A system for congestion management in the Netherlands

Assessment of the options

Rudi Hakvoort, Dan Harris, Jos Meeuwsen, Serena Hesmondhalgh

Zwolle, 24 June 2009

Prepared by:

D-Cision B.V.
P.O. Box 44
NL–8000 AA Zwolle
The Netherlands
Hr.a.hakvoort@d-cision.nlH

The Brattle Group
Halton House, 20-23 Holborn
London EC1N 2JD
United Kingdom
Hdan.harris@brattle.co.ukH

OVERVIEW

MANAGEMENT SUMMARY

Congestion in the Dutch electricity grid

Planned electricity production capacity in the Netherlands has rapidly increased in recent years, especially in areas as Eemshaven, Maasvlakte and Westland. Since reinforcing the transport networks takes longer than building new production capacity, it is expected that generating capacity will exceed the available transmission capacity for some years, and that congestion will result. Accordingly, the Dutch Ministry of Economic Affairs (*Ministerie van Economische Zaken*) has commissioned *D-Cision* and The Brattle Group to develop a national system of congestion management.

Based on the current data, TenneT expects congestion to occur mainly between 2012 and 2017, the average amount of congestion costs estimated to lie in the order of 100 million euros per year. TenneT is keen to stress that these numbers are only rough estimates and subject to significant uncertainty. However, TenneT's estimates are consistent with the costs in other European countries suffering from network congestion. International experience also illustrates that, for various reasons, estimates of congestion costs are subject to large errors.

Objective of the study

D-Cision and The Brattle Group were asked to develop a number of alternative congestion management mechanisms, which should be evaluated against the following criteria:

- *Technical:* Technical criteria concern the network requirements which need to be met by a congestion management system in order to be applicable for the Dutch electricity network.
- *Economic:* Economic criteria refer mainly to the incentives provided by a congestion management system and to the possibility to achieve economic efficiency.
- *Policy:* The policy criteria involve the compliance of the congestion management models with the existing policy on e.g. network connection and network use as well as with the current debate in Parliament on the priority access for renewable energy.
- *Regulatory:* Regulatory criteria mainly concern the consequences of a congestion management model with respect to the proper development of the electricity market and the electricity network.

Stakeholder consultation

During the process to develop the congestion management method two workshops were organized with stakeholders and several interviews were conducted.

In summary, most generators showed a clear preference for system redispatch, mainly based on the argument that since congestion is caused by insufficient investments by network operators, it should be resolved by the network operators as well. Furthermore, generators prefer socialization of the costs of congestion management contrary to consumer organisations who prefer charging these costs to generators. Finally, generators are of the opinion that the Netherlands is too small for locational signals given the limited amount of appropriate locations to site new plant. Given the total capacity of new plant announced – which far exceeds Dutch peak load – at some point the issue of locational signals comes down to the question how much generators serving exports can be hosted by the Netherlands before the congestion costs induced are charged to the parties causing the congestion.

Models analyzed

We assessed six alternative models, but the ones that scored most highly are:

1. *System redispatch*: the system operator is responsible for redispatching plant, which is done via a bidding process. Constrained off plant are fully compensated. The costs of the system are either partly socialised, or could be allocated to generators in the constrained area.
2. *Market redispatch:* Generators are selected to be constrained off on the basis of their declared total costs for solving the constraint. Generators are responsible for the costs of contracting constrained on power downstream of the constraint, and bear the costs of lost profits from being constrained off. Moreover, non-constrained off generators in the constrained zone pay the marginal constraint cost. This not only incentivises generators to 'reveal' the cheapest combination of plants to manage constraints, but also provides a locational incentive for plant investments. The revenue stream can be used to fund network expansion, or be given to generators downstream of the constraint or consumers within the constrained zone. Under conditions of severe market power within the congested area, this revenue stream may be reduced to zero. However, there are no direct costs for the system operator or consumers with or without market power.
3. *Hybrid model:* As the name implies, this is a mix of system redispatch and market redispatch. The system operator manages a market to constrain off plant, and is responsible for contracting plant downstream of the constraint to replace the constrained off power. However, unlike in system redispatch, the generator is not compensated for being constrained off. Non-constrained off generators in the constrained zone pay the marginal constrained off cost. The system operator pays for the costs of constraining on plant, and funds these costs using the revenues from the generators in the constrained zone. Any remaining revenue could be used for network expansion, or be given to generators downstream of the constraint or consumers within the constrained zone. Under conditions of severe market power within the congested area, this revenue stream may be reduced to zero, and the system operator would need to fund the constrained on costs from another source. This could include revenues from an area with more competition upstream of the constraint.

Assessment of the models

While there are many criteria, two emerge as the most important in terms of differentiating between these three models. The first is *cost*, and in particular the costs for consumers that could arise if market power is exercised in the constraint market. Experience in other countries shows that market power is a common problem in constraint markets, and that constraint costs in some countries that use redispatch have risen sharply in recent years. We understand that the Dutch parliament has expressed a view that consumers should not bear the costs of congestion. The second issue is *long-term economic efficiency*, and in particular ensuring that generators have adequate incentives to locate in areas that will not further increase congestion.

The main disadvantage of **system redispatch** – which is the standard congestion management mechanism applied in other jurisdictions – is that it is vulnerable to market power abuse and escalating costs. This could be addressed by, among other things, allocating the costs of the system to generators, so that there is no net cost for consumers. However, even when allocating costs to generators, the incentives to locate outside of the constrained area would be relatively weak, and longer term congestion or increased transmission costs could result.

Under **market redispatch** generators will not exaggerate constraint costs; since they are not compensated for constraint costs there is no point in doing so. However, the market redispatch model has the disadvantage that, since generators must solve congestion bilaterally, transaction cost are relatively high. Moreover, the costs allocated to generators are likely to be more than required to achieve long-term efficiency in terms of plant siting.

The **hybrid model** achieves a good balance between the various evaluation criteria. It avoids market power problems in the constrained off market, and provides balanced incentives for generators to locate outside of the constrained zone, which should improve long-term efficiency. The scheme should not involve net costs to consumers, since generators upstream of the constraint should fund constrained on payments, and transactions costs are low as trades are organised via a central pool.

Recommendation

We recommend the Ministry of Economic Affairs take action to implement the hybrid model for managing congestion in the high-voltage grid.

The hybrid redispatch scheme is – similar to system and market redispatch – also able to accommodate other policy goals, such as giving priority access for renewable energy and avoiding discrimination between new and old plant. While we recognise that reducing the effective price for generators within the congested zone increases regulatory risk, this is proportional to the size of the congestion problem, which requires relatively significant changes if costs are to be controlled, and the errors of other markets to be avoided.

1 Introduction

1.1 Background

Planned electricity production capacity in the Netherlands has rapidly increased in recent years, especially in areas as Eemshaven, Maasvlakte and Westland. Since reinforcing the transport networks takes longer than building new production capacity, it is expected that generating capacity will exceed the available transmission capacity for some years, and that congestion will result. Accordingly, the Dutch Ministry of Economic Affairs (*Ministerie van Economische Zaken*) has commissioned *D-Cision* and The Brattle Group to develop a national system of congestion management.

1.2 Process followed

During the process to develop the congestion management mechanisms which are being assessed in the present report, *D-Cision* and The Brattle Group have performed several stakeholder consultations:

- The major stakeholders, both from the industry and consumer groups or their representative organisations, have been interviewed at the start of the project.
- Two workshops were organized, the first on the assessment criteria to be applied to a congestion management mechanism, the second on the assessment of two opposite approaches (market-based versus system-based congestion management mechanisms).
- Finally, the stakeholders have been invited to submit written responses to the analyses proposed by *D-Cision* and The Brattle Group during the workshops.

The reaction provided by the stakeholders has been duly taken into account in the final assessment of the congestion management models.

1.3 Objectives of the report

This report aims to address the following issues:

a. To assess possible congestion management models and select the most appropriate model(s) for solving national congestion in the Netherlands.

b. To specifically assess any economic constraints on cost-pass through, cost socialization and incentives for generators and network operators.

c. To analyse the applicability of congestion management for all congestion in the Dutch high-voltage electricity grid. Applicability for congestion in the regional networks falls outside the scope of the present report.[1]

d. To discuss complementarities of the congestion management mechanisms with the policy preference to implement priority access for renewable energy.

1.4 Assumptions underlying this study

This study is concerned with the design of a system to *manage* congestion, once it occurs. It is not the aim of the study to review the connection policy or the way in which the network is planned. The Dutch government has adopted a policy to allow plant to connect to the grid without waiting for network reinforcements, so that consumers can optimally benefit from the increased competition that more generating capacity delivers. However, due to this policy, in

---

[1] Given the application of the congestion management method to the high-voltage network only, the network operator managing the congestion is referred to as the 'system operator'.

some parts of the grid congestion may occur. It is not the aim of the present report to assess the benefits or drawbacks of this policy or other matters related to network planning.

The Ministry of Economic Affairs has stipulated that the congestion management mechanism must:

- *Allow priority access for renewable energy.* Any congestion management policy must not interfere with the policy that electricity production from renewable energy sources should have priority grid access.
- *Not discriminate between old and new generators.* The congestion management mechanism should treat existing and new generators the same. For example, it would not be possible to have a congestion management policy that charged new and existing generators in the same location different amounts for congestion.
- *Ensure optimal use of the transmission network.* The congestion management mechanism should facilitate electricity flows in the transmission network reflecting the maximum value of electricity trade while complying with network safety standards.
- *Be quickly implementable.* Given the 'threat' of upcoming congestion quick development of a national system for congestion management implies that participation into the congestion management scheme by foreign generators via cross-border interconnectors is not included.

Moreover, there is a general view in parliament that the costs of congestion management should not be socialised among consumers, but rather targeted at generators or those that directly cause congestion.

2 Congestion and congestion management

2.1 Congestion in the Dutch electricity network

2.1.1 Network transports and network congestion

Congestion arises when the demand for transmission capacity in a (part of a) transmission system is larger than the available transmission capacity. Recently, in many electricity systems in Europe the demand for new transmission capacity appeared to grow faster than originally foreseen by the grid operators. This is mainly caused by (plans for) new large-scale power plants at (often concentrated) locations in the network, and the growth of renewable energy in locations often remote from demand. The network operators could no long provide sufficient transmission capacity in time, basically since the construction time (including the time needed to obtain the necessary permits) for transmission lines generally exceeds the commissioning time for new generators. The result is an (at least temporary) lack of sufficient transmission capacity to accommodate all transports requested by the market.

