), owned by RTE (53.33%), Tennet (24.5%) and Elia (22.17%), and shareholder of Powernext (51%); (ii) the creation, announced in

[93] The European Commission has concluded that the proposed merger would not raise any competition concerns: *"the examination […] showed that there were no horizontal overlaps between the transmission networks of the participating electricity transmission system operators. Moreover, given the limited size and scope of the joint venture's activities, its creation would not enhance the risk of any coordinated behaviour between the verti of the parties to the joint venture belong",* IP/08/1259, available at http://ec.europa.eu/comm/competition/mergers/cases/index.

September 2008 by the French RTE and the Belgian Elia of a regional coordination center (also the German Vattenfall Europe Transmission has expressed an interest in the project); (iii) the merger of EEX and Powernext: a new entity, based in Paris, is expected to operate spot trading for France, Germany, Austria and Switzerland; another platform, based in Lipsia, will operate the derivative markets.

Other priorities for the region include intra-day cross-border allocation and transparency (a transparency report was published in 2007: it indicates how the transparency requirements of the Congestion Management Guidelines should be implemented).

Finally, the parties involved in the MoU have agreed to take into consideration the coherence between their initiative and the initiatives for market coupling with the Nordic region with the objective to find compatible solutions (NorNed cable and the EMCC coupling of Germany and Denmark – see below).

2.1.2 Northern

A Nordic Power Market (Nord Pool) has been in place since the 1990s and it includes Norway, Sweden, Denmark, Finland and a German region (see also Section 3.1).

Since October 2005, when Nord Pool Spot (NPS) opened a new price quotation area in Germany, the Kontek bidding area, trading on the East Denmark-German border, was managed through market splitting. As for the West Denmark-Germany interconnector, explicit auctions were held on an annual, monthly and daily basis. In addition, a market coupling approximation, known as Cross-Border Optimization (CBO), was introduced to improve cross-border efficiency on this interconnector (see Part II of this report).

Very recently, Energinet.dk, E.ON Netz, VE Transmission, EEX and Nord Pool Spot have created a central Auction Office - European Market Coupling Company (EMCC) - to operate a day-ahead market coupling on the Denmark-Germany interconnection (EMCC, 2008). The European Commission has approved the creation of the EMCC in August 2008 and trading started on September 29th 2008. Unexpected deviations in flow and price calculations between the EMCC and the Nord Pool Spot system have led to non-optimal results regarding the utilization of the capacity between market areas. Furthermore, in some cases the power exchange prices for the Danish and German market areas did not correspond to the direction of the flows on the cross-border lines determined by EMCC's bids/offers. Given these unexpected results, EMCC and its project partners jointly decided to suspend market coupling temporarily (EMCC, 2008).

Coupling of the intra-day markets on the same interconnectors is also scheduled, and it will take the form of a continuous trading platform (the intra-day platform DE-DK was operational on June 26th 2008).

The 700 MW cable interconnecting Norway with the Netherlands – the NorNed HVDC link – started operation in May 2008. As of today, the interconnector capacity is allocated through explicit daily auctioning of the ATC.[94] No annual or monthly auction will be organized. Explicit daily auctioning is considered a temporary solution: TenneT (the TSO in the Netherlands) and Statnett (the TSO in Norway) intend to introduce a market coupling solution based on an implicit auction.

Capacity allocation issues in this region focus also on two merchant interconnectors (Baltic Cable and Swe-Pol Link). A dialogue between transmission capacity owners, regulators, TSOs, and PXs started concerning how to make this capacity available to the market, considering also the constraint due to the lack of liquidity on the Polish market.

Another significant progress in the region regards the harmonization of transparency rules. Regulators have published a report which outlines the type and timing of data to be published on the website of TSOs and PXs; these rules are to be implemented by 2008. This report is considered an example of "good practice": two other REMs (Central-East and Central-West) are to adopt a similar approach.

[94] Participants in the auction have to be a Programme Responsible Party in the Netherlands and/or a Balance Responsible Party in Norway. The product being auctioned consists of hourly Physical Transmission Right. Statnett and TenneT inform the market about the capacity for the next day at 9:15 am; bids can be filed until 09.45am and the results of the auction are published no later than 10.15am. Nominations on TenneT's side (the Netherlands) must be made before 2.00pm; after that time, the "Use it or lose it" principle applies. For all imports into the Netherlands the resulting trade on the NorNed link must take place on the APX day-ahead spot market. In Norway, both imports and exports must be submitted to the Nord Pool Spot exchange. The allocated capacity is binding for the TSO's (except for Force Majeure) and any imbalances will be taken by the TSO's. Each participant will receive a settlement as if the cable was in operation and thus will have no imbalance due to faults in the cable (NorNed, 2008).

2.1.3 France-UK-Ireland

A Single Electricity Market (SEM) of the Republic of Ireland and Northern Ireland was created on November 1st, 2007.

The main priorities in this region are three: improving arrangements for cross-border balancing, congestion management and market transparency.

As for cross-border balancing, TSOs have published a timetable for the implementation of a balancing model on the French-English interconnector by the summer of 2009. The region foresees the development of a common IT system for capacity allocation, to be ready by the end of 2008. The region has formulated a menu of three options in compliance with the mandatory Congestion Management (CM) Guidelines and the coordinated allocation of interconnection capacities. TSOs are free to choose any of these options, as they present the same range of costs for TSOs and users and the same potential for harmonization with the rest of Europe. Consultation has been carried out on transparency issues.

2.1.4 Central-South

Priorities in the Central-South (CS) region include: *(i)* harmonization and improvement of explicit auctions for cross-border capacity allocation; *(ii)* implementation of market coupling; *(iii)* transparency.

As for congestion management, TSOs of the CS region have published a common document with harmonized explicit auction rules for all borders (some specific rules still apply to each border). As a result, allocation rules and methods, definitions, products and nomination procedures are (almost) common since January 2008. Moreover, regulators, TSOs and PXs have agreed on working on the establishment of a Single Auction Office/Common Platform for the region as an efficient and practical way to achieve further harmonization. Regulators have requested TSOs to provide a proposal for a detailed action plan for this project.

The issue of market coupling is now being analyzed, together with TSOs and PXs, with the objective to identify a feasible approach, as well as borders where it can be tested.

Finally, regulators are also working on a Transparency Report.

2.1.5 South West

An Iberian market (MIBEL), including Spain and Portugal, was established in 2007.

The main priority in the region is increasing the cross-border capacity. Another issue is the improvement and harmonization of congestion management. Other priorities include implementing market coupling between MIBEL and the Central-West region as soon as possible, and developing reciprocal access to balancing markets.

Interconnection capacity between Spain and Portugal has doubled during the last years and it is planned to double again by 2014 (thus reaching 3000 MW in both directions). Indeed, the lack of sufficient interconnection capacity between France and Spain is deemed the major obstacle to market integration. An agreement was signed in 2008 by the French and Spanish governments for the construction of a 1200 MW interconnector.

Harmonization and improvement of cross-border congestion management started with the proposal of a common explicit auction mechanism for the allocation of capacity in the long and medium term. A new version of the rules for the allocation of capacity at the Spanish-French interconnection (IFE rules) was presented in September 2008 (see Section 3.3).

Finally, it emerged that the Spanish regulatory authority lacks the competences and powers on cross-border trade. This major regulatory issue has been solved: the regulator is now the competent body to propose the final version of cross-border trade rules for approval to the Spanish Ministry of industry, Tourism and Trade (Royal Decree 871/2007).

2.1.6 Central-East

Priorities in the region are two: to implement the relevant provisions of the Electricity Regulation (1228/2003) and the CM Guidelines and to devise solutions in support of the development of the regional market.

Explicit auctions for capacity allocations have been implemented on all borders (and are coordinated on five). Full coordination remains a key objective. To this end, TSOs established a regional Auction Office in Freising for the CEE Region (near Munich). Decisions on the future General Manager and IT tendering have been made. The formal establishment of the company qualified as a “notifiable merger” and it required the consent of the relevant competition authorities. The office will coordinate capacity allocation at all electricity interconnections in the CEE region on a daily, monthly and yearly basis, relying on the actual physical electric power flows, replacing thus different solutions in the region.

For the allocations a load flow based approach is foreseen: since 2006, TSOs have been working on a joint flow-based capacity calculation scheme that will include a common grid model. Recently, TSOs decided to follow the so called "maximum flow" approach. A model is under elaboration and test calculations are expected to be ready by 2008 to enable the TSOs and market participants to assess the potential market impacts.

Finally, transparency is also a relevant issue. An assessment of the current situation resulted in a plan to introduce stepwise transparency improvements during 2008.

2.1.7 Baltic

The main priority in the region is enhancing co-operation among TSOs (congestion is not an issue in the region).

BALTSO, the organization of the TSOs in the region, focused on technical co-operation on transit flows, inter-TSOs compensations and dispatch. A common balancing methodology exists between Latvia and Estonia. A study is under way for the creation of a common balancing market at regional level.

2.2 Coherence and convergence issues

The ERGEG conclusion paper on “coherence and convergence” indicates that ERIs could benefit from a stronger co-ordination (ERGEG, 2008a). The existence of binding Guidelines is a strong support for the development of common approaches on various issues; vice-versa, less work is being done in areas were legal requirements do not exist. A lack of inter-regional co-ordination has been signaled as the main hindrance for progress (in particular in the area of market coupling). Following closely the ERGEG conclusion paper (2008a), we provide here further details on these issues.

2.2.1 Capacity calculation

The CM Guidelines legally require TSOs within each of the seven electricity REMs to coordinate and implement common capacity calculation methods.

Common capacity calculation for the day-ahead time frame, based on a common transmission model, is a priority in both the Central-West and the Central East regions (Nordic countries have used a common transmission model since the 1990s). In other words, convergence towards a common calculation model (eventually flow-based) seems to appear only in "central regions" (because of the meshed structure of the networks). As for the other regions, the question has not been addressed yet.

On the one hand, given the requirement of the CM Guidelines, and ERGEG concerns regarding a lack of transparency in current calculation methods, changes might be expected in the way national TSOs perform the task of capacity calculation. On the other hand, no institutional changes are foreseen at this level. The literature has proposed the creation of Regional/European TSOs, with the aim of enhancing the efficiency of capacity calculation and investment decisions (Glachant et al., 2006). The latter topic has not been addressed in the Regional Initiatives.[95]

As for the Central-South region, the methodology for the assessment of the Net Transfer Capacity (NTC) of the Italian northern interconnections is harmonized between Terna and the neighboring TSOs - ELES (Slovenia), RTE (France), Swissgrid (Switzerland) and Verbund-APG (Austria) - according to ETSO guidelines and UCTE Operational Handbook. The calculation makes use of a common transmission model and is repeated once a year.[96]

The CS region is working in two directions: (i) assessing possible benefits deriving from periodic NTC calculations during the year, instead of the present "yearly-fixed" methodology; (ii) further harmonization and transparency of NTC calculation on the borders

[95] Even so, European TSOs are working toward higher coordination in network operation and investment.
[96] TSOs take part in a Working Table that carries out all the activities needed to check NTC consistency and reliability as well as to ensure transparency towards supervisory bodies.

within the CS region that are not included in the common methodology, in particular the Switzerland-Germany, Switzerland-France, Switzerland-Austria borders and the border between Italy and Greece (the latter is in principle very simple because the two countries are interconnected by a DC cable).