In principle, (expected) network congestions in the Netherlands are caused by a combination of mutually interrelated issues, among which:

- *Natural discrepancy between realisation times of new generators on the one hand and new transmission capacity on the other*: In general, a new power plant can be realised much faster than new transmission capacity. This mainly applies to (extra) high-voltage transmission lines and to a lesser extent to medium-voltage distribution lines.
- *Sudden demand decreases:* Incidentally, as in Zeeland, congestion is (partly) caused by a sudden decrease in consumption of large industrial customers, apparently the consequence of the present economic crisis. Congestion occurs when the network capacity is insufficient to transport the resulting excess supply to other parts of the system.
- *Network companies facilitating market development:* According to the existing legislative framework, network operators facilitate market developments. The question is when network investments should be made: at the moment generators announce their intentions to invest (in which case the network is reinforced before certainty is obtained on the necessity of the investment) or at the moment the generator is in the process of construction (in which case the network reinforcements will come too late).
- *Cost-effectiveness of network operation:* Network operators are incentivized to operate their grids in a cost-effective way, which is implemented by the regulatory framework. For regional network operators, investments in new network capacity will only be remunerated when corresponding network use increases. This implies that network operators feel a financial incentive to deter investments until sufficient certainty on the new generation is obtained. In this way, the time-lag between network reinforcement and generator investments can easily increase.
- *Existing tariff structure in the Netherlands:* In the present Dutch network tarification approach, a robust locational incentive for new generation capacity is missing. From a network point of view, some locations are preferred to connect new capacity given the availability of transport capacity. Nevertheless, these differences are not reflected in the transport tariffs.

2.1.2 Objective of a congestion management mechanism

In principle, congestion management is a method which aims to maintain the stability of the transmission system in a situation where network flows resulting from market transactions would exceed the system stability threshold (in which case line overloading causes cascade tripping of protection systems).

Congestion can be solved by operational measures on a short term (by applying congestion management) and structurally by network reinforcements (which take time to implement).

The *operational* approach is to solve congestion by decreasing the generation in the congested area and simultaneously increasing the generation outside the congested area in order to keep the power balance in equilibrium. Such operational actions are called 'redispatch' since generators at both sides of the congestion receive instructions to change their dispatch. The generators to be redispatched can be selected in an economic manner (e.g. based on their marginal costs) or in some other way (e.g. *pro rata*). Furthermore, the 'redispatch' can be applied *ex post* (these approaches are generally referred to as 'redispatch models') or *ex ante* by taking the congestion into account when allocating transmission capacity (e.g. by auctioning transmission capacity or under market splitting).

Structural solutions for solving the congestion problems include reinforcing the grid or decommissioning generation in the congestion area. Although the present policy approach prescribes that all congestion needs to be solved, if not by changes in the generation pattern (which can only be done by the market parties themselves) then by network investments, it is not proven that these (huge) investments in new transmission capacity will always be compensated by the benefits of lower market prices (at least for Dutch consumers). However, this question falls outside the scope of the present study.

2.1.3 Congested regions in the Dutch electricity grid

The likelihood of network congestion is increasing in the Netherlands. At first, only the regions Westland, Northern Netherlands and the province of Zuid-Holland were assigned as (potential) congestion areas, but recently other regions have been added as well. At present, the following possible congestion areas can be (or have been) identified:

1. Northern part of Netherlands, specifically:
   - the region of the Noordoostpolder, and
   - Eemshaven;
2. Province Zuid-Holland, specifically:
   - the region Westland;
3. Provinces Zeeland and West-Brabant;
4. Province Noord-Holland, and
5. The national 380 kV ring in general.

Most of the congestions will take place in the (extra) high-voltage grid, which is operated by TenneT. In general, it can be concluded that a simultaneous development of small-scale generation in the distribution networks and large-scale generation in the transmission network mainly causes congestion problems at the higher voltage levels or at the connection points between the networks of TenneT and the regional distribution system operators.[2]

The amount of congestion highly depends on the number of new plans for generation capacity at certain locations. Since 2007, many new requests for network access have been received by the grid operators. Mainly three categories can be distinguished:

1. New large-scale production units (mainly in the coastal region);

[2] In the case of Westland, the commissioning of new generating capacity creates a congestion problem at the connection point between the networks of TenneT and Westland Infrastructuur.

2. Development of combined heat and power units (greenhouses);
3. Development of wind mill parks, especially in the coastal region (onshore and offshore).

According to TenneT, the total amount of announced new generation capacity exceeds 30 GW.[3] Especially compared with the present installed capacity of about 25 GW and the system peak load of approximately 19 GW[4], this represents an enormous amount of new generating capacity.

It is rather difficult to establish a reliable forecast on the expected time of occurrence of future congestion, which depends on many uncertain factors:

- Commissioning of new generation capacity announced is uncertain, especially for new coal fired power plants.[5]
- The merit order of new and existing generation units in the Netherlands and in neighbouring countries will determine which units will actually produce electricity in the future.
- Large industrial consumers show a changed load pattern in certain regions which may have an impact on congestion to occur.
- (Inter)national developments in generation planning, like installing large onshore or offshore wind mill parks may induce large transit flows and impact available transmission capacities.
- International developments in transmission planning, as the commissioning of international HVDC sea cables like NorNed, BritNed, Cobra and the interconnector between France and the United Kingdom, may change national transfer capacities as well.
- Varying weather conditions will have consequences for load demand as well as for wind production.

Presently, there is only one area which structurally exhibits congestion, namely the Westland region. The amount of congestion is nevertheless restricted to several tens of MW's for a limited period of time. The installation of additional transformers in existing and new substations will reduce the congestion problem significantly. In May 2009 an additional transformer in substation Westerlee is expected to bring relief. In 2010, a new substation De Lier is expected to be commissioned, which should completely resolve the congestion problem. Furthermore, congestion is being relieved by ongoing investments and operational measures in the 25 kV network as well.

According to TenneT the first serious congestion situations in the extra high-voltage grid may arise in the following regions (see Figure 1 and Figure 2):[6]

- Zuid-Holland: third quarter of 2010 until third quarter of 2012, and
- Eemshaven: first quarter of 2012 until the beginning of 2017.

From the figures, it can be observed that the amount of capacity shortage will vary over time. On the one hand the capacity shortage tends to increase due the commissioning of new generating capacity. On the other hand it will decrease as a consequence of appropriate capacity additions in the transmission network. Delay in commissioning of new plant will relieve

3 Presentation of TenneT at a discussion meeting on congestion management, Arnhem, 23 March 2009.

4 TenneT, *Rapport Monitoring Leveringszekerheid 2007-2023*, Arnhem, June 2008.

5 See for example the questions of the Raad van State about the provision of environmental permits to new coal-fired power plants. A list of questions has been sent to the European Court of Justice in Luxembourg asking whether the new plant contributes to the National Emission Ceilings (NECs) for sulphur dioxide (SO_2) and nitrogen compounds (NO_x) or not. As a consequence a commissioning delay of 1,5 to 2 years is expected.

6 Presentation of TenneT at a discussion meeting on congestion management, Arnhem, 23 March 2009.

congestion. Delay in the commissioning of network reinforcements will increase the estimated congestion.

The amount of congestion costs involved in congestion management (assuming system redispatch) has been estimated by TenneT. Based on the current estimates, the average amount of congestion costs may lie in the order of 100 million euros per year.[7] Obviously, this estimate depends on the amount of capacity shortage (which varies over time) as well as the units dispatched and the assumed cost for redispatched generation.

TenneT is keen to stress that these numbers are only rough estimates and subject to significant uncertainty. A more detailed assessment would be required to make a more accurate estimate of possible congestion costs. However, TenneT's estimates are consistent with the costs experienced in other European countries that are experiencing network congestion.[8] International experience also illustrates that, for various reasons, estimates of congestion costs are subject to large errors.

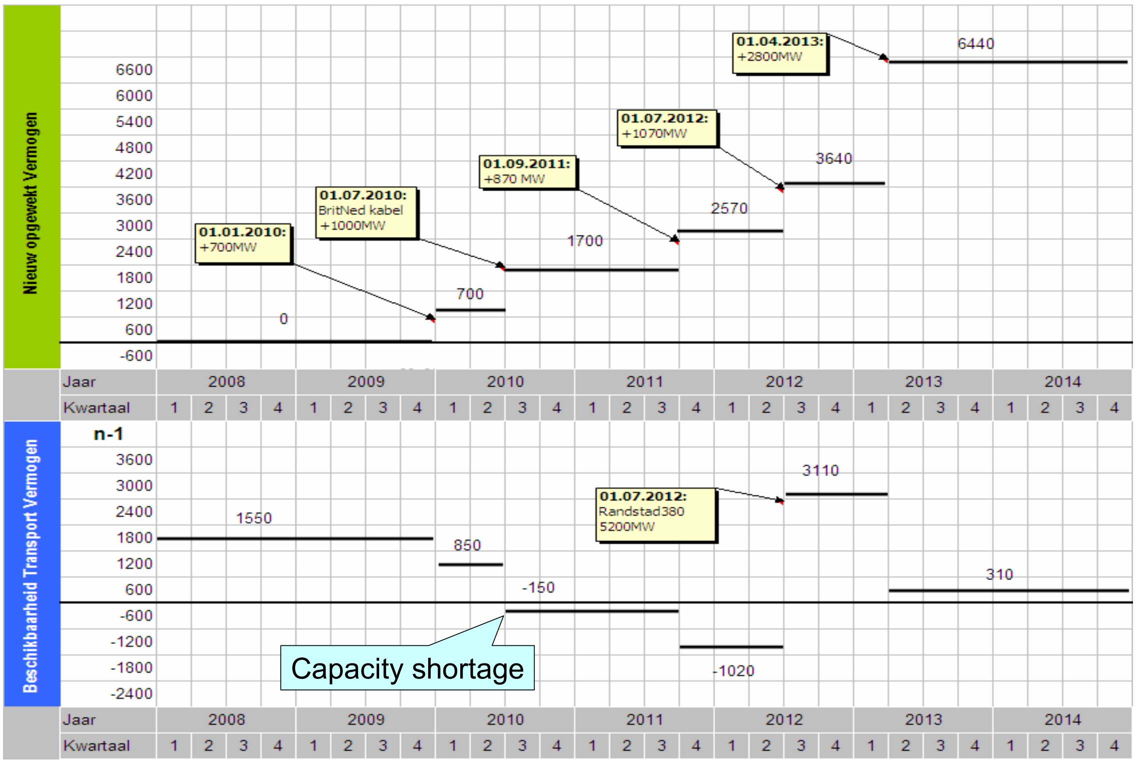


Figure 1. Expected timing of congestion problems in the province Zuid-Holland.