2.2.2 Long and medium term capacity allocation

The CM Guidelines legally require TSOs to coordinate and implement common capacity allocation methods for each region.

As for long and medium term allocation a strong convergence arises: explicit auctions are likely to be the standard cross-border allocation mechanism for the near future in Europe. The only exception is the Nordic region, where long term hedging instruments are available for this purpose (all transmission capacity is allocated in the day-ahead market). Several projects were started on harmonization of auction rules and some of them foresee the creation of a single auction platform. They regard the Central-West, France-UK-Ireland, Central-South, South-West, and Central-East regions. These projects, and in particular the creation of regional auction platforms, might also help to harmonize the characteristics of the traded products (yearly/quarterly, firmness, notification schemes, and so on).

Explicit auctions are normally carried out by national TSOs. As for the Central-South region, a first harmonization effort has already been completed. Operators have indicated that further harmonization in auction procedures and products would provide significant benefits in lowering transaction costs and increasing market liquidity (in particular, the degree of firmness of capacity is deemed extremely important to reduce trading risk).

The creation of a single regional Auction Office (and trading platform) has been proposed for the Central-South region. Regulators have requested TSOs to provide a proposal for a detailed action plan for this project, where they will clarify all the main administrative measures to be adopted and gather all the necessary information in order to decide the best legal vehicle to set up the single office. Both governance and ownership are relevant issues in this matter.

It is important to note that ERGEG long-term vision on long and medium term capacity allocation is quite different from the status quo. ERGEG envisions a convergence towards implementing market coupling with financial transmission rights issued by the TSOs for hedging risk.

2.2.3 Day-ahead capacity allocation

As for the allocation of day-ahead cross-border capacities, market coupling (implicit allocation) has emerged as the preferred option in three regions (plus the Nordic market).

On the one hand, ERGEG notes that allocation methods are converging: the Central-West region has already implemented a three-country market coupling and the plan is to extend it to Germany and Luxembourg; the South-West region has implemented a two-country market splitting mechanism and plans to couple the MIBEL with the Central-West market; a day-ahead market coupling has been introduced on the German-Danish interconnectors; the Central-South region has indicated market coupling among its first priorities. On the other hand, convergence is seriously challenged by the extent to which countries in two or more regions, or neighboring regions, might participate in coupling projects.

Day-ahead market coupling, differently from market splitting, does not require national PXs to implement major modifications, in particular in the case of a NTC-based volume coupling (indicated by ERGEG as a flexible, transitional solution on the way to harmonized price coupling). Complete harmonization in products traded in each market and trading platforms might not be required. Nevertheless, harmonization is necessary on a number of issues (such as gate closure times and transmission right notification), in addition to the development of a suitable coupling algorithm. These issues, with respect to the Central-South region, are discussed in Section 4. Note that ERGEG long-term agenda foresees a gradual merger of PXs for physical spot trading and implementation of market splitting (ERGEG, 2008a).[97]

[97] It is remarkable in this sense, the recent (June 2008) merger between Powernext and EEX on both spot and derivatives markets, concluded within the framework of the Pentalateral Forum. French and German markets represent together almost 1/3 of the European electricity production. Moreover, because EEX manages both the Swiss and Austrian spot markets, this merger sets the path for a strong convergence in the direction indicated by ERGEG.

2.2.4 Intra-day capacity allocation

Regarding intra-day capacity allocation, a continuous trading platform is functioning in the Nordic region and a similar arrangement is being discussed in two regions: Central-West and South-West. Intra-day market coupling is implemented also in MIBEL (it includes both auctions and continuous trading). Cross-border coordinated intra-day allocations are organized also on the French-Belgian, French-German, French-Swiss, French-Spanish, and German-Swiss borders.

Indeed, integration of intra-day markets does not emerge as a priority in many regions. Where this issue has been addressed, regions have opted for continuous trading.

ERGEG (2008a) concludes that with regard to the design of intra-day markets (discrete or continuous trading), the question remains open and that both options may coexist. In terms of organizational issues, ERGEG position is that PXs should be in charge of intra-day trading. As for the Central-South region, at present the issue is not a priority.

2.2.5 Balancing

The integration of balancing markets might be highly beneficial, according to the EC's Sector Inquiry (2007), given the small size of balancing zones which favors incumbents and potentially increases costs. The topic is also being addressed at the EU level via the ERGEG voluntary Guidelines of Good Practice for Balancing Market Integration (to be finalized at the beginning of 2009). There are, as of today, no legally binding rules on the matter.

Balancing market integration has not been targeted as a first priority issue in many regions. The Nordic market applies a coordinated balancing regime, the France-UK-Ireland is planning to introduce one and both regions have opted for a TSO-to-TSO based model. In addition, only the Baltic region has indicated balancing market integration as a first priority issue (it is, however, a priority in the South-West region).

ERGEG is hoping that the France-UK-Ireland proposal will foster further initiatives in integrating balancing markets. Nonetheless, several aspects remain open. In particular, in developing the Guidelines ERGEG is looking at the following issues: harmonization of remuneration schemes for balancing orders, share of automatically activated reserve vs.

manually activated reserve, TSO-TSO vs. Actor-TSO models and the interaction between balancing and intra-day markets. As for the Central-South region, this issue is not a priority.

2.2.6 Transparency

The lack of transparency is another issue identified in the EC's Sector Inquiry (2007) as a major barrier to integration. Efforts to improve transparency were targeted in most of the regions as a first priority issue. In particular, the discussion focused on the implementation of the mandatory transparency requirements of the CM Guidelines and of the recommendations in ERGEG's Guidelines on the same topic. Numerous respondents to ERGEG consultation on coherence and convergence indicated transparency as the basis for creating a common level playing field where all relevant data are available for all market participants in due time.

The Northern region report for the harmonization of transparency rules is based on both the CM mandatory and the ERGEG voluntary guidelines. The report clearly specifies publication requirements for all relevant data: what has to be published, when it should be published, which time frame it shall refer to, how long information shall be available and if it is necessary to regularly update the data. It also specifies deadlines for implementation.

The report has provided the basis for similar transparency reports in the Central-West and Central-East regions. In both cases the Northern report was only slightly modified. It also provided the basis for discussion on transparency in the Central-South and South-West regions. A strong convergence is thus emerging in terms of transparency, with a region setting the path for the others. Moreover, ERGEG is strongly encouraging all regions to harmonize definitions and publication details.

Regulators in the Central-South region are preparing a Transparency Report. The document now under consultation is based on the Central-West region Transparency Report. The adoption of the proposed information disclosure requirements would imply changes, mainly on the part of TSOs and market operators. As for PXs, wholesale market data can be published voluntarily and, in any case, this information is already available to a large extent.

3. Experiences in Europe

As of today, implicit allocation of cross-border capacity is implemented in three regions in Europe: the Nordic region, the Central Western European region (integration regards, at present, three countries: France, Belgium and The Netherlands) and the South West region (the Iberian market includes today Spain and Portugal). In addition, other integration models have been proposed over time: the German model of an Open Market Coupling (2006), the joint ETSO-EuroPEX proposal for a Flow-based Market Coupling (2004), and more recently the Central Western Europe Market Coupling Project (2008).

In what follows, we describe and discuss the main features (Section 3.1 and Section 3.2) of the first two existing regional markets, focusing mainly on the day-ahead segment of the market. Explicit auctions are, at present, the most common arrangements for cross-border capacity allocation in Europe. In Section 3.3 we describe two experiences, in particular the France/Spain model and the France/England model.

3.1 Nordic region

The Nordic Power Market (Nord Pool) is the oldest regional market in Europe. Nord Pool started as the Norwegian power exchange in 1995; its day-ahead auction trading system was extended to Sweden in 1996, Finland in 1998, Denmark in 2000 and partly also to Germany in 2005. The day-ahead auction uses the internal border capacities to optimize the clearing of orders introduced in the different countries (every country corresponds to one or more price zones). Harmonization of the intra-day market and real-time markets are also quite advanced (Meeus and Belmans, 2008).

The total volume traded at NPS (Nord Pool Spot) in 2007 was over 70 per cent of the total Nordic electricity consumption, close to 290 TWh (an increase of 16 per cent from 2006).

In what follows, we briefly describe and discuss: the mechanism for congestion management (Section 3.1.1.); the products traded in the Nordic market (Section 3.1.2); and the economic effects of integration (Section 3.1.3).

3.1.1 Congestion management

Nord Pool Spot organizes the physical trade of electricity in the Nordic countries and Kontek with a day-ahead market (Elspot). In addition, NPS operates an intra-day market (Elbas) in Finland, Sweden, Denmark and Germany (Norway is expected to be included in January 2009).

In order to handle grid congestions, the Nordic exchange area is geographically divided into bidding areas. All trading capacity between the bidding areas is dedicated to NPS for implicit auction in the Elspot price calculation. In general the bidding areas coincide with the transmission network of each TSO, thus Sweden, Finland and the German area Kontek are each one bidding area. The grids in Jutland (West Denmark) and Zealand (East Denmark) are not physically connected, so they are two bidding areas in Denmark. The Norwegian grid is usually divided into two bidding areas but having three or more areas within Norway is possible.[98] As mentioned above, the only exception from this principle occurred on the West Denmark-Kontek interconnection, where the CBO method was applied until September 2008.[99]

Elspot calculates an unconstrained market clearing price, the System Price, which is obtained ignoring the trading capacities between the bidding areas. If the contractual flow of power between bidding areas at System Price exceeds the trading capacity, Area Prices are calculated.

The Elbas market provides continuous power trading (the trade is concluded when bid and ask match) 24 hours a day, 7 days a week covering individual hours, up to one hour prior to delivery (Elbas opens for trading when the deadline for filing complaints on Elspot is closed – at 14.00 CET). The time span between the closure of the day-ahead Elspot market and the actual delivery hour of concluded contracts is between 12 and 36 hours forward in time.

Also in the Elbas market price areas are separated when bottlenecks occur.

[98] It is the Norwegian TSO – Statnett – who determines how to split the Norwegian grid into bidding areas based on physical conditions.

[99] An assessment of cross-border auctions on this interconnection can be found in Kristiansen (2007a, 2007b). The author finds that daily, monthly and annual explicit capacity auctions do not reflect the value of the underlying asset as specified by the valuation of price differentials between the two countries. Therefore, he welcomes the introduction of new coupling arrangements on this border.

Grid congestions within a bidding area are handled by the TSOs using counter-trading. In other words, the TSO buys increments/decrements from producers/consumers in order to solve internal congestions. In performing this task a TSO has to take into account how the different agents affect the constraint in question, in order to minimize the cost of redispatching. This means that a TSO must be allowed to select bids submitted to “regulation markets” according to both the merit order and the effectiveness in solving the congestion - solving an optimal dispatching problem with a complete network model (Bjorndal and Jornsten, 2007).

In any case, counter-trading affects the costs sustained by the network system operator and, depending on how they are charged to market participants, they will also affect the efficiency of the market. Bjorndal and Jornsten (2007) have shown that TSOs in the Nordic market might prefer to “remove” internal capacity constraints by setting more restrictive transmission capacity limits between the bidding areas. This “indirect” method of congestion management is inefficient: it may be more costly and result in larger price differences than necessary.

Bjorndal and Jornsten (2007) suggest that a higher degree of coordination among TSOs in the Nordic market might be beneficial and could be carried out, for instance, through Nordel (the organization of the Nordic network operators). Indeed, considering the Nordic power market as a whole may lead to a better utilization of the existing grid. For instance, potential benefits might derive, in terms of economic efficiency, from considering the network physical bottlenecks as the basis for zonal separation, instead of following, as of today, the borders of the TSO control areas. Also Glachant e Pignon (2005), referring to the Nordic market, argue that TSOs have an incentive to over-constrain inter-zonal transfers in order to avoid intra-zonal congestions and thus hide the need for additional investments.

3.1.2 Traded products

There are two main different types of bids available on Elspot: the hourly bid and the block bid. Block bids are not a product traded on the Italian Power Exchange. On the contrary, they are often found on other European PXs. It is therefore interesting to synthetically describe the Elspot block bids to understand why they are useful for market participants. For a discussion on which difficulties they introduce in the auction problem see Part II of this report.[100]

The block bid is an aggregate bid for several hours, with a fixed price and volume throughout these hours. The block bid gives the participant the opportunity to set an "all or nothing" condition for all the hours within the block. This is particularly useful when the cost of starting and stopping power production is high. Furthermore, inelastic production, consumption and contracts can be handled efficiently with block bids.

While hourly bids can be partially accepted, a block bid must be accepted in its entirety. The block bid price is compared with the average Elspot price for the hours to which the block bid is applied and the bid is accepted if the following conditions are met:

- If the bid price of a sales block is lower than the average Elspot area price;
- If the bid price of a purchase block is higher than the average Elspot area price.

On Elspot the participants can freely pick the start and stop hour of a block, but this must consist of at least four consecutive hours. Each participant may post up to fifty block bids per delivery day.

On Elspot it is also possible to define links between block bids, meaning that the evaluation and acceptance of one block bid (daughter block) is dependent on the acceptation of another block bid (mother block). It is possible to link up to three block bids together. The third bid is then dependent on acceptance of both the first and the second bid.

The linking of block bids is useful for instance when the cost of starting one generator depends on whether another generator is already started or not. Or if the start up of a generator at night is favorable only if the same generator is planned to run at daytime as well.

[100] On the Iberian market, bids can include "complex conditions", such as maximum amount of energy accepted in total (useful for hydro/pumping plants with a predefined amount of energy that they want to sell or buy) or minimum number of adjacent hours accepted. These conditions introduce complexities in the auction algorithm in a way similar to block orders.

3.1.3 Effects of integration in the Nordic market

Under the assumptions that different production technologies are employed in the involved markets and an efficient pricing scheme is found on both sides, the potential benefits of market integration are: (i) a better use of generation and transmission resources, increased reliability of supply, and optimal investments in infrastructure; (ii) as a result of these effects, a decrease in prices on the wholesale market (Kube and Wadhwa, 2007).

Kube and Wadhwa (2007) empirically observe the effect of integration on the development of wholesale prices on the Nordic spot market, Elspot, between 1996 and 2004. By means of a regression analysis (that includes the main supply and demand determinants of wholesale prices and a proxy for the degree of market integration) they find that market integration leads to a decrease in prices due to efficiency gains. The most probable drivers behind these gains are the diversification of the generation technology portfolio and the reduction of transaction costs. This work offers strong support for the current drive towards a single European electricity market: there are efficiency gains to be achieved via an integration of national markets.

Lundgren et al. (2008) empirically explore the electricity price dynamics in the Nordic power market during the years 1996-2006. They conclude that a larger electricity market seems to reduce the probability of sudden price jumps. That is, the multi-national electricity market integration seems to have created a market that handles external shocks to supply and demand more efficiently than the separate national electricity markets previously did.

The association of Nordic energy regulators observes that each of the national wholesale markets has a certain degree of ownership concentration when considered separately and large national producers are often dominant in their national markets (the concentration in all bidding areas is to be regarded as high, with Norway as the least concentrated market and Denmark as the most concentrated). From this perspective, the existence of a Nordic market implies a market size effect, where market shares for national producers are decreased. This effect is maximized when there is a uniform Nordic spot price. This occurred approximately one third of the time in 2005, 2006 and 2007 (NordREG 2007, 2008).

3.2 Central Western European region

Market integration in the Central Western European (CWE) market began in 2006 with the TriLateral Coupling (TLC) project that includes France, Belgium and The Netherlands. When TLC started, both France and The Netherlands had already well-established PXs: Powernext in France, functioning since 2001 and APX in The Netherlands, since 1999. Belgium, on the contrary, was only in the process of starting an exchange. Belpex, launched in 2006, was created by a cooperation agreement which included Elia (the Belgian TSO), RTE (the French TSO), Tennet (the Dutch TSO), APX and Powernext (and they all own a share of Belpex).

Unlike the Nordic market, the three PXs have chosen to coordinate their existing trading platforms and not to rely on a single system (each area had its own Market Operator and different market designs). In other words, while the Norwegian power exchange extended its model to other countries over the years, TLC was started as an agreement between the three PXs, in order to coordinate operations. Note also that the initiative for market integration came directly from the involved players (regulators ensured that there was space for the project to develop).

The three PXs, that offer several different services at the national level, have chosen to couple their day-ahead operations. The coupling mechanism involves handling simultaneously their supply and demand curves, but there is no common order book. For the members of the individual PXs, bidding methodologies have remained practically unchanged and the PXs continue to exist as legally separate entities, with their own clearing and settlement arrangements (Meeus and Belmans, 2008).

In practice, some degree of centralization is necessary. Currently, the Dutch power exchange, APX, carries out the required centralized functions (Coordination module - see below) and collects the aggregate supply and demand curves from the other PXs (in the form of Net Export Curves - also see below).

The internal border capacities are used to optimize the clearing of the orders introduced to the day-ahead auctions. However, in contrast with the Nordic market, a fraction of the available cross-border capacity is still auctioned in explicit auctions (about 20%).

The three markets have jointly a 111 GW peak load and an annual consumption of 684 TWh. Taken together the three countries account for about 1/4 of the EU electricity production (Meeus and Belmans, 2008).

In what follows, we briefly describe and discuss the coupling mechanisms (Section 3.2.1); moreover, we provide some evidence of the effects of integration on prices and cross-border capacity utilization (Section 3.2.2).

3.2.1 Trilateral market coupling algorithm

The market coupling algorithm adopted by APX, Belpex and Powernext (2006) makes use of:

- the Available Transfer Capacities (ATC) between each area for each flow direction for each Settlement Period (each hour) of the following day;[101]
- the Net Export Curves of each market and hour;
- the Block Orders submitted by participants in each market;

and it provides, as an output for each market and hour:

- the Net Position;
- the price;
- the set of accepted Block Orders.

A fundamental element of the coupling algorithm is the *Net Export Curve (NEC)*. A NEC reflects a market's import or export volume sensitivity to price and it is obtained by determining the volume differences between the supply curve and the demand curve for each hour of the day. Depending on the market, two types of NEC are possible: either a stepwise curve, as in APX and Belpex, or a (piecewise) linear curve, as in Powernext.[102] This does not create particular problems to the coupling algorithm.

[101] The transmission model in TLC is tree-like, with the three countries lined up; allocation of cross-border capacity is based on ATC – it is not a flow-based model.

[102] If hourly bids and offers are defined with price-quantity range pairs - i.e. for each price a range of feasible quantities is defined - the curve is obtained by joining the segments to each other and the NEC is a stepwise curve. If hourly bids and offers are defined with price-quantity couples, the curve is obtained by joining the points to each other, and the NEC is a piecewise linear curve.

Coupling two markets is easy using the NECs. Since the market with the lower isolated marginal price exports to the market with the higher isolated price, the export of one market is equal to the import of the other and the equilibrium is found at the intersection of the NEC of a market, and the inverted NEC of the other. As illustrated in Part I of this report, if the ATC is large enough, the equilibrium and net positions are given by the intersection of the two curves. If congestion occurs, the power transmission is equal to the ATC; the net positions are equal to the ATC and the price of each market is given by the point on its NEC corresponding to the ATC.

A market may construct several NECs for every Settlement Period (hour), considering every possible combination of accepted and rejected Block Orders. These combinations of accepted Block Orders are called Winning Subsets.

A particular instance of the NEC is the Block-Free NEC, constructed strictly from Divisible Hourly Orders, thus excluding all Block Orders. All other possible NECs of the same Settlement Period can be derived from the Block-Free NEC, as they only differ from each other by the Winning Subset. For the NEC construction, the Block Orders within the Winning Subsets are represented by price-inelastic Divisible Hourly Orders. The net volume resulting from the Winning Subset for a particular Settlement Period is called the Net Block Volume (NBV). Thus each possible NEC for a particular Settlement Period is a horizontal translation of the Block-Free NEC by a NBV (APX, Belpex and Powernext, 2006; Dupuy, 2008).

The Trilateral market coupling process adopts a decentralized approach. In each power exchange, Block Orders are submitted to a so called Block Selector and Divisible Hourly Orders to a NEC creator. A NEC is built on the basis of a hypothetical set of accepted Block Orders (the algorithm is initialized with a first Winning Subset that includes all Block Orders). Then, NECs from each market are used to determine the price and the net position of each market (in the so called Coordination Module). The set of prices obtained might not be compatible with the assumed Winning Subset, thus the Winning Subset is updated and the prices calculated again. These calculations are iterated until the algorithm reaches a stable solution (APX, Belpex and Powernext, 2006).

Consequently, because of the Block Orders, the algorithm involves iterations between two modules:

- The Block Selector of each power exchange, in charge of the decentralized selection of block offers (returns the NBV);
- The Coordination module, in charge of the centralized calculation of the prices and net positions, using NECs and ATCs.

As for the *Coordination Module*, the algorithm is based on the idea of performing two bilateral couplings in a sequence.

The first step is to order the three markets according to their isolated marginal price. Let us assume that market A has the lowest isolated price and market C the highest. Thus, market A is going to export to market C (potentially via market B, if market B is Belgium), which will raise the price in market A and reduce the price in market C. Two situations can then occur:

- if a market uses all its ATC, it becomes isolated and cannot export or import anymore; then a bilateral coupling takes place between the two other markets;
- the price of market A or C reaches the price of market B, and the algorithm aggregates the two markets together (the two NECs are added together); then a bilateral market coupling takes place between the market resulting from the first coupling, and the other one (Dupuy, 2008).

Examples of this procedure can be found in APX, Belpex and Powernext (2006) and Dupuy (2008). From these examples it emerges that the problem solved by the algorithm in case of ATC limitations is greatly simplified by the fact that the three markets are lined up (in case of congestion, one market gets isolated and a bilateral coupling takes place between the other two). For instance, in a scenario, where the ATC between Belgium (exporting market) and the Netherlands (importing market) is limiting cross-border transactions, the Netherlands becomes isolated. This is due to the absence of interconnections between the Netherlands and France; otherwise the algorithm would have to consider the possibility of import from France as well.