---

7 Presentation of TenneT at a discussion meeting on congestion management, Arnhem, 23 March 2009.

8 See Appendices I and II.

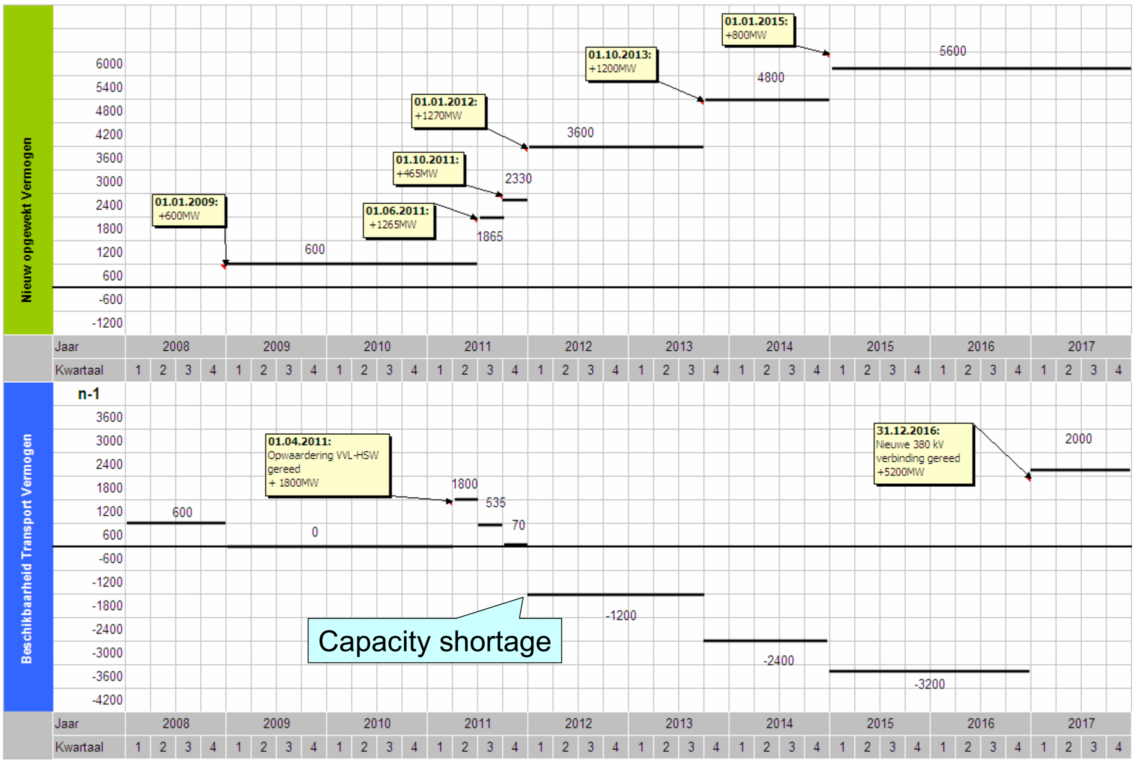


Figure 2. Expected timing of congestion problems in the region northern part of the Netherlands.

2.2 Economics of congestion management

2.2.1 Introduction

Congestion occurs when the volume of electricity that generators want to transport exceeds the available transmission capacity. 'Congestion management' refers to approaches that can be adopted to reduce the network flow over the congested tie lines (which is necessary to prevent the system to become unstable as a result of overloading and cascade tripping of protection mechanisms). In literature, many mechanisms and concepts have been described to solve congestion issues in electric transmission systems, and many different designs have been tested or implemented worldwide.

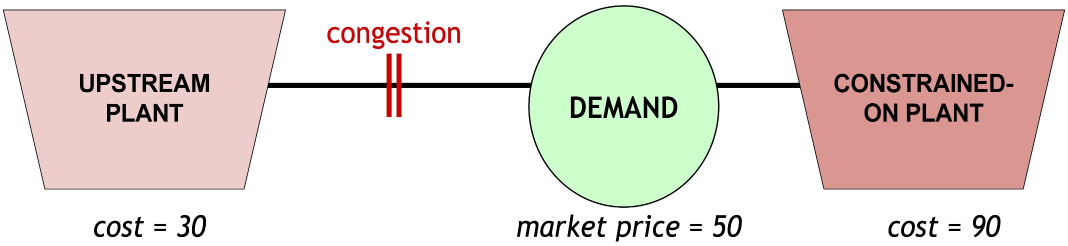
In practice, the physical results of all congestion management methods are the same: generation in the congested area is being decreased and generation outside the congested area is simultaneously increased with the same amount in order to keep the power balance in the system in equilibrium. However, the mechanisms to realise the increase and decrease of generation as well as to allocate the costs may differ significantly.

Typically, the generators will inform the system operator of their production plans the day before despatch. The system operator will then check if there is sufficient transmission capacity available for the planned production. If there is not, then generators 'upstream' of the transmission constraint will need to reduce their production, and generators downstream of the constraint will increase their production in compensation. In this section we describe the economics involved in managing congestion.

2.2.2 Redispatch and counter-trading

The most commonly applied solution to congestion is redispatch and counter-trading – which for the remainder of this report we will refer to as simply redispatch. Figure 3 illustrates a simple example. The upstream plant will be asked to reduce its output or 'constrained off'. The downstream plant will be asked to increase its output – 'constrained on' – so that the total amount of electricity is unchanged.

Figure 3. Simple example of congestion management.



Normally the additional constrained on production is credited to the upstream plant so that it maintains its original (pre-constraint) electricity production. In other words, if the upstream plant originally planned to produce 500 MW during a certain hour, it will still be able to sell 500 MW on the market, even if it has actually been constrained off and is only producing 300 MW. If the electricity was not credited to the constrained off party, the system operator or the constrained on generator would need to sell the power. Crediting the electricity to the constrained off generator keeps things contractually simple, and avoids the need for another transaction in the market.

Accordingly, in a competitive redispatch market generators will pay the system operator to be constrained off. Being constrained off enables the generator to sell e.g. 500 MW of electricity while only incurring production costs for e.g. 300 MW of electricity. Generators should be willing to pay the system operator their avoided variable costs to be constrained off, since doing so will leave them financially indifferent between producing or being constrained off. For example, in Figure 3, before accounting for constraints the upstream generator will make a short-run profit of 20 €/MWh, being the difference between the market price of 50 €/MWh and the generator's costs of 30 €/MWh. If the generator is constrained off, it saves the 30 €/MWh. Offering up to 30 €/MWh to be constrained off leaves the generator with a profit of 50 €/MWh – as before. If the generator had to pay more than this it would rather produce than be constrained off.

The costs of the constrained on producer should, by definition, be higher than the prevailing market price. If the costs of the constrained on producer was lower than the market price, the generator should already be running. In the example below, the upstream plant would offer the system operator 30 €/MWh to be constrained off, and the downstream plant would ask for 90 €/MWh to be constrained on. The net cost of resolving the constraint is the difference between the money paid by the constrained off generator and the money paid to the constrained on producer – in this example the cost is 60 €/MWh. As we discuss later, this cost can be recovered either from consumers, generators or a combination of both.

Typically the system operator organises a central market where generators submit bids to be constrained off and offers to be constrained on. The system operator then calls on these bids and offers using the most cost effective first.

2.2.3 Efficiency

Efficiency can be defined in both the short-term and the long-term. *Short-term efficiency* refers to the situation in which all the investments in generation and transmission are fixed. In the short-term, an efficient congestion management system should ensure that electricity is produced at the lowest cost possible. This is achieved by constraining off the most expensive plant and constraining on the cheapest plant first. An efficient redispatch market should achieve this – the upstream generators with the highest costs will offer the most to the system operator be constrained off, since they save the most money by reducing their generation while still being credited with electricity production. The cheapest plant downstream of the constraint should make the lowest offers to be constrained on.

Long-term efficiency refers to minimizing the total cost of transmission and generation, including all of the resources that generators use such as land and gas pipelines. To obtain long-term efficiency, generators should be charged for the cost of the resources they use. In this way, generators should make decisions that minimize overall costs. In reality, achieving long-term efficiency is complex. For example, generators making an investment in a plant with a 30 year lifetime should make decisions based on the total costs of the resources over the lifetime of the investment, not just at the time the decision is made.[9] Hence, predicting which policies will lead to long-term efficiency is difficult. However, in general policies that target costs to those that cause them should lead to greater long-term efficiency than policies where those that create costs do not bear them.

2.2.4 Price effect and costs of congestion management

When considering congestion management policy it is important to understand the link between the costs of redispatch and the price for electricity in a market with a single price zone. Figure 2 illustrates a simplified supply curve for the Netherlands and the demand line, with the price equal to P1. The cost of redispatch is the volume of electricity constrained off (which equals the volume of electricity constrained on), multiplied by the difference between the cost of the constrained off plant and the cost or price of the constrained on plant. The red rectangle, area A, represents this cost.

Figure 4. Illustration of the cost of redispatch.

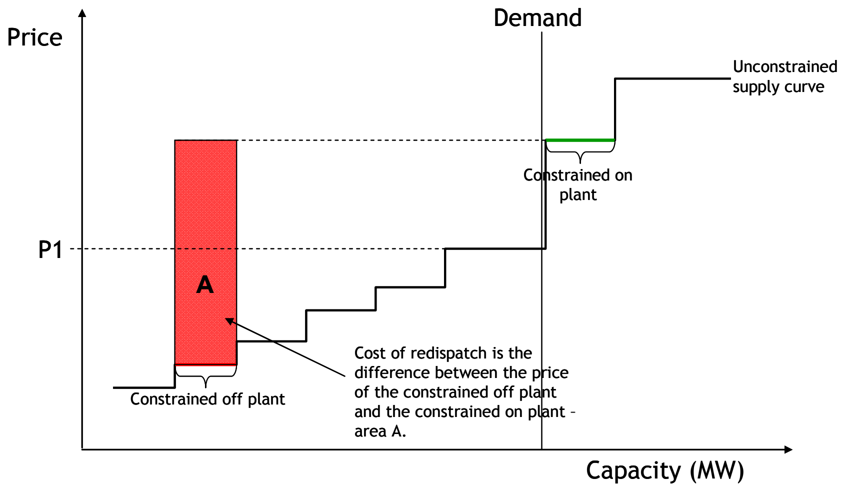
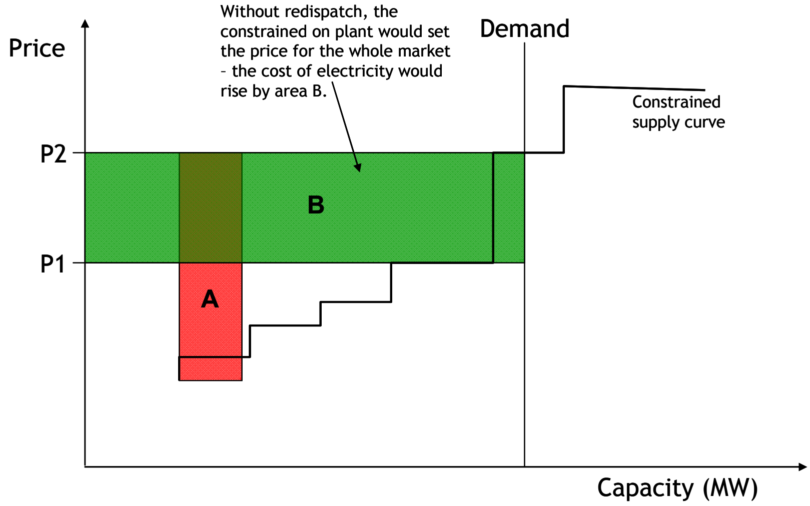


Figure 5 illustrates that redispatch reduces the market price of electricity. The alternative to redispatch would be to simply stop upstream generators from making offers in excess of the available transmission capacity. The withdrawal of these offers would shift the supply curve to the left, and increase prices to level P2. The result is that the plant that would have been constrained on ends up setting the price for the whole market. The increase in costs for consumers is the increase in price (P2 less P1) multiplied by demand – the green rectangle, area B, represents the increase in consumers' costs.

[9] For example, one could change transmission prices to represent the 'efficient' level, but making the change could affect generators' expectations of future prices and undermine confidence in the regulatory regime, thereby reducing efficiency.

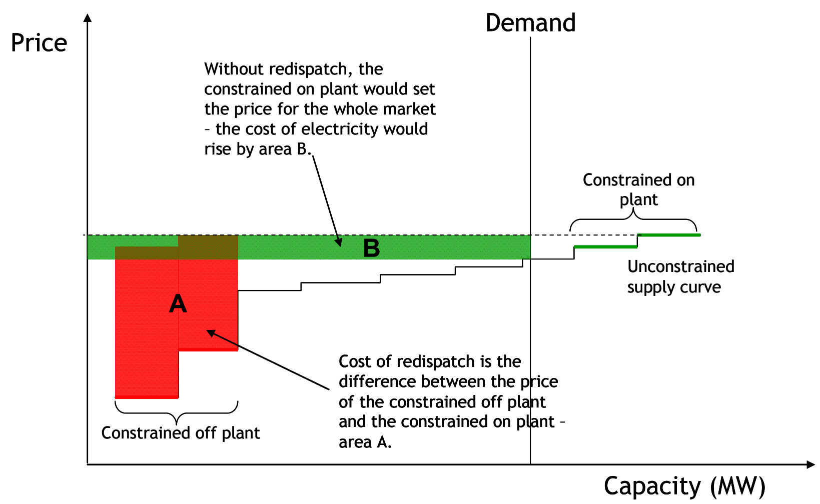
In the example in Figure 5, the increase in consumers' costs (area B) is actually larger than the cost of congestion (area A). In other words, in this example consumers would rather pay for the costs of redispatch and avoid the price increase associated with the constrained plant withdrawing its offer. This is result is due to the single price zone – whereas congestion costs depend only on the constrained volumes, if the price-setting plant changes this will affect costs for the entire market volume.

Figure 5. Situation in which the price effect is larger than the constraint costs.



As constraints increase, the trade off between congestion costs and prices changes. Figure 6 illustrates a situation where the cost of congestion is now higher than the reduction in prices. In this situation, consumers would not want to pay for congestion costs. This is likely to happen once a large volume of new efficient plant has pushed older, higher cost plant off the system.

Figure 6. Situation in which the price effect is smaller than the constraint costs.



These examples illustrate two main points:

- Even if they do not run, generators in the constrained of zone lower prices by making offers into the market.
- There is an 'optimal' amount of congestion management, in the sense that the reduction in prices balances the costs of congestion management. This suggests that a congestion management mechanism should have the ability to avoid congestion costs increasing beyond this optimum level.

2.2.5 Market power and gaming in congestion markets

The examples above assume competitive congestion markets, in which generators compete with one another to be constrained off or on. In reality there are circumstances in which generators may not behave competitively. In its July 2007 report for TenneT on connection policy, The Brattle Group noted that constraint markets have historically been subject to market power abuse, and cited the concerns expressed by the Great Britain regulator regarding constrained on payments as early as 1992.[10] As we discuss in more detail section 2.3.3, since then there have been many more cases of market power abuse, either proven or alleged.

There are two main ways in which generators could exercise market power in the constraint market:

- Lack of competition in the constrained off market – there may only be only one or two generators upstream of the constraint, that can offer to be constrained off. This is especially true if RE generators are excluded from the congestion management mechanism, or are reluctant to be constrained off as they will lose valuable subsidies by doing so. In this case generators may not offer their avoided costs to be constrained off, but may even request money from the system operator.
- Generator serving a 'load pocket' – in some circumstances there may only be one generator that serves a load (perhaps a city) at the end of a transmission line. If the transmission line is smaller than the demand of the load, then some output from the generator will be required to serve the load. The generator can create a constraint on the transmission line by refusing to offer sufficient power in the day-ahead market. The system operator is then forced to pay the generator a high price to be constrained on and solve the artificial constraint. This is the most common mechanism for exercising market power in the congestion market that we are aware of in practice.

The design of a congestion management mechanism should account for the fact that congestion market may not be competitive, and include measures to avoid the problems described above.

2.3 Overview of international experience

2.3.1 Introduction

Before developing alternative congestion management mechanisms for the Netherlands, we have reviewed international experience with congestion management in other large EU electricity markets. The main purpose of this review is to see what lessons can be learned for application of a congestion management policy in the Netherlands, bearing in mind specific Dutch policy objectives. We have looked at the electricity markets of Great Britain, Italy, Germany Spain and France.

All markets studied use redispatch or counter-trading to manage congestion, with costs either divided between generators and consumers or allocated only to consumers. However, this does not mean the Netherlands should adopt a policy of system redispatch with partial cost socialisation by default. Most countries designed their congestion management policy at a time when transmission capacity was relatively plentiful. Pre-market liberalisation, vertical integration between supply transmission simplified planning, and a lack of transparency and cost focus meant that many countries had no shortage of transmission capacity. Since market liberalisation, in many countries the balance between transmission capacity and generation has become tighter, resulting in more constraints. It could be that, were some countries to re-

---

[10] The Brattle Group, *A review of TenneT's connection policy*, July 2007.

examine their congestion management policy in the light of current conditions, they would apply a different system. For example Great Britain has experienced rapidly increasing congestion costs, and in March 2009 asked the system operator (National Grid) to take steps to reduce congestion costs, by limiting prices paid for congestion and perhaps targeting costs at the generators that cause congestion. At the time of writing, Italy is also making reforms to its congestion market, since costs there are also very high. Accordingly, the message is not that the Netherlands should simply copy the congestion management systems of other countries, but rather learn from the lessons of others whose congestion issues are more advanced.

Connection policies vary between countries and we would expect this to have a significant effect on congestion costs. For example in France generators must wait for network reinforcements to be ready before connecting, and this lowers congestion costs. A ‘wait to connect’ policy was also the case in Great Britain, although decisions taken during the creation of a single electricity market in 2005 has in effect led to generators connecting before there is sufficient transmission capacity available, and congestion costs have increased. We summarise the main lessons for Dutch congestion management policy below. We give more details of international congestion management experience in Appendices I, II and III.

2.3.2 Costs

Congestion costs can be significant. Even after scaling for the relatively small size of the Dutch electricity market, congestion costs of around €100 million a year would be typical for Great Britain and Spain. In Italy, even after adjusting for the larger size of the market, congestion costs are nearly €500 million in Dutch terms. These international numbers indicate that TenneT’s initial estimates of Dutch congestion costs of around €100 million per year, or roughly €15 per Dutch household per year, are not entirely unrealistic.

Congestion costs are also difficult to predict, and can increase rapidly. Between 2005 and 2007, actual constraint costs in Great Britain were more than twice the forecast level. Some of the underestimates of constraint costs arose due to the new circumstances of the combined market, but constraint costs are inherently difficult to predict since they are the difference between two things – the avoided costs of the constrained off plant, and the costs of the constrained on plant. A relatively small change in the cost of e.g. constrained on plant can lead to a large change in constraint costs. Forecast costs have increased from £77 million to £262 million from one year to the next.[11]

The international experience of relatively high and volatile constraint costs points to a congestion management mechanism that can control congestion costs, or at least the costs born by consumers.

2.3.3 Market power

Market power has been a consistent problem in constraint markets, and ex post competition law can be difficult to apply. Ofgem recently abandoned an investigation into alleged market power abuse in Scotland as it felt it as unlikely to be able to provide a sufficient level of proof. Nevertheless, Ofgem has taken steps to reform the way that generators are paid for congestion to try and avoid such problems in future. Spain has had several incidents of market power in constraint markets, but seems to have been more successful in prosecuting those involved – the most recent case was in April 2008. Italy has very high constraint costs, and the Italian regulator and competition authority have found the incumbent generator dominant in several of Italy’s

[11] Managing constraints on the GB Transmission System, 17 February 2009, letter from Ofgem to National Grid.

markets including the market for imbalance, reserve and redispatch.[12] It therefore seems likely that Italy's high congestion costs could be related to market power issues.

Several countries have developed ways to try and mitigate market power in congestion services. In Italy the system operator can sign 'must run' contracts with designated essential plants and force them to run at the market price. In Germany, generators costs are audited to check that they are bidding close to competitive levels.

We have not studied the configuration of the Dutch transmission network and the location of generators to see how likely it is that market power problems will emerge in the Dutch congestion market. Both TenneT and the Energiekamer have expressed concerns about the competitiveness of the balancing market in general. The market for relieving local transmission constraints, especially within a congested area, would be even less competitive. Therefore, it seems sensible to design a congestion management mechanism which anticipates market power problems and deals with them ex ante, rather than relying on ex post competition law which is cumbersome to enforce in such circumstances.

2.3.4 Information requirements

Several countries require generators to submit a production plan on a plant by plant basis, though the plans are not always binding or the basis for imbalance charges. For example, in Great Britain generators have to tell the system operator the planned production for each plant, but deviations from the plan does not attract imbalances charges. However, generators have a license obligation to submit 'accurate' programmes, and so consistent deviation from the programme would be a breach of the license and could result in a large fine. This system seems to work well in the UK, and there have been no reports of gaming by consistent deviation from nominated plant programmes.

[12] AG congestion management, AEEG, Indagine conoscitiva sullo stato della liberalizzazione dei settori dell'energia elettrica e del gas naturale (IC22).

3 Criteria for evaluating congestion management mechanisms

3.1 Categories of assessment criteria

The congestion management mechanism to be implemented needs to comply with several criteria. The criteria include:

- *Technical criteria:* Technical criteria concern the network requirements which need to be met by a congestion management system in order to be applicable in the Dutch electricity system.
- *Economic criteria:* Economic criteria refer mainly to the incentives provided by a congestion management system and to the possibility to achieve economic efficiency.
- *Policy criteria:* The policy criteria involve the compliance of the congestion management models with the existing policy on e.g. network connection and network use as well as with the current debate in Parliament on the promotion of renewable energy.
- *Regulatory criteria:* Regulatory criteria mainly concern the consequences of a congestion management model with respect to the proper development of the electricity market and the electricity network.