Dupuy (2008) develops a sequential algorithm that uses NECs as inputs for the solution of a coupling problem (i.e. the full supply and demand curves from each market are not required) between three countries all interconnected to each other. Indeed, a fundamental aspect of the TLC algorithm is that the input data are NECs: this enables power exchanges to limit the information they need to share in the Coordination Module. However, as indicated also by the simplified market model solved by Dupuy, this interesting feature introduces not negligible additional difficulties when the number of countries and the complexity of the network increase. Dupuy infers that when dealing with more than three countries, modeling the coupled auctions as an optimization problem is a preferable solution. The optimization problem, however, requires power exchanges to share the full supply and demand curve in a centralized auction office.[103]

Another important element of the coupling algorithm is the need for a clear criterion to solve indeterminacies in price and quantities. In some cases the TLC algorithm produces outcomes where a range of prices or a range of quantities are possible, while one single combination of price and quantity is required as outcome for each individual period.

As for quantity determination in the TLC algorithm, when several sets of Net Positions are possible for the same equilibrium prices:

- the set of Net Positions are chosen so as to maximize total executed volume;
- in case several volume-maximizing sets of Net Positions are possible, the volume is spread evenly between the markets (APX, Belpex and Powernext, 2006).

Similarly, a range of feasible prices may be possible for linear as well as for stepwise NECs. When such indeterminacy arises, the middle of the feasible range is chosen. If several vertical segments have common prices, the middle of the common range is chosen (APX, Belpex and Powernext, 2006).

The importance of price coordination between power exchanges is well illustrated in Meeus et al. (2008). First of all, the authors observe that which price is chosen on a price range is a relevant issue only if the coupling algorithm is often facing this problem and if these price ranges are significant. In the first two months of Belpex operation, price ranges were not observed only in 30% of the hours. Although in 80% of the hours the range was

[103] Dupuy (2008) observes that with two markets only, it would be possible to solve an optimization problem using NECs only. In this case it is possible to compute, from NECs, the change in welfare for a market when moving from an isolated to a coupled situation. Thus, the objective of the optimisation would be to maximise the increase in welfare. Obviously, a difficulty is encountered when working with more than two markets. It would then be necessary to examine all possible aggregations of different markets before finding an optimal two-area model (an iteration process would be needed to test all the possible situations).

smaller than 20 €/MWh, there are also few observations with ranges close to 400 €/MWh. Therefore, given a typical wholesale price of 50 €/MWh, the problem appears quite relevant.

Secondly and more importantly, since cross-border price differences generate congestion revenues and provide locational signals to market operators, it is crucial to ensure that prices are set correctly. Meeus at al. (2008) note that the mechanism employed in TLC to set the price in case of indeterminacy (i.e. taking the middle price of an overlap between price ranges, subject to the property that power always flows from a high price to a low price region) is specific for three aligned markets. If market coupling is to be extended to other countries (and meshed networks) a different solution is necessary. In their paper, the authors argue that a good and straightforward way to choose between alternative prices is to minimize congestion revenues. In choosing between different sets of Linear Marginal Prices, LMPs, the authors observe that the set that minimizes congestion rents captures the value of the positive effects on total gains from trading of a 1 MW increase in transmission capacity (while the set that maximizes congestion rents captures the negative effect on total gains from trading of a 1 MW reduction in transmission capacity). In principle, the highest and the lowest value are as relevant, but in a European context with scare interconnection capacity between countries, the question is rather which interconnector to further expand than which to maintain. This is one argument in favor of minimizing the congestion rents when choosing between sets of LMPs. Another argument is that one of the main concerns at the moment in Europe is that only a small fraction of the congestion rents is used to invest in the network (Meeus at al., 2008).

As for the *selection of Block Orders* each power exchange selects blocks independently, but with the same logic. In fact, for every iteration, each Block Selector finds the Winning Subset using the prices returned by the Coordination Module in the following way:

- if the average market price is higher than or equal to a Sale Block Order’s Price Limit, then the block is accepted;
- if the average market price is lower than or equal to a Purchase Block Order’s Price Limit, then the block is accepted;
- all other blocks are rejected for the current iteration.

In addition, in order to prevent the algorithm from cycling and to ensure convergence in a reasonable time, a number of rules are implemented to allow paradoxical rejection of some Block Orders (Paradoxically Rejected Blocks, PBR). The main assumption of such rules is that Block Orders rejected at the previous iteration and accepted at the current one are considered as potential PRBs. In each iteration, it is made increasingly difficult for such blocks to re-enter the Winning Subset so that cycles are less likely to occur as the number of iterations increases (APX, Belpex and Powernext, 2006).

The choice of a decentralized versus a centralized approach to block selection is discussed in Meeus (2006). Ideally, block selection in a relatively small-scale scenario with a small number of blocks, could be performed by enumerating all potential solutions and selecting the optimal one (within a reasonable time).

In a scheme where block order information is decentralized, the enumeration of all possible solution is clearly not possible. In addition, since different selectors are used it is not even possible to compare different solutions. Although in principle all feasible solutions found with a decentralized block selector could be stored, it would not be possible to choose between several feasible solutions. In fact, blocks are included in NECs as price taking orders so that gains from trade or the number of PRBs are not known centrally (only national PX knows the order's price limit).

As a consequence, a decentralized selection is always inferior to a centralized one even if enough time was available to try all block combinations - and time is obviously a constraint in realistic scenarios (Meeus, 2006).

Meeus (2006) carries out a test using two very simple algorithms, a centralized and a decentralized one, to have some indications on their relative performance. Results show that in terms of gains from trade and number of PRBs decentralizing the selection of block orders has a negative impact on performance. Moreover, under-performance is larger as the market develops and more regions are coupled.

3.2.2 Effects of integration in TriLateral Market Coupling

TLC is now just two years old (it started operation in November 2006). Several different parameters have been used to measure the "success" of this market coupling arrangement.[104]

First, market coupling has lead to a more efficient use of the daily capacity on the interconnections. Figure 3.1, illustrates the nominated cross-border net exchange between France and the Netherlands in 2005. Inefficient flow directions (from high to low price areas) were registered in 42% of the hours (II and IV quadrant) (see also Section 5). By contrast, Figure 3.2 shows that, since the creation of the TLC, cross-border capacity has always been used in the right direction (from low to high price areas).

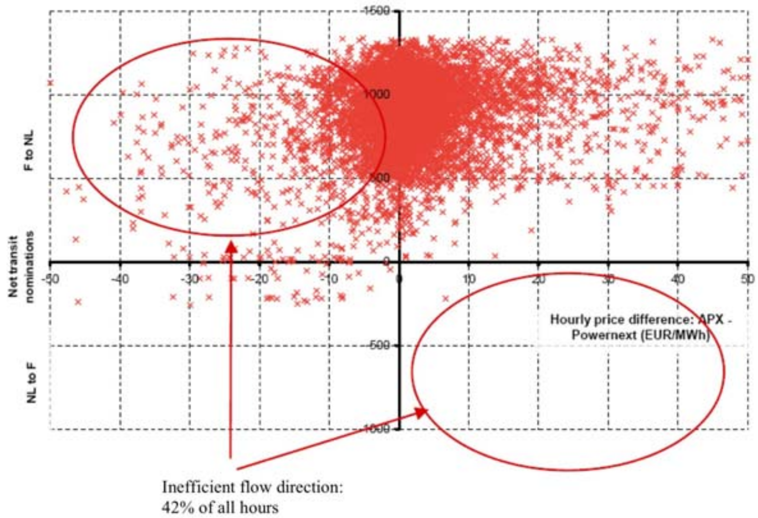

Figure 3.1 - Nominated cross-border net exchange between France and The Netherlands in 2005 (Meeus, 2008)

[104] Data reported here are taken from a lecture given by L. Meeus in Florence in April 2008, at the Florence School of Regulation - Advanced Training Course on Electricity Markets.

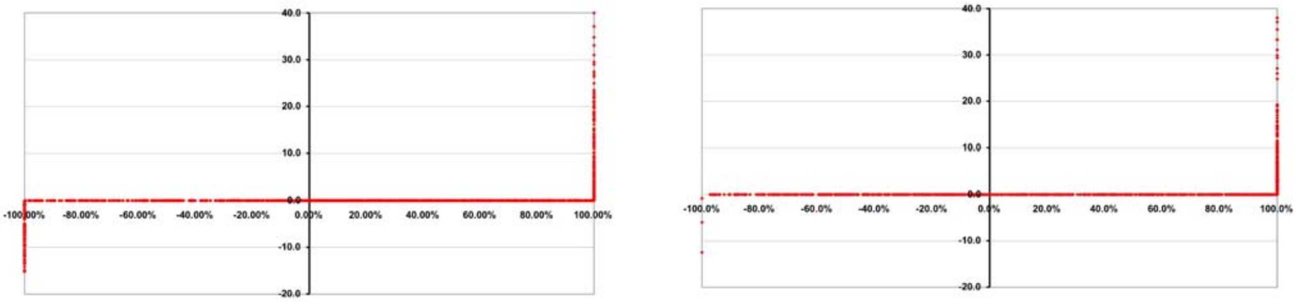

Figure 3.2 - Use of France-Belgium (left) and Belgium-The Netherlands (right) capacity (% of average capacity) vs. price difference (€) during TLC (Meeus, 2008)

Second, market coupling has lead to a uniform price for the whole area for a considerable number of hours. In the period November 2006 - November 2007, a uniform TLC price was registered in 60% of the hours. The Netherlands were isolated in 26% of the hours and France in 12%. Three different prices were registered in only 2% of the hours in the same period. Figure 3.3 reports monthly congestion revenues for the border Belgium - The Netherlands. It emerges that for the year 2006, congestion revenues have been quite large. However, they have significantly decreased since the creation of TLC.

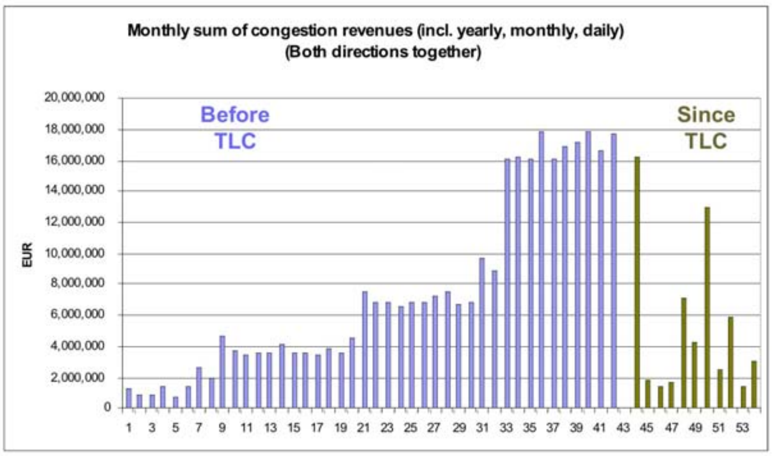

Figure 3.3 - Monthly sum of congestion revenues (including yearly, monthly, daily): Netherlands and Belgium, both directions (Meeus, 2008)

Moreover, as illustrated in Figure 3.4, the occurrence of large price differentials between Powernext and APX has significantly decreased.