The criteria will be detailed in the following sections, where also choices are made for the manner in which these criteria will be applied when assessing the possible congestion management mechanisms. The assessment itself will be presented in chapter 5.

3.2 Technical criteria

The following technical criteria are included in the assessment of the congestion management options:

3.2.1 Applicability

The first criteria to be applied concern the applicability of the congestion management model in the Dutch electricity system. The congestion management system must be broadly applicable in the Dutch high-voltage electricity network[13], irrespective in which geographic region congestion occurs (e.g. Eemshaven, Maasvlakte or some other electrically congested area).

Moreover, the congestion management system must be able to operate irrespective of the amount of congestion. The focus here is on the technical applicability of the system (which may involve information requirements or the availability of sufficient dispatchable generators), not on the economic constraints to the system (e.g. too high costs of the system).

3.2.2 Effectivity

The effectivity of a congestion management mechanism in essence says that the mechanism performs as intended: The system should be able to sufficiently manage all expected congestion, at least on the high-voltage grid.

3.2.3 Network safety

Since a congestion management mechanism aims to shift network flows (in order to relieve congestion), the mechanism should be 'network safe'. This means that the operational network safety criteria need to be complied with after application of the mechanism. The reverse should be true as well: Application of the congestion management mechanism should not lead to a

---

[13] The high-voltage grid consists of all lines with a voltage level higher than 110 kV. The applicability of congestion management for networks of a lower voltage level falls outside the scope of this study.

reduction of available transport capacity (which might be the consequence of some congestion management methods in order to guarantee network safety).

3.2.4 Incentives for information supply

A final technical criterion relates to the incentives for market parties to supply accurate information on planned network use. Although ‘gaming’ will be separately discussed being an economic criterion, the (associated) provision of adequate technical information to the network company is also of importance.

Some congestion management mechanisms may incentivize market parties to overestimate transport needs (e.g. a *pro rata* auctioning approach to network capacity). Likewise, some mechanisms may provide incentives to underestimate network use (e.g. when not-nominated transports are excluded from the congestion management mechanism and not financially charged). The congestion management system should be such that there is an intrinsic driver for market parties to submit the best data on network use.

3.3 Economic criteria

The following economic criteria are included in the assessment of the congestion management options:

3.3.1 Economic efficiency

Economic efficiency, of cost efficiency, concerns the possibility of the congestion management method to achieve ‘allocative efficiency’. Allocative efficiency reflects the extent to which firms use inputs (for example capital and labour) in optimal proportions for a given set of input prices and a given technology.

Allocative efficiency is measured by the total social surplus, which is the sum of producer and consumer surplus. Producer surplus is given simply by profits i.e. the difference between revenues (the price at which the Electricity is sold) and production costs. Consumer surplus is provided by the difference between consumer willingness to pay and the going market price. As long as the willingness to pay by consumers is higher than the marginal costs of producing the unit, an increase in output increases total surplus and consequently allocative efficiency. Maximum allocative (or Pareto) efficiency (given the existing demand and supply conditions) is reached when costs and willingness to pay are equal at the margin. In the Pareto optimal state there does not exist another feasible state where one party is better off (on a higher utility level) and no other party is made worse off. Then, an optimal allocation of scarce resources is reached and society as a whole will be best off.

A congestion management mechanism needs to contribute to an economic efficient dispatch of generation. This means that, after congestion has been taken into account and plant has been redispatched, market equilibrium is achieved at the lowest total costs. This means that the cheapest plant available should run. This criteria is an essential one, otherwise the congestion management scheme would distort the market, with winners and losers.

3.3.2 Robustness (gaming possibilities)

Given that many congestion regions suffer from market power, the congestion management model should be robust for gaming. This implies that the outcome of the congestion management system cannot be influenced by strategic behaviour of individual market parties. Given the experience in other countries, robustness against gaming is an important criterion for assessing the different congestion management models.

The difference between these criteria and the 'incentives for information supply' (under the technical criteria) lies in the economic benefits. Under some schemes, especially system-based redispatch models, market parties might quote to low cost for constrained-off power (i.e. the avoided cost of generation which they are to pay to the system operator) or to quote too high prices for constrained-on power (i.e. the cost for compensatory generation outside the congestion area, which must be contracted to restore the balance between supply and demand).

The reverse can be true as well, especially in market-based congestion management schemes, where generators might place bids which deviate from their real costs in order to reduce the total congestion costs for all their plant.

3.3.3 Economic incentives

A third economic criteria concerns the (short-term) economic incentives applied in a congestion management mechanism.[14] The congestion management system should be able to give the correct economic 'signals' to network users. Such signals indicate where congestion occurs. They should favour transactions that reduce or bypass the congestion in contrast to transactions increasing the congestion.

The economic incentives provided ideally match the 'cost causation' principle. This principle says that the costs incurred by congestion are to be paid by the parties causing the congestion.

3.3.4 Transaction costs

Transaction costs involve the cost of market parties for participating in the congestion management scheme. These e.g. include the *ex ante* costs of placing bids with the system operator and establishing the contract (in case they are constrained off or on) and the *ex post* cost for carrying out the transaction, including financial settlement.

Often, transaction costs are not very transparent, since these are internal to the company. However, different congestion management schemes may involve different transaction costs, mainly depending on the complexity of the congestion market. Of importance are also the costs of information requirements and the costs of IT systems to be able to participate in the congestion management system.

3.4 Policy criteria

The following policy criteria are included in the assessment of the congestion management options:

3.4.1 Proportionality

The criterion of proportionality represents a balance between the 'size' of the congestion to be solved and the complexity of the congestion management system. In other words: the 'impact' of the congestion management system should match the seriousness of the congestion. If congestion is rather limited, congestion management by a system operator may be preferred (although this involves net costs for the system operator), whereas more significant and structural congestion might best be addressed by changes in the market design.

Since for highly congested systems, any congestion management system is probably justified and therefore proportional (almost by definition) the criterion will be applied pertaining to slightly

[14] Long-term economic signals are covered in §3.5.3.

congested systems. The question is then whether the efforts needed to implement different congestion management models are balanced by their benefits.

3.4.2 Priority access for renewable energy

A major constraint for any congestion management model is that it should allow priority transports for electricity generated from renewable sources. This priority can be understood in several ways:

1. Renewable plant is entirely excluded from the congestion management system, so it is not allowed to be constrained off and will not pay any congestion-related charges.
2. Renewable plant is excluded from constraining off. However, given that the effective market price in the congestion zone may be affected, renewable generators are *not* exempt from paying the associated congestion costs (or fee) just as non-constrained off conventional generators.
3. Renewable plant may participate in the constrained off congestion management system, although it is not obliged to do so.

An assessment of the (dis)advantages and effects of these options is outside the scope of the present report (see section 6.1). However, the congestion management model should allow implementation of each of the three options.

3.4.3 Accuracy of cost assessments

For a policy decision on congestion management and given the international experience, it is important to have some idea of the total costs of the congestion management system. Especially if these costs are to be paid by consumers, an accurate assessment beforehand is relevant. After all, it is unwelcome if the costs involved in congestion management significantly exceed preliminary estimates.

Although the congestion revenues are difficult to assess beforehand as well, this is generally not considered a problem. Therefore, this criterion will be applied primarily on the assessment of the (net) costs resulting from a congestion management model.

3.4.4 Speed of implementation

Finally, for a congestion management system to be of interest, it needs to be quickly implementable. Some systems may need significant changes in the market design, which makes speedy implementation questionable. Furthermore, since congestion management involves new information exchanges, the amount of investments in IT and information systems, changes in network procedures and market rules needs to be assessed. Any legal hindrances for implementation have not been included under this criterion, since it is supposed that these will be solved by adapting the Electricity Act.

Given the expected amount of network congestion (and since congestion management is considered a temporary measure to be replaced by structural solutions as network investments), the congestion management method needs to be implementable in the course of 2010.

3.5 Regulatory criteria

The final set of criteria involves regulatory criteria:

3.5.1 Non-discrimination

The congestion management system should not discriminate between network users, of course unless required by legislation. The latter may occur when renewable generators are to be excluded from participation. Otherwise, all generators in the congestion area should be treated in the same way. Specifically and in accordance with the assumptions underlying this study (see §1.4), no distinction should be made between existing generators and new plant.

3.5.2 Simplicity and transparency

An important criterion is the simplicity and transparency of the congestion management system. Not only do market parties need to have a good understanding of how the system works (and how costs and prices are inferred from it), but also the effect of the congestion management system on the market itself needs to be transparent.

3.5.3 Incentives for investment

Short-term economic incentives have already been discussed under the economic criteria. However, also long-term incentives must be included in the assessment. These mainly relate to investment incentives for network companies and generators.

- *Incentives for network companies:* Network companies need to have proper incentives for structurally solving network congestion by reinforcing congested links. When some of the congestion costs are allocated network companies, they will be able balance operational measures (like congestion management) with structural measures (network reinforcements).
- *Incentives for generators:* Generators need to get incentives in order to site new plant near demand (and not at locations which will increase network congestion). The same holds for large industrial consumers.

Given that incentives to network companies can be given in several ways (particularly through the regulatory mechanism) and since much of the network congestion results from the difference in building time (including permitting) between generation and transmission lines (the latter being slower), in the present assessment the investment incentives are interpreted to apply to generators.[15] A positive score with respect to this criterion must be considered to represent a situation in which new generation investments in the congested zone will be deterred up to the economic optimum.

3.5.4 Incentives for decommissioning

A related incentive concerns the effect of the congestion management system on decommissioning of obsolete plant. When generators earn by solving congestion, there may be an incentive to keep old plant in a congestion area in operation instead of decommissioning. A congestion management mechanism needs to discourage postponing decommissioning of inefficient plant.

3.6 Prioritizing the criteria

The congestion management models will be assessed according to the criteria given above. Given the criteria, much of the assessment will be of a qualitative nature. However, for some of the economic criteria, an assessment based on a model situation will be provided.

[15] For a more extensive defence of this position, see section 2.2.3.

With respect to the importance of the criteria, it is suggested to distinguish between ‘essential’ criteria, ‘major’ criteria and ‘minor’ criteria:[16]

- *Essential criteria:* The congestion management mechanism needs to comply with the essential criteria in order to be eligible for implementation.
- *Major criteria:* A congestion management model needs to comply with most, although not all, of the major criteria. In case of non-compliance, suggestions to mitigate the negative effects of the model need to be assessed.
- *Minor criteria:* A congestion management preferentially complies with the minor criteria as well, although they are no ‘show-stoppers’.

The ranking of the criteria as considered in this report is presented in Table 1.

Table 1. Ranking of the evaluation criteria for congestion management models according to their relevance.