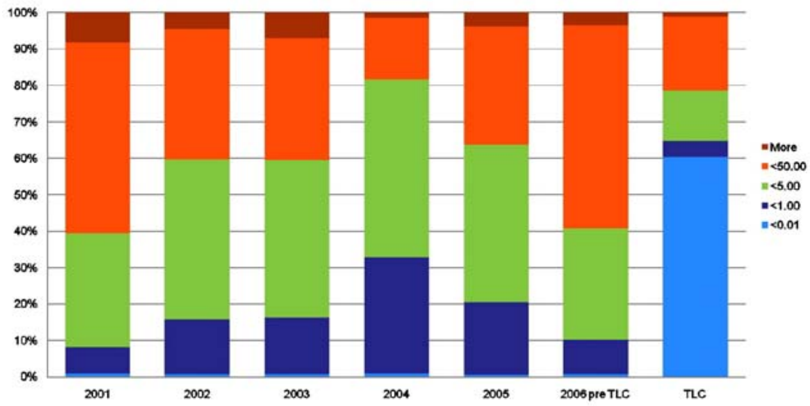

Figure 3.4 - Price differences: Powernext-APX; hourly prices Nov. 2006 - Nov. 2007 (Meeus, 2008)

Third, probably one of the most important lessons learned from the creation of TLC is that liquidity is not a prerequisite for market coupling. Figure 3.5 illustrates the liquidity growth on Belpex from November 2006 to March 2008.

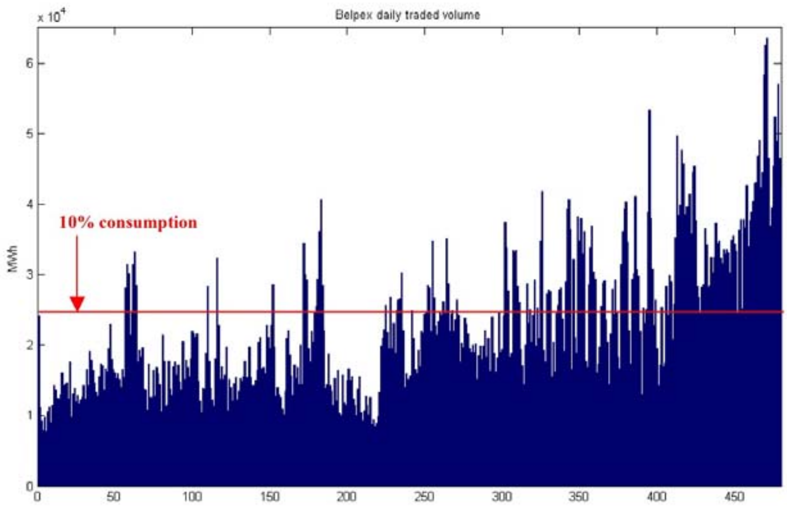

Figure 3.5 - Belpex daily traded volumes: Nov. 2006 - March 2008 (Meeus, 2008)

In a very recent contribution, Jonghe et al. (2008) confirm that TLC has resulted in sharp increase in the frequency of hourly price convergence between APX and Powernext after the coupling of the markets (Figure 3.4). Their work explores then the impact of market coupling on price volatility. They find evidence of a significant price volatility reduction on APX; however, not on Powernext (note that only one year of data were available at the time of the study). This reduction is explained using the concept of "volatility reduction potential", based on the non-simultaneous occurrence of supply and demand shocks in different exchanges. Indeed, data analysis indicated that this reduction potential was higher on APX than on Powernext.

3.3 Examples of explicit auctions

This section discusses the main features of the explicit auction for cross-border capacity on the France-Spain interconnection (Section 3.3.1) and on the France-England interconnection (Section 3.3.2).

3.3.1 France-Spain interconnection (IFE Rules)

A coordinated mechanism of capacity allocation for the France-Spain interconnection was launched in June 2006. The current version of the "Capacity Allocation Rules" (IFE Rules) was published in June 2007 (RTE and RED, 2007). This is the document we refer to in describing the allocation mechanism. However, following a consultation phase, these rules are now under revision, according to ERGEG's recommendations (ERGEG, 2008b). The proposed modifications are summarized below.

For the France-Spain interconnection separate auctions are implemented for the two directions. Products are offered on different time horizons in the form of Physical Transmission Rights (PTRs) on the basis of ATC:

- Continuous annual PTRs and non continuous annual PTRs (a number of days are excluded according to specifications);

- Continuous monthly PTRs and non continuous monthly PTRs (a number of days are excluded according to specifications);
- Daily PTRs and intra-day PTRs (there are two intra-day auctions; one is held the day-ahead in the afternoon, and one on the day of delivery, in the morning).

TSOs guarantee the firmness of PTRs; however, for reasons of power system safety PTRs may be reduced before the Programming Authorizations are sent (at 4.00 pm of D-2), on a pro-rata basis. Prior to the sending of the Programming Authorization, PTRs can be reduced if necessary. In this case participants are compensated with 110% of the price of the capacity right. After nomination (7.45 am the day-ahead for annual and monthly PTRs; 3.00 pm the day-ahead for daily auctions; several deadlines apply for intra-day PTRs) the resulting Exchange programs are firm except in case of *Force Majeure*. In this case, curtailments are pro-rata and do not depend on the horizon in which the PTR has been obtained.

PTRs allocated at annual and monthly auctions that are unused are re-allocated in the daily auction and similarly unused daily PTRs are re-allocated in intra-day auctions. A secondary PTR market is available for reselling of yearly PTRs at monthly and daily auctions and monthly PTRs at daily auctions.

Nomination and designation of the counterparty are obligatory irrespective of the energy trades in the bilateral and/or organized markets. A PTR which is not nominated is lost, without financial compensation ("use-it-or-lose-it" provision). Netting of nominations is applied and considered in the calculation of the ATC for the subsequent explicit auctions.

These auction rules are considered temporary: a market coupling mechanism is expected to replace the explicit daily auctions (RTE and RED, 2007). On a shorter time frame, three important changes will be introduced in the IFE rules and they regard the firmness of the capacity rights (ERGEG, 2008b): *(i)* capacity acquired in the day-ahead and intra-day auctions cannot be curtailed, except for *Force Majeure*; *(ii)* long-term (annual and monthly) capacity is "financially" firm: compensation in case of curtailment or cancellation of the day-ahead auction is at full market spread (with a cap, to limit the effect of this provision on tariffs); *(iii)* not-nominated capacity are automatically offered in the day-ahead auction: agents who do not nominate the daily PTRs receive back the price of the day-ahead auction (ERGEG, 2008b).

3.3.2 France-England interconnection (IFA Rules)

The French and the British transmission system are connected by a 2000 MW DC interconnector (IFA), operated by National Grid and RTE. At the end of 2000, capacity rights on the England-France Interconnector were made available to third parties for the first time. As of today, capacity is offered for sale separately in the direction England to France and France to England through tenders and auctions, for different contractual periods, according to the IFA Access Rules (RTE and NGIL, 2005a).

A first (and only) tender was held in January 2001, for the capacity in the direction from France to England, in units of 50 MW with a term of 3 years.

Auctions are held periodically for capacity rights with different time horizons (from “contract year” to “contract day”) and each direction is auctioned separately. The following “categories of capacity” are currently available (in both directions and in 1 MW tranches): annual, seasonal (6 months), quarterly (3 months), monthly, weekends (Saturday and Sunday), and daily (one “contract day”).

A user who has acquired capacity is entitled to submit nominations to the TSOs for transfers of energy. Nominations for capacity rights that are longer than one “contract day” must be made by 6 am on the day ahead. Not nominated capacity will be made available for the daily auction. The “use-it-or-lose-it” principle applies: the holder of the unused capacity right is still required to pay for the capacity and does not receive any proceeds from the auction. Reassignment to another user or reallocation by TSOs to another user are allowed under the IFA Rules.

Capacity is sold on a non-firm basis and users are responsible for any trading consequences as a result of changes to the interconnector capability (for imbalances incurred in the wholesale markets). Curtailments, due to planned and unplanned outages, are pro-rata for all users. An availability reconciliation applies to all capacity sold in long and medium term auctions. An availability target is set for each MW when sold; ex-post, actual availability is compared with target availability and a reconciliation is carried out, on an individual basis, taking into account curtailments and rebates paid by the TSOs (partial refunds that can be applied in case of curtailments) or additional capacity payments paid by the users, depending on whether the actual availability exceeds or falls short of the target. Normally daily capacity is sold on the basis of a 100% target availability unless there is a planned outage during the day, in which case it is sold on the basis of an availability profile (RTE and NGIL, 2005b).

An analysis of the existing allocation process conducted by RTE and National Grid for the UK-France-Ireland ERI has revealed that some aspects of it do not fully comply with the Congestion Management Guidelines (annex to the Regulation 1228/2003). Proposed changes include the application of “use-it-or-sell-it” rules on unused capacity, the introduction of hourly allocation on the day-ahead, an update of the current auction platform, the implementation of an intraday allocation and nomination mechanism, and the development of financial compensation rules in case of curtailments. RTE and National Grid are currently working towards a gradual implementation of these proposals (RTE and NGIL, 2007).

4. Market coupling in the Central-South region

This section focuses on the national trading arrangements for the day-ahead and intra-day timeframes in the Central-South (CS) region. The design of national electricity markets is not uniform. A comparative assessment of their main features can provide useful information regarding the areas where harmonization is required or difficulties may arise. It can also guide the choice of the most suitable coupling solution.

4.1 Features of day-ahead markets

All countries in the CS region have established a national PX, with the exception of Switzerland. EEX, the German PX, organizes a day-ahead market for the Swissgrid balancing area.[105] The trading arrangements in Greece are in a transitional period: our comments will refer to the market design expected to be in place in Greece starting January 2009. Borzen has published on September 9th, 2008 a document for public consideration regarding new "Electricity Market Operation Rules" (Borzen, 2008). We refer here to the arrangements in place before this date.

Those markets are compared on the basis of technical aspects (mainly traded products, clearing algorithms and timing) as well as institutional aspects. While the latter do not appear to be significantly different, a variety of approaches is found in terms of technical solutions and this aspect needs to be carefully addressed. Nonetheless, technical issues do not appear so critical as to prevent a progressive integration of the Italian market with neighboring markets, in particular assuming that a form of volume coupling is implemented.

4.1.1 Institutional framework

All PXs in the CS region organize a day-ahead auction, with physical delivery (orders refer to a specific point of injection/withdrawal). The only exception is France: products

[105] In September 2008, Swissgrid and the German EnBW Transportnetze have created a joint-owned company (Central European System Operation Coordinator - CESOC) to monitor and operate their respective networks in a coordinated manner.

traded in Powernext do not include such indication. Note that all PXs co-exist with bilateral physical trading, except in Greece, where the market will be organized as a mandatory pool.

Authorities involved in the control of day-ahead markets may include both the energy regulatory authorities and the governments through the Ministries of economic affairs or Directorates for energy. National PXs and TSOs interact closely in organizing the electricity markets.

Since January 2008, all countries in the CS Region (with the exception of Germany) have adopted common rules for the explicit auctions that allocate cross-border capacity on a yearly, monthly and daily basis. In this regard, two aspects need further investigations:

- Currently the allocation of cross-border capacity (via explicit auctions) is carried out by the Italian and by the neighboring countries' TSOs. Considering the neighboring regions, implicit allocation of cross-border capacity is carried out only at the France-Belgium border. As to internal transmission capacity Slovenia, Austria, and France are single price markets. Internal congestions, if any, are managed by the TSOs through redispatching. In turn, the Italian and Greek PXs implement market splitting in the day-ahead market. We conjecture that some legal elements governing the relationship between TSOs and PXs will have to be updated in some countries for the introduction of market coupling. In particular, the contractual framework between PXs and TSOs should be updated in order to cover new issues (shipping agent, data flows, payments and guarantees, etc.).
- Arrangements for market participants to notify import-export schedules to TSO's: these arrangements are not uniform and they will likely have to be revised, in some countries, as part of the implementation of market coupling.