		ESSENTIAL	MAJOR	MINOR
Technical criteria	*Applicability*		✓	
	Effectivity		✓	
	Network safety	✓		
	Incentives for information supply			✓
Economic criteria	*Economic efficiency*	✓		
	Robustness (gaming possibilities)		✓	
	Economic incentives		✓	
	Transaction costs			✓
Policy criteria	*Proportionality*		✓	
	Priority access for renewable energy	✓		
	Accuracy of cost assessments		✓	
	Speed of implementation	✓		
Regulatory criteria	*Non-discrimination*		✓	
	Simplicity and transparency			✓
	Incentives for investment		✓	
	Incentives for decommissioning			✓

[16] This distinction is partly based on the responses from stakeholders.

4 Alternative congestion management mechanisms

4.1 Introduction

In this chapter we provide a systematic description of several alternative models of congestion management. We offer a description of the models themselves, provide a numerical example illustrating the model and discuss the main economic advantages and disadvantages of each model. We will provide a more general evaluation of each model against the selected criteria in chapter 5.

The following six models will be discussed:

- Market agent approach;
- Basic system redispatch model;
- System redispatch model with cost pass-through to generators;
- Market redispatch model;
- Hybrid redispatch model, and
- Market splitting approach.

In order to facilitate the reader, a basic characterization of these models is presented in Table 2.

Table 2. Basic characterization of the models discussed.

	Market agent	System redispatch	Hybrid redispatch	Market redispatch	Market splitting
Basic approach	*ex post*	*ex post*	*ex post*	*ex post*	*ex ante*
Constraining-off	*Pro rata*	System	Market	Market	Power exchange
Constraining-on	Market	System	System	Market	Power exchange
Number or price zones	one	one	one	one	two
Costs/revenues for system operator	-	costs	probably revenues	revenues	revenues

Reference model

For illustration of the different congestion management mechanisms, we present a simplified two-zone system with three generators in each zone (see Figure 7). Depending on the generation costs in A and B, there may be a net flow from A to B which is larger than the available transport capacity from A to B, so that congestion occurs.

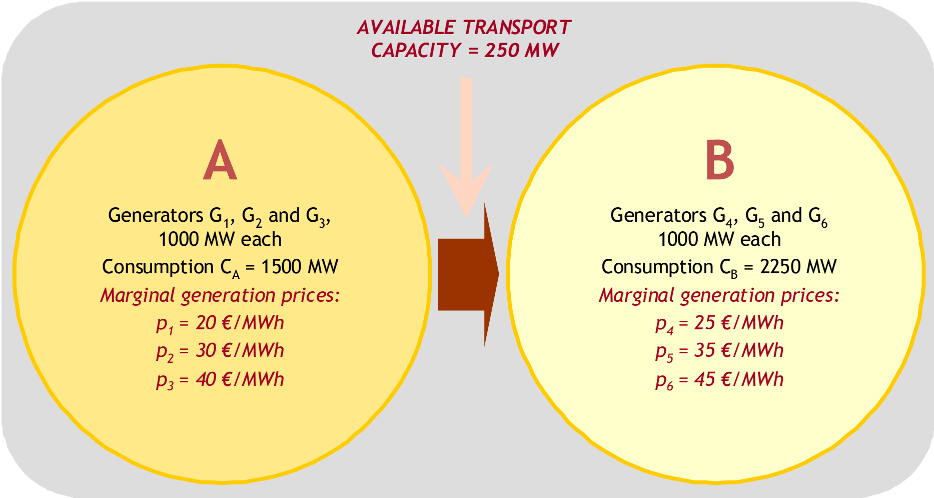


Figure 7. Simplified model to illustrate the effect of congestion. Since average generation costs in A are lower than in B, region A exports to region B. The available transport capacity from A→B, 250 MW in the example, may be insufficient to accommodate all transports required by the market.

Assuming 250 MW transport capacity between A and B, there is more demand for cheaper generation in A (from consumers in B) than the available transport capacity may accommodate. Therefore congestion occurs.

The most economically efficient way to clear the market is to maximize the flow from A to B when clearing the market. This will result in the following dispatch:

REGION A		250 MW transport capacity available	REGION B	
Effective demand	1750 MW	*effective transport:* 250 MW	Effective demand	2000 MW
Marginal generator	G2		Marginal generator	G5
Market price	30 €/MWh		Market price	35 €/MWh
Generation costs in A	€ 42.500	Total for A and B: € 102.500	Generation costs in B	€ 60.000

In this table, the 'effective demand' represents the net demand which must be met by generators, either with or without transport between A and B. In this case, since there is no transport capacity available, effective demand equals consumption in each region.

The 'marginal generator' represents the most expensive generator which needs to run to supply the market. This generator sets the market price. In this example we assume a competitive market, so that the price is equal to the cost of the marginal generator). The 'generation costs' represents the total (hourly) costs of generation in the two regions.

As can be observed, the total generation costs are € 102.500. If there would be more transport capacity available from A to B, the total generation costs can be reduced.[17] However, given the limited transport capacity, there is no dispatch possible which leads to lower total costs: Additional generation in zone A cannot be transferred to zone B (given the congestion), whereas a different dispatch in zone B (i.e. G6 instead of G4 or G5) would increase total costs. Therefore, the dispatch suggested is short-term efficient.

[17] In Appendix IV we illustrate the costs and prices for cases with unconstrained transmission capacity and no transmission capacity between the zones, to illustrate the effect of transmission capacity on generation costs.

The model presented above will be used to illustrate the effect of different approaches towards congestion management.

4.2 Market agent approach

4.2.1 Description of the model

The market agent approach to congestion considers congestion as an external limitation to trade originating from the network. Instead of arranging for an *ex ante* methodology for prioritizing transports or *ex post* mitigation of the negative effects of congestion, this approach transfers the responsibility to deal with congestion entirely to market parties.

The most straightforward manner to implement this system (although alternative systems can be conceived as well) is to reduce the transport rights of all generators in A *pro rata* until the congestion is solved. Since the market transactions are not changed, the generators need to compensate for the reduced generation by themselves, either by reducing their sales in B or by purchasing power in B. No financial compensation is given to generators who have to reduce their output. Therefore all costs (cost for compensating measures in B minus avoided costs of running constrained plant in A) rest with the generators in A.

Illustration by means of the reference model

After noticing that the projected transports from region A to region B exceed the available transmission capacity of 250 MW, all generators in region A are called to pro rata reduce their output. Effectively, the output of both G1 and G2 is being reduced from 1000 MW to 875 MW each. Since this does not represent the least-cost dispatch (G1 runs cheaper than G2), the generators could trade bilaterally to re-adjust their output (e.g. increasing G1 to 1000 MW and further reducing G2 to 750 MW). However, this would involve additional transaction costs, and involve bargaining issues in how to allocate the savings. We do not illustrate the case of bilateral trading in this example.

In region B, there is a residual power deficit of 250 MW (equalling the constraining of G1 and G2). Generators G1 and G2 (or their program responsible parties) need to purchase this additional 250 MW in B (or arrange for corresponding demand decrease), since there is now a mismatch between generation and consumption in B. This generation deficit must be corrected by the market agents themselves. The cheapest available generator in region B is G5 (ignoring possible demand reduction).

Schematically, the market agent approach can be represented as follows:

REGION A		BEFORE	REGION B	
Output G1	1000 MW	TRANSPORT CAPACITY: 250 MW	Output G4	1000 MW
Output G2	1000 MW		Output G5	750 MW
Output G3	0 MW	TRANSPORT A→B 500 MW	Output G6	0 MW
Demand in A	1500 MW		Demand in B	2250 MW
Total generation in A	2000 MW		Total generation in B	1750 MW

REGION A		AFTER	REGION B	
Output G1	875 MW	REDUCTION OF 250 MW APPLIED	Output G4	1000 MW
Output G2	875 MW		Output G5	750 MW
Output G3	0 MW	TRANSPORT A→B 250 MW	Output G6	0 MW
Demand in A	1500 MW		Demand in B	2250 MW
Total generation in A	1750 MW		Total generation in B	1750 MW
CONSTRAINED OFF	250 MW		MISMATCH	250 MW

4.2.2 Economic evaluation

Advantages

The main advantage of the market agent approach is that it avoids market power problems. Since generators are not compensated for constraints, there are no payments or prices to manipulate.

An additional potential advantage of this approach is that it allows the market to 'discover' the most efficient way to carry out the transactions to solve congestion – this could either be via a centralized market, or via bilateral transactions. Initially, there would be no centralized market for solving congestion. The burden of managing the congestion is allocated to the generators in region A, who may themselves set up market arrangements for minimizing their constrained-off costs. Plant with higher opportunity costs may prefer to stick to its original program and trade its obligation to reduce output to cheaper generators. Hence, if a centralized market is more efficient than bilateral transactions, and when sufficient independent generators exist in zone A, a market for constrained-off generation may develop.

However, the 'discovery' process could be time consuming and involve inefficiencies. The market agent approach may achieve economically efficient dispatch only with secondary trading of some form. If bilateral trading is slow to develop, the inefficiencies could be large.

Disadvantages

The main disadvantage of this model is that it is inefficient (at least in first instance) since cheap generators as well as relatively expensive generators are constrained off indiscriminately. In principle, this inefficiency could be overcome by a secondary market within the congested area, although this might involve high transaction costs. It is doubtful how liquid this market would be, and there could be market power in bilateral deals between generators. This may especially become a problem when a generator operates under must-run conditions and must sell its output reduction at any price.

Furthermore, an 'emotional' disadvantage may be involved in the mandatory curtailment of all generators, irrespective of where the consumers are located. If G2 only supplies contracted customers in region A, it is nevertheless obliged to reduce its output and participate in the scheme. Despite the supply to customers of G2 not being at stake (given the energy balance in region A), G2 is nevertheless held responsible for purchasing compensatory power in region B. This however could be overcome by use of 'regional e-programmes', where generators could identify where their customers were and their net export requirements from the region. Generators could then be constrained off based on their net export requirements from the region.[18]

4.3 Basic system redispatch model

4.3.1 Description of the model

System redispatch is the 'standard' redispatch model (as described in section 2.2.2). The system operator organises a market for both constrained-off services and constrained-on services. In region A, generators bid for which price they are willing to reduce their output. In region B, generators bid for which price they are willing to increase production.

Constrained-off generators in A are net compensated, since although they pay to be constrained off they are credited with electricity which they sell at the market price. Constrained on generators are also paid their costs. The system operator organises all redispatch, and acts as a central counter party for redispatch trading. The costs of congestion management would be allocated to both consumers and generators.

Illustration by means of the reference model

In a competitive market, generators in region A bid into the constrained-off market according to their marginal costs in the same way as described in section 2.2.2.

In region B, generators who have additional generation capacity available will bid at their marginal costs. The system operator, aiming to purchase the cheapest generation available, will contract 250 MW with G5 to compensate for the reduced output of G2.

In summary, the system operator, knowing the marginal costs (assessed by the bids of generators) in both region A and region B, reduces output of G2 from 1000 MW to 750 MW in order to relieve the congestion. Simultaneously, it purchases 250 MW of the cheapest available generation in B, which is 250 MW of G5.