4.1.2 Traded products

Products traded on the day-ahead auction are not homogeneous across markets. In Italy and Slovenia the day-ahead auction includes hourly products only. In Austria, France and Germany both hourly and block orders are included in the day-ahead auction. Block orders are generally standardized and restrictions may apply on the number and volumes traded.[106]

Market clearing in the presence of block orders implies:

- either a mechanism based on iterations between the central auction office and the different PXs (as in TLC); or
- implementing of block constraints in the central solver algorithm; or
- implementing some form of volume coupling.

Most markets enable trading of block and hourly products in continuous day-ahead and intra-day markets. This practice as well as the trading of financial products does not appear to be incompatible with coupling day-ahead markets.

Price and volume steps of sale/purchase orders are rather uniform. Prices have minimum intervals of 0.01 €/MWh. The smallest volume step is found in Italy (0.001 MWh) and the highest in Slovenia and France (1 MWh). Some form of coordination of the format of the bids/offers sent off by PX's to the Central (coupling) office will be required for market coupling.

Finally, price caps or values of lost loads are quite diverse. The lowest price cap is implemented in Greece (150 €/MWh) and the highest is found in France (3000 €/MWh) and Italy (where a recent provision introduced a 3000 €/MWh value of lost load). No specific limitation exists in Slovenia and Austria. The Greek price cap of 150 €/MWh is likely to make export to Italy attractive in tight supply conditions, as prices in the Italian Brindisi zone might result higher in that condition. Maximum prices of the order of 3000 €/MWh represent the value of lost load, rather than an attempt to limit price spikes in the market (as in Greece). For the functioning of market coupling, coordination on the value of lost load might be crucial in times of generation tightness.

[106] For instance, on the Austrian EXAA, 11 standardized products are traded and the maximum volume tradable per block is 500 MW. On Powernext, participants have the possibility to submit block orders that link a minimum of four hours of the day together. Each trading day, the participants are allowed to submit a limited number of block bids and the maximum tradable quantity is fixed at 100 MW per block.

4.1.3 Timing

1

In terms of timing, market coupling requires all involved countries to synchronize at least in the following deadlines:

- collection of demand and supply bids by all involved markets must end before the coupling algorithm is run;
- publication of domestic market outcomes by each national PX must take place after the market coupling results are calculated.

Figure 4.1 shows a generic timing for market coupling implementation and Figure 4.2 the timing of the relevant countries. There is an apparent heterogeneity both in terms of gate closure times and in terms of computation time. A strong effort in the direction of the harmonization of timing is very likely to be necessary in order to proceed with market coupling in the Central-South region. As a reference, the computation time in the TLC is 15 minutes.

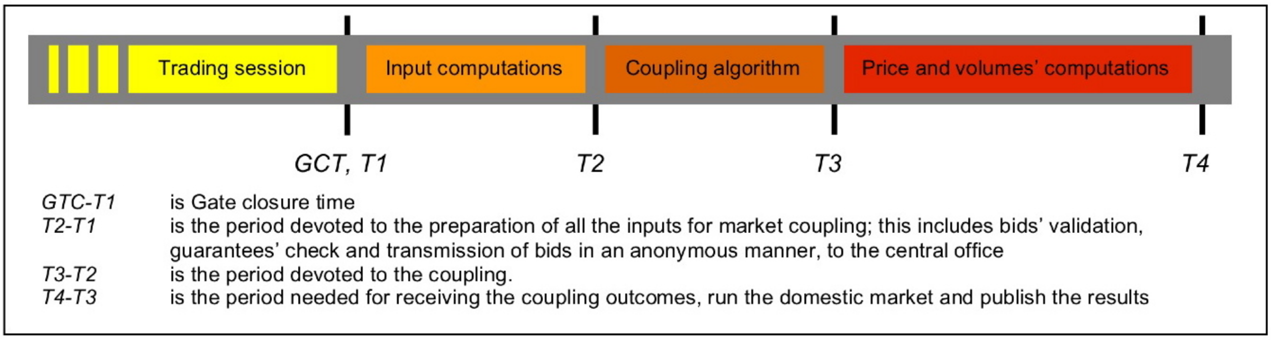

Figure 4.1 - Generic timing of market coupling

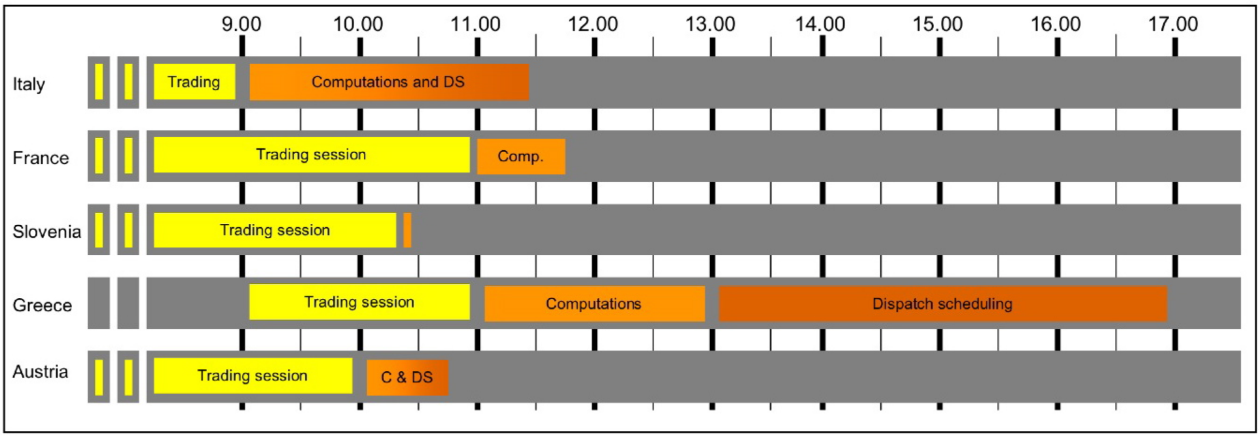

Figure 4.2 - Timing for the different countries in the Central-South region

4.1.4 Algorithm

Market clearing algorithms vary substantially from country to country. These differences might impact on the choice of the market coupling model to be implemented in the CS region. In particular, if these national features are to be maintained, it appears that the most viable solution for market coupling in the short-term is some form of volume coupling. Coordination with other coupling mechanisms, for instance TLC which includes France, is a further question that will have to be dealt with. We focus here on the first issue and leave the second to further research.

As for the relevant differences in clearing algorithms, we have already highlighted the impact of block bids, currently implemented in several markets in the region.

A second important feature of the market design of some of the involved national markets is the use of locational prices to allocate the available domestic transmission capacity (market splitting). This is the case for Italy and Greece, while the other day-ahead markets do not provide locational signals. When considering the introduction of price and volume coupling the main question is whether the current market splitting would induce problems in terms of coherence of cross-border flows. Indeed, running separately a national market splitting would imply the recalculation of the bidding areas prices which were calculated by the central coupling algorithm. Such a recalculation could lead to deviations between the final national prices set by the internal market splitting, and the prices calculated by the central coupling algorithm, from which cross-border flows are determined.

A third design feature of the involved markets may have a significant impact on the coupling model: in Italy and Greece the price applied to power withdrawals is nationwide uniform, calculated as the weighted average of Zonal prices (in Italy: PUN, Prezzo Unico Nazionale); Zonal prices clear the supply side of the market.[107] This is an issue because price deviations might arise between the central and national solutions and also because of a constraint on the time to run the local market clearing algorithms. The market clearing algorithm implemented in Italy takes around 20 minutes to run, which would make it

[107] In Italy the PUN is applied only to withdrawal points in national geographic areas, thus excluding import and export orders that are traded at zonal prices. This is consistent with the idea that the PUN is a cross-subsidization method introduced for the benefit of national consumers and, for this reason, it should not concern trades that occur cross-borders (indeed, they are not included in the calculation of PUN). Similarly, in the Greek market, imports are paid the zonal price at the relevant interconnection. Therefore, these arrangements do not create discrimination as far as foreign trade is concerned.

impossible for Italy to generate a full-information NEC in a time consistent with the functioning of a full-fledged coupling solution. Based on the information that the expected time to solve the unit commitment problem in Greece is around two hours, we guess that Greece would face a similar issue.

The issues of price discrepancies and timing deserve further analysis. The Italian reference price for the market coupling is the price of the neighboring Italian zone to the border in which the coupling mechanism is implemented (North zone for France, Switzerland, Austria and Slovenia, Brindisi zone for Greece). Therefore, the NEC that the IPEX will calculate is the curve describing the prices of this relevant zone with different quantities of import/export from/to the coupled market. The IPEX NEC could be calculated by considering the domestic congestions and using a "pure" market splitting algorithm (without considering the impact of the PUN) in order to reduce the time of the computations.

Even so, at the end of the coupling process, when the IPEX algorithm is run and the PUN is calculated, minor price discrepancies may arise between the relevant Italian zonal price calculated by the market coupling and the zonal price calculated by the IPEX. This is due to the fact that the PUN is a price that must fulfill, at the same time, two requirements:

- to be the weighted average of zonal prices;
- only demand bids whose price is not lower than the PUN are accepted.

To ensure that the above mentioned requirements are satisfied, the PUN cannot be calculated ex-post, as the weighted average of the zonal prices calculated by a "pure" market splitting algorithm, but it needs to be embedded in the algorithm itself.[108]

Our conjecture is that those price discrepancies might be rather small, since:

- the Italian Net Export Curve sent to the coupling office would embed the solution of domestic congestions;

[108] Let's assume to run a market with two zones, A and B and let's assume that by running a "pure" market splitting algorithm the outcomes are: (1) 40 €/MWh in the zone A; (2) 50 €/MWh in zone B. Let's assume that the weighted average of the zonal prices is 48 €/MWh. Therefore, if the PUN were calculated ex-post, it would be equal to 48 €/MWh. In case we had a demand bid submitted in the zone A at a price that is higher than 40 €/MWh, but lower than 48 €/MWh, this bid would be accepted on the basis of the pure market splitting algorithm, but it would be forced to pay a price higher (48 €/MWh) than the one it was willing to pay. Vice-versa, if we had a demand bid submitted in the zone B at a price that is lower than 50 €/MWh, but higher than 48 €/MWh, this bid would not be accepted on the basis of the pure market splitting algorithm, even if the average price calculated ex-post is lower (48 €/MWh) than the one it was willing to pay. Therefore, to avoid those paradoxical outcomes, the PUN cannot be calculated ex-post and the market splitting algorithm must be able to find a heuristic solution that detects the optimal PUN.

- demand is typically entered in the Italian market at prices far above equilibrium level (most often bids are at Value of lost load), so we don't expect differences in the equilibrium output determined by the central office and by GME .

For this reason most of the times the Italian PUN is equal to the weighted average zonal price that would be calculated ex-post if IPEX used a "pure" market splitting algorithm.

In addition, we envisage also that some form of mechanism might be implemented to address issues arising from the fact that more than two different Italian market zones (North and South) would host import/export flows.

4.1.5 Guarantees and settlements

Guarantees required to participate in the national markets are similar in all countries. In general, operators must submit a bank letter of credit that will cover the counter-value of the outstanding debt. As an alternative, the Slovenian PX allows participants to provide a cash deposit (100% coverage of purchase bids). The Austrian PX accepts a bank guarantee, cash deposits or securities. In France, a guarantee deposit is required before any trading takes place. The guarantee deposit is daily adjusted, according to the average amounts of electricity purchased on the market.

Settlement times are very diverse. In Slovenia and France, settlement occurs few days after the trading day. By contrast, payments in Italy are settled within the 16^{th} working day of the second month following the month of trading. In Greece and Austria, payments are settled once a month. Harmonizing settlements to the shorter time frames would require PXs in Italy, Austria and Greece to settle cross-border trades in advance of domestic trades. Consequently, these PXs would have to take non zero net financial positions.

4.1.6 Market monitoring, transparency and liquidity

Market monitoring is generally a function carried out by the energy regulatory authorities, with the support of PXs. As for transparency, publication of aggregated supply and demand curves, traded volumes and prices and others indices is a common practice. Although it might not be a major issue in view of market coupling, it is worth to notice that

Greece intends to publish all data made available to the TSO and concerning the day-ahead market. In particular, all auction participants will be able to see all submitted bids and the complete results of the market (not only traded volumes and prices).

Our evaluation is that market coupling could be implemented without imposing any change on the current national rules for publication of information regarding national transactions.

Market liquidity provides an indication of the amount of trades that occurs on the organized markets vs. bilateral trading. Italy reports a high liquidity (65% day-ahead trade/total trade, including bilaterals). For France and Austria the liquidity is around 7% and 4% of consumption respectively. The Slovenian market is known to present an extremely low liquidity in the day-ahead auction (0.009% of consumption in 2006). Our evaluation is that the market coupling is going to increase the liquidity of the involved markets.

4.2 Features of intra-day markets

Two intra-day markets are operational in the CS region: the EEX Intraday market in Germany and the Powernext Intraday market in France.[109]

On EEX individual hours and block orders are exchanged in a continuous trading system. The intra-day market is open 24 hours a day, 7 days a week and participants can trade the individual hours of the current day up to 75 minutes prior to the beginning of delivery; trading for the following day begins at 3.00 pm. The delivery area comprises the four balance areas in Germany. The trading participants have to present a balance agreement in a TSO zone of their choice in order to be able to participate in the Intraday Market.

On Powernext hourly and block products (11 standardized blocks or user-defined blocks linking at least 2 consecutive hours) are exchanged in a continuous trading system. Trading times are from 7:30am to 11:00pm, 7 days a week; trading for the following day begins at 11:30 am. Volume and price steps are respectively 1 MWh and 0.01 €/MWh. The value of the settlement is equal to the clearing price multiplied by the volume traded; the

[109] In April 2008 ETSO has published a report that introduces a reference model for continuous-trading-based intraday markets (ETSO, 2008).

central counterparty ensures the financial security of the transactions and is responsible for collecting payment from net buyers and distributing it to the net sellers. Powernext nominates all physical commitments of the balance responsible members and deliveries are made to RTE.

5. Capacity allocation on the Italian borders

The current mechanisms for cross-border capacity allocation between Italy and neighboring countries are based on an explicit auction mechanism. One of the positive outcomes of the work carried out in the ERI framework was the publication, in January 2008, of a common document with harmonized Auction Rules for all borders (Terna, 2007), approved by the transmission operators of Italy, France, Switzerland, Austria, Slovenia and Greece.

Energy regulators have clearly stated that, for the day-ahead time frame, *"implicit auctions are more efficient than explicit auctions and should be the target mechanism for all regions"* (ERGEG, 2007, p. 12). Indeed, as illustrated in Section 5.1, studies of explicit cross-border auctions have found different types of inefficiencies. A preliminary study of inefficiencies on for the Italian market was carried out by Carboni (2008) and Section 5.2 reports the main findings. One must be careful, however, not to identify inefficiencies in explicit auctions with efficiency gains from market coupling. One thing is to eliminate inefficiencies in cross-border allocation; another to evaluate the costs and benefits of modifying the design of a national market, with the objective of enlarging its geographical scope, or to identify the optimal market integration model. As illustrated in Section 5.3, the study of potential efficiency gains deriving from market coupling has been addressed in the literature; as for the case of the Italian market it remains an issue for further research.

5.1 Inefficiencies in explicit auctions: literature review

Studies of explicit cross-border auctions have addressed different European markets and have highlighted different inefficiencies.

Newbery and McDaniel (2002) observe the inter-connector auctions on the German-Dutch border and on the England-French border. They find two types of inefficiencies. The first is related to the definition of transmission capacity: on both interconnections the capacity to be sold is defined on the gross flows, instead of netting imports and exports. The second concern imperfect arbitrage. The authors find that yearly and monthly auctions on the German-Dutch border are efficient (the 2001 auctions cleared at close to the price difference of year-ahead base-load OTC contracts in the two markets). On the contrary, the daily market is not well arbitraged (the daily auction cleared on average at a lower price than the spot market price difference). The authors suggest that this reflects the risk that the hourly spot prices are not known at the time of bidding for transmission capacity, in contrast to the yearly and monthly auctions where financial contracts can guide price discovery. Similarly, imperfect arbitrage is found comparing annual and quarterly auction costs for the year 2002 on the England-French border, with the corresponding spot price difference. Average auction costs from France to England are lower that average price differences, suggesting either risk aversion, high transaction costs or illiquidity as reasons for the imperfect arbitrage. The authors conclude that, to improve the efficiency of the auctions, energy and transmission capacity markets should clear simultaneously.

Tornquist (2006) studies cross-border capacity allocation, from 2002 to 2005, on the borders Belgium-The Netherlands, Denmark-Germany, and France-UK. He finds a first inefficiency in the fact that the level of firmness of capacity depends on the time duration over which the capacity was allocated (TSOs normally curtailed daily capacity first, then monthly capacity and finally yearly capacity). Moreover, the compensation, in the event capacity is curtailed, is not at the full market spread but rather on prices paid for the capacity. Finally, the author calculates average prices for transmission capacity and shows that, in most cases, annual and monthly auctions appear to be valued more highly than daily auction. He concludes that there is a need to offer more capacity of longer-term duration, and then allow trading of these capacity rights in secondary markets.

Kristiansen (2007a; 2007b) analyses auction prices for the cross-border auctions between West Denmark and Germany and between East Denmark and Germany (the Kontek cable). Against his expectations, but in line with Tornquist (2006), the author finds that annual and monthly capacity is generally valued more than daily capacity (but traded quantities are higher in daily auctions). Moreover, he finds that the daily, monthly and annual capacity auctions do not reflect the value of the underlying asset as specified by the appropriate valuation of the energy price differentials between West Denmark and Germany and East Denmark and Germany. As such the explicit auction procedure is not cost efficient.

The European Commission "DG Competition Report on Energy Sector Inquiry" notes that, although on many congested interconnectors TSOs make use of explicit auctions for cross-border capacity allocation, *"This mechanism is considered not to be satisfactory by a* *number of respondents in the Sector Inquiry, because it suffers from the time lag between capacity allocation and wholesale market clearance" (EC, 2007, p. 183).* The Report analyses the case of the interconnection between the Netherlands and Germany. The study reveals that in many hours during 2004 (40 percent of all observed hours) capacity was nominated from Germany to the Netherlands while prices in Germany where higher than in the Netherlands. This result is clearly not rational. One of the explanations for these economically inefficient outcomes is that the deadline for the day-ahead interconnector auction ended before the German (EEX) and Dutch (APX) energy market cleared. The financial loss resulting from underutilization plus incorrect utilization (wrong sign nominations) of interconnector capacity ("Estimated value of unused cross-border capacity") was found to be worth almost 50 million Euro for the Dutch - German border in 2004 (equal to 46% of the total value of the interconnector capacity). Higher price volatility on APX in 2003 resulted in higher financial losses.

According to the Report, the analysis illustrates that *"although explicit auctioning is theoretically and with perfect foresight an efficient mechanism and it is in practice compatible with Regulation 1228/2003, it has efficiency deficits compared to implicit auctioning especially where intraday and balancing markets are illiquid" (EC, 2007, p.186).* An important, additional advantage of implicit auctions is that netting becomes more feasible. Hence, introducing implicit auctions may increase the available capacity significantly. Finally, due to the arbitrage errors made by market participants, incorrect signals prevail regarding the value of interconnector capacity. This also leads to incorrect incentives to attract new investments into interconnector capacity (EC, 2007).

5.2 Inefficiencies in explicit auctions: Italian borders

Carboni (2008) has identified three areas where inefficiencies arise in the current arrangements for capacity allocation on the Italian borders:

- operational risk;
- trading risk/cost;
- use of transmission capacity.

Operational risk may arise because the day-ahead stage of the explicit auction occurs in a very tight time frame (Figure 5.1). In this time period Terna and GME need to exchange information to coordinate the capacity market (managed by the TSO) with the energy market (managed by the PX). In the same time interval, market operators are requested to prepare and submit their bids and offers, taking into account the outcome of the transmission capacity auction. Allocating all the day-ahead cross-border capacity via an implicit auction will reduce this complexity and thus the operational risk (Carboni, 2008).

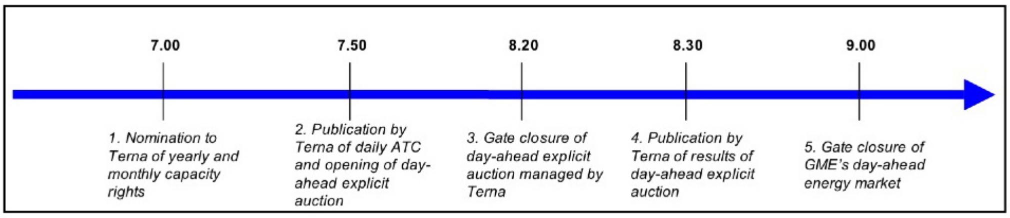

Figure 5.1 - Time schedule of day-ahead auctions managed by Terna (Carboni, 2008)

Secondly, having separate day-ahead capacity and energy markets forces market operators to coordinate their capacity and energy positions, taking into account the risk that arises from trading these products on two different markets. Moreover, operators incur trading costs (guarantees, IT, and so on) for two different trading platforms. These risks and costs may thus be reduced by eliminating the day-ahead capacity auction (GME, 2008).

Thirdly, explicit auctions introduce inefficiencies in the use of cross-border capacity. Carboni (2008) analyzed cross-border transactions on the Italian borders for the year 2007, using price and flow data for countries having a liquid index of hourly day-ahead energy prices: France (Powernext), Austria (EXAA) and Switzerland (EEX-CH). The reference Italian price was the price in the IPEX Northern zone.[110]

Figure 5.2 shows, for each hour in 2007, the spot price differences between Italy (IPEX Northern zone) and France (Powernext) on the horizontal axis and cross-border schedules on the vertical axis (imports from France in the upper semi-plane). It reveals that in many hours during 2007 capacity was scheduled from France to Italy while prices in Italy were lower than in France. This result, shown in the IV quadrant, is not rational, but it can be

[110] In 2007, cross-border schedules on the borders with France, Austria and Switzerland amounted to 45 TWh: 14% of the energy sold on the IPEX, the Italian PX (including OTC) and 33% of the energy sold on the Northern zone of IPEX, including OTC (Carboni, 2008).

explained by the fact that prices in Italy are generally expected to be higher than in France.[111] The deadline for the day-ahead auction ends before the energy market clears and market participants have to place auctions bids based on expected market prices. The consequence is that explicit auctions do not lead to an optimal use of scarce interconnection capacity (EC, 2007). A similar not rational behavior is found also on the borders with Austria and Switzerland (Carboni, 2008).