The total (hourly) costs for system redispatch now add up to 250 MW × 35 €/MWh (representing the cost for constraining-on G5) minus 250 MW × 30 €/MWh (the avoided costs for constraining-off G2, i.e. the price paid by G2 to the system operator), which results in a net cost to the system operator of € 1250.

Schematically, the system redispatch model can be represented as follows:

[18] The argument given could be applied in other models as well, although generators may indicate their 'willingness' to participate in a congestion scheme through their bids. Generators could identify customers within the zone, and have a net export programme (the APX counting as outside the zone). They could then be charged on the basis of net exports, rather than generation. The advantage is that generators would compete for customers within region A, and reduce prices for customers there, giving a signal for customers to locate there. The disadvantage is that it necessitates splitting the E-program.

REGION A		BEFORE	REGION B	
Output G1	1000 MW	TRANSPORT CAPACITY: 250 MW	Output G4	1000 MW
Output G2	1000 MW		Output G5	750 MW
Output G3	0 MW	TRANSPORT A→B 500 MW	Output G6	0 MW
Demand in A	1500 MW		Demand in B	2250 MW
Total generation in A	2000 MW		Total generation in B	1750 MW

CONSTRAINED-OFF BIDS IN REGION A		BIDS IN CENTRALIZED MARKET	CONSTRAINED-ON BIDS IN REGION B	
G1	20 €/MWh		G4 (0 MW)	-
G2	30 €/MWh		G5 (250 MW)	35 €/MWh
G3	-		G6 (1000 MW)	45 €/MWh

REGION A		AFTER	REGION B	
Output G1	1000 MW	REDUCTION OF 250 MW APPLIED	Output G4	1000 MW
Output G2	750 MW		Output G5	1000 MW
Output G3	0 MW	TRANSPORT A→B 250 MW	Output G6	0 MW
Demand in A	1500 MW		Demand in B	2250 MW
Total generation in A	1750 MW		Total generation in B	2000 MW
CONSTRAINED OFF	250 MW		CONSTRAINED ON	250 MW

4.3.2 Economic evaluation

Advantages

A system redispatch model is short-term efficient, in that it ensures that electricity is generated in the cheapest way possible by constraining off the most expensive plant and constraining on the cheapest extra-marginal plant.

Transaction costs are relatively low, since they take part via a centralised pool. While generators may have to make bids and offers every hour, the format is standardised, and so once the system is established the transaction costs may be low. Credit and counter-party risk is minimal since the generators are dealing directly with the system operator.

Finally, system redispatch also has the advantage of being 'tried and tested', since it is the default scheme for managing constraints in power markets around the world.

Disadvantages

In a 'standard' redispatch model, generators are fully compensated for the cost of constraints, and so have no incentive to locate their plant outside of the constrained area, or reduce output during constrained hours. Combined with a policy of connecting all generators without waiting for network reinforcements, a system redispatch model could lead to very high congestion costs. If consumers bore all or most of the costs, then from the consumers' point of view congestion management costs could become excessive. The reduction in price from the offers of constrained off generators would be less than the congestion management costs.

A related disadvantage is that system redispatch does not provide incentives for new plant to move outside the congestion area. Due to potential market power effects (see below), generators in zone A might even be rewarded for contributing to congestion, which may attract new plant or postpone decommissioning, increasing the congestion. This could result in long-term inefficiency.

As mentioned, one of the main disadvantages of system redispatch is that it is vulnerable to gaming or market power (section 2.3.3 describes the way in which market power is exercised). Such market power will have the effect of offering too high prices for constrained-off and constrained-on power to the system operator (so that the system operator pays more than the marginal costs in region B for constrained-on power and received too little for constrained-off power in region A).

4.4 System redispatch model with cost pass-through to generators

4.4.1 Description of the model

The present model build on the system redispatch model as described above in section 4.3.1. The important difference is that all the costs of congestion management would be allocated to generators upstream of the constraint who are not constrained off. Excusing the generators that are actually constrained off from bearing any costs would be one way to motivate generators to participate in the constraint market and to offer reasonable constrained off costs.

Illustration by means of the reference model

The economics of the system redispatch model have already been illustrated in section 0.

The total congestion costs in the example amounted to € 1250. These costs are now allocated to the generators in region A, which are G1 and G2. Given their combined output of 1750 MW, they will be charged a fee of 0,715 €/MWh (which can be invoiced separately or through the network tariff). G1 will therefore pay € 715 to the system operator (for 1000 MW), whereas G2 will be charged € 535 (for 750 MW).

4.4.2 Economic evaluation

Advantages

This variant of system redispatch shares the advantages of the 'standard' model described above, with the exception that the allocation of costs is no longer flexible. The system has the additional advantage that congestion management costs for consumers cannot get out of control.

Disadvantages

The disadvantages resemble those of the system redispatch model, the main disadvantage being that the system does not eliminate incentives to exercise market power. Although the system reduces the incentive to exercise market power in the constrained off market, it does not eliminate them. In the standard system redispatch design, there is only an upside to paying the system operator less than actual costs to be constrained off, since the generator loses nothing if its offer is rejected. In this revised scheme, generators that are not constrained off have to bear the costs of congestion management, which may give them some incentive to bid a bit more competitively (although the costs related to the socialization of the redispatch costs do not outweigh the individual revenues of continued overbidding). Consequently, market power mitigation measures would still be required, as described above.

Generators in region A continue to receive a higher (net) price for their electricity than in some other CM schemes, and therefore have less incentive to locate outside of the congested zone. This could reduce long-term efficiency. In our example case the net price generators in region A receive is slightly more than 34 €/MWh (35 €/MWh minus 0,715 €/MWh), whereas the economic value of generation in region A amounts to 30 €/MWh.

A potential disadvantage is that there is no flexibility to allocate the costs of congestion management.

4.5 Market redispatch model

4.5.1 Description of the model

According to this model, the system operator reduces the output of generators in region A according to the bids of generators in a congestion rights market. The constrained-off generators in region A are obliged to contract compensatory power in region B themselves (possibly through their program responsible parties).

The effectiveness of the model is dependent on the design of the congestion rights market. One approach could be to have generators bid into a monthly auction for congestion rights.[19] Their bidding price should reflect the (hourly average) costs for redispatching plant themselves (i.e. reducing 1 MW of output in region A and simultaneously purchasing 1 MW of additional generation in region B).

The system operator places the bids on a bidding ladder, but no payment is yet involved. When congestion is expected to occur, the system operator calculates how much plant needs to be redispatched. Then it informs the generators representing the lowest bids in the congestion rights market (corresponding to the parties who incur the lowest cost to shift generation from region A to region B) to reduce their plant output in A and purchase redispatch capacity in B. The costs rest with the generators which are called off (i.e. each generator pays its own redispatch cost).

The bid of the most expensive generator which is asked to redispatch defines the marginal congestion price (for that hour[20]). All plant who have bid into the auction for a price higher than this marginal price are not asked to redispatch. These generators are charged the marginal congestion price times their output instead.

Illustration by means of the reference model

When placing the bids in the congestion auction, each generator in region A estimates its costs for redispatching plant. The bids represent the net costs for which a generator can transfer generation from region A to region B, which reflects the sum of the lost profits from reducing generation in A and the costs of contracting compensatory generation in B (for which the generator will try to get the best deals in the

[19] The choice for monthly auctions contains a trade-off since the system works for any auction period, even on an hourly basis. The real-time economic value of congestion can be more accurately assessed in an hourly auction. However, the transaction costs for generators may be high (although these can be reduced by standardizing the formats), given the necessity to prepare hourly bids. Weekly or monthly auctions require less bids (and therefore involve lower transaction costs), but now the bid must represent the average cost of redispatch for all the (expected) congestion hours in the respective week or month. Moreover, the system operator now needs additional information for assessing which units are in operation in a certain hour (requesting non-running plant to redispatch does not make sense), although this information must be available for accurately assessing the congestion in all models.

[20] Since the amount of congestion may vary over time, the congestion price may vary as well, even given monthly auctions.

market). In this simple example, the marginal price setting plant has excess capacity - and can be constrained on at the market price. Therefore the constrained on costs in this example are zero. More normally, there would be a constrained on cost, because the constrained on plant would be more expensive than the market price.

Given the market price of 35 €/MWh, G1 will bid in the auction at 15 €/MWh, which represents the lost profits if the plant is constrained off (35 €/MWh minus its marginal costs of 20 €/MWh) and G2 at 5 MWh. As discussed above, in this example there are no constrained on costs, though these would normally be added to the generators' bids. G3 does not place a bid in this example.[21]

The system operator now constrains off G2 by 250 MW, since G2 has announced the lowest costs for doing this. G2 will reduce its output from 1000 MW to 750 MW and contract G5 for an increase of 250 MW. The costs involved are 250 MW × 5 €/MWh = € 1250 and are paid by G2. Furthermore, for the present hour, the congestion price is settled at 5 €/MWh.

G1 and G3, who were not selected, as well as G2 for the remaining 750 MW must pay a congestion fee to the system operator of 1000 MW × 5 €/MWh = € 5000 for G1, 750 MW × 5 €/MWh = € 3750 for G2, and 0 MW × 5 €/MWh = € 0 for G3. Note that for G1 this payment is lower than having to redispatch itself, which would have resulted in a total cost of 250 MW (redispatched plant) × 15 €/MWh (redispatch cost of G1) + 750 MW (non-redispatched plant) × 5 €/MWh (congestion price) = € 7500. The total income for the system operator amounts to 1750 MW × 5 €/MWh = € 8750. Effectively, application of market redispatch lowers the income of generators in region A by 5 €/MWh.

Schematically, the market redispatch model can be represented as follows:

REGION A		BEFORE	REGION B	
Output G1	1000 MW	TRANSPORT CAPACITY: 250 MW	Output G4	1000 MW
Output G2	1000 MW		Output G5	750 MW
Output G3	0 MW	TRANSPORT A→B 500 MW	Output G6	0 MW
Demand in A	1500 MW		Demand in B	2250 MW
Total generation in A	2000 MW		Total generation in B	1750 MW

CONGESTION BIDS IN REGION A			REGION B	
G1 (35–20 €/MWh)	15 €/MWh		Market price = extra-marginal price in B (set by G5)	35 €/MWh
G2 (35–30 €/MWh)	5 €/MWh			
G3	-			

REGION A		AFTER	REGION B	
G1 (not selected)	1000 MW	REDUCTION OF 250 MW APPLIED	Output G4	1000 MW
G2 (selected)	750 MW		Output G5	1000 MW
Output G3	0 MW	TRANSPORT A→B 250 MW	Output G6	0 MW
Demand in A	1500 MW		Demand in B	2250 MW
Total generation in A	1750 MW		Total generation in B	2000 MW
CONSTRAINED OFF	250 MW		CONSTRAINED ON	250 MW

[21] Although the plant must bid in a monthly auction, since it will run at some time; nevertheless, its bid is meaningless for the present hour, since it does not run.