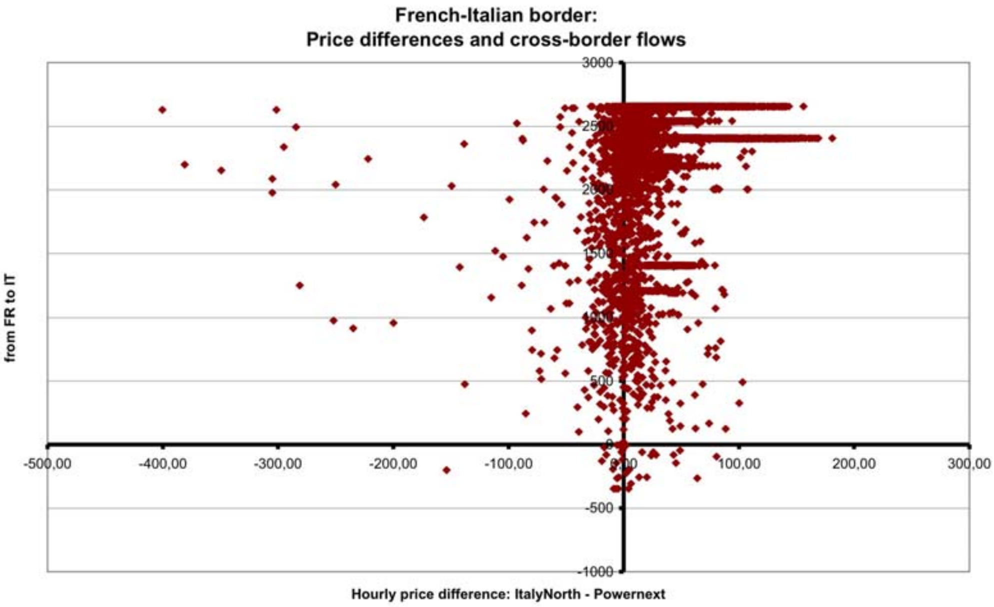

Figure 5.2 - Price difference (€) vs. cross-border schedule (MW) (source of data: Carboni, 2008)

Inefficiencies in cross-border schedules were also quantified by means of the "estimated value of the unused cross-border capacity" (EC, 2007). This value is defined in Carboni (2008) as the product of *(i)* the difference between hourly NTC and hourly day-ahead net scheduled flows times *(ii)* the price differential between the two market areas. Note that this value represents only a proxy of the inefficient use of the cross-border capacity: it relies on the rather strong assumption that prices, liquidity and bidding behavior of the operators would not change in case of market coupling.

The value of the unused cross-border capacity on the Italian borders France, Austria and Switzerland was estimated to be 30 Mln € for the period January-September 2007 and

[111] Also the II quadrant represent an irrational outcome.

132 Mln € for the period October-December 2007. A comparison with the price differentials between the market prices shows clearly that in the last quarter of 2007 (where more that 80% of the estimated value of the unused cross-border capacity was concentrated) price differences changed signed more frequently than in the rest of the year (Carboni, 2008).

In summary, the analysis of the Italian, Austrian, French and Swiss cross-border schedules illustrates that inefficiencies are significant when market operators are not able to correctly predict the sign of the price difference. On the contrary, when expectations of market prices are verified, the estimated value of unused capacity is negligible (Carboni, 2008).

With implicit auctions cross-border capacity is allocated as a function of the price differential in the two market areas: cross-schedules always go from the low price areas to the high price area and congestion revenues do not arise if there are no price differentials. An additional advantage from introducing implicit auctions for the day a-head time frame is the netting of the flows in opposite directions, which significantly increases the cross-border capacity made available to the market (and consequently it reduces cross-border congestion rents).

5.3 Efficiency gains from market coupling: literature review

A few studies have explored the efficiency gains that might derive from coupling national electricity markets. This work, which is mainly based on large scale Cournot market models, provide interesting results; however it is also sensitive to modeling details.

Hobbs et al. (2005) note that stronger market linkages are generally viewed as being economic beneficial for several reasons. These include: granting access to lower cost imports, enhancing export opportunities for producers, and, in the long run, decreasing the amount of reserve required to maintain a given level of reliability. Further, closer linkages can dilute local market power. Of course, the benefits of linking markets are not uniformly distributed; for instance, consumers in regions that export more power may suffer price increases.

In their work, the authors study the expected efficiency gains for the, at the time only proposed, coupling of the Belgian and Dutch markets. They argue that market coupling could lead to *(i)* more efficient use of inter-country transmission by counting only net flows against transmission limits, *(ii)* by improving access to the Belgian market, and *(iii)* by eliminating the mismatch in timing between interface auctions and the energy spot market. They used a large-scale Cournot market model, which accounted for the region's transmission pricing rules and limitations, to simulate market outcomes with and without market coupling and quantified the first two of the three expected impacts of coupling. Their results indicate that an improvement in social surplus is attainable, unless market coupling encourages the largest producer in the region to switch from a price-taking strategy in Belgium to a Cournot strategy due to a perceived diminishment of the threat of regulatory intervention. Whether market coupling may benefit Dutch consumers also depends on the behavior of this company.

Among other similar studies, Petrov et al. (2003) use a supply function equilibrium approach to simulate a coupled Benelux market, and conclude that prices would decrease. However, their analysis did not explicitly represent transmission constraints. In another study on the Benelux region, Neuhoff (2003) finds that consumers in all countries are better off under market coupling, in contrast to Hobbes at al. (2005) conclusion that Dutch consumers could be worse off. Indeed, Neuhoff et al. (2004) note that differences in the treatment of transmission allocation and pricing models can make a significant difference in the conclusions of different models, even if they all are Cournot in energy sales.

6. References

APX, Belpex, and Powernext (2006), “Trilateral market coupling”, *Technical report, APX, Belpex, Powernext*.

Bjorndal M. and Jornsten K. (2007), “Benefits from coordinating congestion management - the Nordic power market”, *Energy policy*, 35,1978–1991.

Borzen (2008), “Rules for the operation of the electricity market (draft)”, *Technical report, Borzen*.

Carboni F. (2008).

CW MC Project (August 2008), “Implementation Study. Technical Report”, Available from: http://www.apxgroup.com/index.php?id=186.

De Jonghe C., Meeus L. and Belmans R. (2008), “Power exchange price volatility analysis after one year of Trilateral Market Coupling”, *5th International Conference on European Electricity Markets*, Lisbon, 28-30 May 2008.

Dupuy M. (2008), “Electricity markets: Balancing mechanisms and congestion management”, *Master’s thesis, Royal Institute of Technology, Stockholm, Sweden*.

EC (2007), “DG Competition Report on Energy Sector Inquiry”, Available from: ec.europa.eu.

EMCC (2008), “European market coupling company”, http://www.marketcoupling.eu.

ERGEG (2008a), “ERI coherence and convergence report”, *Conclusion paper, ERGEG*.

ERGEG (2008b), “New version of IFE rules – Public consultation”, *Conclusion paper, ERGEG*.

ERGEG (March 2007), “ERGEG Regional Initiatives Annual Report”, *Technical report, ERGEG*.

ETSO (April 2008), “Reference model for cross-border intraday markets”, *Technical report*, *ETSO*.

EuroPEX and ETSO (2004), “Flow-based market coupling”, *EuroPEX and ETSO Technical report*.

Glachant J. M. and Pignon V. (2005), "Nordic congestion's arrangements as a model for Europe? Physical constraints vs. economic incentives", *Utilities policy*, 13, 153–162.

Glachant J. M., Belmans R. and Meeus L. (2006), "Implementing the European Internal Energy Market in 2005-2009", *European Review of Energy Markets*, 1(3), 51–85.

Haubrich H. J. (2006), "Technical issues regarding open market coupling", *Technical report, Institute of power systems and power economics. Energy Research Institute*, *Aachen, Germany.*

Hobbs B. F. and Rijkers F. A. M. (2005), "The more cooperation, the more competition? A Cournot analysis of the benefits of electric market coupling", *Cambridge Working Papers in Economics, CWPE 0509*, *University of Cambridge*.

Kristiansen T. (2007a), "A preliminary assessment of the market coupling arrangement on the Kontek cable", *Energy policy*, 35, 3247–3255.

Kristiansen T. (2007b), "An assessment of the Danish-German cross-border auction", *Energy policy*, 35, 3369–3382.

Lundgren J., Hellstrom J. and Rudholm N. (2008), „Multinational electricity market integration and electricity price dynamics", *5th International Conference on European Electricity Markets*, Lisbon, 28-30 May 2008.

Meeus L. (2006), "Power Exchange - Auction trading platform design", *PhD thesis*, *Katholieke Universiteit Leuven, Belgium*.

Meeus L. (April 2008), "Regional issues in cross-border trading", *Lecture given at the Advanced Training Course on Electricity Markets. Florence School of Regulation, Florence.*

Meeus L. and Belmans R. (2008), "Electricity market integration in Europe". *Mimeo*.

Meeus L., Vandezande L., Cole S. and Belmans R. (2008), "Marker coupling and the importance of price coordination between power exchanges", *Energy*, forthcoming.

Neuhoff K. (2003), "Effect of integrating Dutch and Belgian electricity market on Dutch prices", *Mimeo, Dept. of Applied Economics, Cambridge University, July 24.*

Neuhoff K., Barquin J., Boots M.G., Ehrenmann A., Hobbs B.F. and Rijkers F.A.M. (2004), "Network-constrained models of liberalized electricity markets: the devil is in the details", *European IAEE 2004 Annual Meeting*, Zurich.

Newbery D. and McDaniel T. (2002), "Auctions and trading in energy markets: an economic analysis", *DAE Working Paper WP 0233, University of Cambridge.*

NordREG (2007), "Congestion management in the Nordic region. a common regulatory opinion on congestion management", *Report 2/2007, Nordic Energy Regulators*.

NordREG (2008), "Nordic market report 2008", *Report 3/2008, Nordic Energy Regulators*.

NorNed (2008), "Auction rules. Technical Report", *NorNed*, Available: http://www.norned-auction.org.

Petrov K., Scarsi G.C. and van der Veen W. (2003), "Modeling Strategic Bidding Behavior in Power Markets", *Research Symposium on European Electricity Markets*, The Hague, Sept.

RTE and NGIL (2005a), "IFA Rules", *Technical report, Issue 6.0, 31*.

RTE and NGIL (2005b), "IFA User Guide", *Technical report, 7*.

RTE and NGIL (2007), "IFA: congestion management proposals", *UK-FR-IE Electricity Regional Initiatives report*.

RTE and RED (June 2007), "Capacity allocation rules for the France-Spain interconnection (IFE rules)", *Technical report, Version 2.0*.

Terna (2007), "Access rules to France-Italy, Switzerland-Italy, Austria-Italy, Slovenia-Italy, Greece-Italy interconnections", *Technical report*.

Tornquist J. (2006), "More transmission capacity for European cross border electricity transactions", *5th Conference on Applied Infrastructure Research, Berlin, 2006.*