4.5.2 Economic evaluation

Advantages

A significant advantage of market redispatch compared to system redispatch is that since generators are not compensated for solving congestion, there are no market power problems in the congestion market. Congestion costs cannot get 'out of control' since there are no explicit congestion costs that can be allocated to consumers.

Since generators pay the marginal cost of solving congestion, they have an incentive to bid their actual costs of solving congestion. If they can solve congestion more cheaply than other plant then generators are better off declaring this, rather than paying a higher marginal cost. Generators whose redispatch costs are higher than the congestion price are better off by just paying the congestion price.

The market redispatch model allows solving congestion at the least cost, i.e. it will achieve short-term economic efficiency as well.

Additionally, since the congestion fee effectively reduces the market price in region A, the market redispatch model provides efficient incentives for both consumers and generators to locate new generation in places which relieve congestion (since the effective market price they will be able to collect will be higher). The need for transmission network reinforcements is therefore reduced. Effectively, the system more or less works as reverse market splitting, as far as generators in the congested zone are concerned. It results in net revenues (instead of net costs) for the system operator.

Finally, it is an advantage that smaller market parties who find it more difficult to obtain compensatory power in region B may bid higher in the congestion rights auction. Consequently, they will be called off later and will be able to benefit from redispatch by other plant.

Disadvantages

From a generator perspective, market redispatch is obviously less attractive than system redispatch. In system redispatch, the generator may get some revenues from being redispatched (depending on the difference between the redispatch price paid by the system operator and its marginal costs). In market redispatch, the best a generator can do is to minimize its costs (by either redispatching itself or paying the congestion fee).

Market redispatch changes the effective prices that generators in different parts of the Netherlands receive. This would create 'winners' (generators with a majority of plant downstream of constraints) and 'losers' (generators with a majority of plant behind a constraint). The associated price adjustment could create resistance to implementation from the 'losers', although the price difference makes economic sense.[22]

The fees paid by generators in the congested region may be more than required to ensure long-term efficiency, especially in the case of a steep marginal cost curve of generators in region B. For illustration, assume that the output of G5 equals 750 MW instead of 1000 MW. Now the congestion bids of G1 and G2 would have been 25 and 15 €/MWh respectively (since both would

[22] In fact, the present situation in which generators in a congestion area receive the market price belonging to an integrated, single-price market is a bit weird, since this market price does not reflect the technical-economic conditions of generation in the congestion area. Phrased differently, due to this anomaly consumers in a congestion zone pay too much for their electricity, since generation is cheaper than the price they have to pay for it. The latter effect is however not corrected by the market redispatch and hybrid redispatch models, unless some of the congestion revenues are redistributed among consumers in the congested zone.

need to contract in region B with G6 at 45 €/MWh instead of G5 which has no longer any additional power for sale). This would have resulted in a congestion price of 15 €/MWh (instead of 5 €/MWh) to be paid by all generators in region A.

The underlying reason for these higher redispatch costs is that the bids in the congestion market include both the constrained-off price (in region A) and the constrained-on costs (in region B). When compared with a market splitting model (see section 4.7), market redispatch will in the end increase prices in A (decreasing congestion), whereas market splitting will increase prices in B (thereby internalizing the cost of congestion in the market price).

Transaction costs under market redispatch will be relatively high, since generators have to arrange for their own constrained on power in region B bilaterally, though larger firms could probably arrange this from within their own portfolio of plant.

4.6 Hybrid redispatch model

4.6.1 Description of the model

The hybrid redispatch model can be considered as a mixture of the market redispatch model (for the congestion area) and the system redispatch model (for the constrained-on area).

Generators in the congestion area (region A) bid into a congestion auction market where the bids represent the lost profit for generators if they are constrained off. When congestion occurs, the system operator requests the generators representing the lowest bids first (i.e. the most expensive and least profitable plant) to reduce their output. The system operator still credits these generators with the constrained off electricity so that they are able to honour contracts and remain in balance, but charges them the APX-price for the power[23]. The highest bid accepted by the system operator for a specific hour sets the congestion price for that hour. For each MW of remaining generation delivered to the grid (which has not been reduced by the system operator), the congestion price must be paid (by generators in A).

The main difference with the market redispatch model is the manner in which compensatory power is being purchased in region B. This responsibility now no longer rests with the generators in A (as in the market redispatch model) but with the system operator, who runs a centralized market for constrained-on power. Therefore, generators in A only need to assess their costs for downward regulation (and take these into account when bidding in the auction), since the system operator takes care of upward regulation in B (outside the congestion area) and pays the associated additional costs (i.e. difference between the extra-marginal cost of the constrained-on generators in region B and the APX price). Generators for downward regulation in A are selected through the auction (which may be held monthly), whereas generators for upward regulation in B are selected on a centralized market. The latter market needs to be an hourly market.

This approach incurs net income to the system operator in zone A (all non-constrained units will pay a congestion fee) whereas the contracting of compensatory power in B incurs costs to the system operator. The net costs to the system operator depend on the amount of congestion. If the amount of redispatched generation is not too high relative to the volume of plant in region A, and constrained on costs are not excessive, the system is very likely to yield additional revenues for the system operator. However, since in this model the generators pay only the constrained off costs, and the system operator funds constrained on costs, the revenues for the

[23] The APX price is used as a proxy for the market price in each hour. Since the constrained-off generator avoid generation themselves, their net 'costs' more or less equals the generation surplus they have bid in (assuming they have sold their output at the APX price).

system operator and net costs for the generators will be much lower than under the market despatch model.

Illustration by means of the reference model

When placing the bids in the congestion auction, each generator in region A estimates its surplus (which equals the APX price minus the avoided for generation). These bids represent the surplus of each generator. G1 will therefore bid in at 35 €/MWh (the APX price) minus 20 €/MWh (its marginal cost) = 15 €/MWh and G2 at 5 MWh. G3 does not place a bid in this example.

For sake of clarity, a key distinction between the market redispatch model and the hybrid redispatch model involves the different bids of generators in region A into the auction (although the numbers coincide in the present example):

- In a market redispatch model, generators in A bid according to the lowest estimated costs of contracting generation in B (i.e. *costs of marginal plant in B*) minus their avoided cost for not running in A (i.e. the marginal cost of generation in A). For placing a bid, generators in A need therefore to have price information on generation in B.
- In a hybrid redispatch model, generators in A bid according to the value they attach to not having to generate power in A (with the contractual delivery be carried out by the system operator), i.e. the market value of generation minus the avoided costs for not running in A. This value is well known since it is internal to the company.

After collecting the bids on a bidding ladder, the system operator will conclude that G2 represents the most expensive plant and therefore provides the 'cheapest' output reduction. The system operator therefore assigns the constrained-off capacity to the amount of 250 MW to G2, who will consequently reduce its output from 1000 MW to 750 MW. The system operator however takes over G2's supply obligation and G2 will therefore pay the market price (e.g. established by the APX price) to the system operator.

In region B, the system operator also organizes a market (which must be an hourly market). When all generators in B are obliged to submit bids to the system operator for their non-used generation capacity24, the system operator will be able to call the generators with the lowest marginal costs first. In the example, G4 already runs at its maximum output. G5 may bid its remaining 250 MW at a marginal costs of 35 €/MWh and G6 will bid 1000 MW at 45 €/MWh. The system operator will now contract 250 MW with G5 (the cheapest plant in B) to compensate for the reduction of 250 MW in zone A (which was needed to solve the congestion).

In summary, for solving a congestion of 250 MW, G2 will be called to ramp down and G5 will be called to ramp up. G2 will now pay 250 MW × 35 €/MWh = € 8750 to the system operator. For this, the system operator takes over the supply obligation from G2 to customers in region B. The 'loss' for G2 is its surplus of 250 × 5 €/MWh, being the 250 MW × 35 €/MWh (the price to be paid to the system operator) minus 250 × 30 €/MWh (its avoided costs).

The congestion price is set at the highest called-off bid, i.e. 5 €/MWh for this hour. This congestion fee is paid by all remaining generation. Therefore G1 and G3 (who were not selected) as well as G2 (for the 750 MW at which it will continue to run) must now pay a congestion fee to the system operator of 1000 MW × 5 €/MWh = € 5000 for G1, 750 MW × 5 €/MWh = € 3750 for G2 and 0 MW × 5 €/MWh = € 0 for G3, which amounts to € 8750.

[24] This single-buyer market will reflect the imbalance market, where all generators need to bid in non-used generation. However, whereas for the imbalance market only those generators are relevant which may respond quickly (within 15 minutes), for the congestion market also generators who need more time to respond and block bids should be allowed, which will make the market more liquid. It may be advisable to make bidding into this congestion market (in zone B) mandatory in order to increase the liquidity of the market.

For constrained on power in zone B, the system operator contracts 250 MW at 35 €/MWh with G5, which (in the example) equals the market price. The costs for this are € 8750, but in this example these costs are exactly offset by the revenue received from the constrained off generator. Now the total balance for the system operator is the following:

Region A:	Income for taking over supply obligation from G2 (250 MW at APX price)	+€ 8750
	Income due to congestion fee (1750 MW at congestion price)	+€ 8750
Region B:	Costs of constrained-on generation (250 MW at extra-marginal price in B)	-€ 8750
Net congestion revenue for system operator		**+€ 8750**

Just as in the market redispatch model, application of the hybrid redispatch model lowers the net income of generators in region A by 5 €/MWh. However, it should be noted that there is a basic difference since the in market redispatch generators in A also pay for the constrained on costs in region B (which in the example happen to be at the market price). When the constrained on costs in region B are higher than the market price (which will often be the case), the additional costs are allocated to the generators in A in the market redispatch model whereas these extra costs are paid by the system operator in the hybrid redispatch model.

Schematically, the hybrid redispatch model can be represented as follows:

REGION A		BEFORE	REGION B	
Output G1	1000 MW	TRANSPORT CAPACITY: 250 MW	Output G4	1000 MW
Output G2	1000 MW		Output G5	750 MW
Output G3	0 MW	TRANSPORT A→B 500 MW	Output G6	0 MW
Demand in A	1500 MW		Demand in B	2250 MW
Total generation in A	2000 MW		Total generation in B	1750 MW

CONGESTION BIDS IN REGION A			BIDS IN CENTRALIZED MARKET IN B	
G1	15 €/MWh	Market price: 35 €/MWh	G4	-
G2	5 €/MWh		G5 (250 MW)	35 €/MWh
G3	-		G6 (1000 MW)	45 €/MWh

REGION A		AFTER	REGION B	
G1 (not selected)	1000 MW	REDUCTION OF 250 MW APPLIED	Output G4	1000 MW
G2 (selected)	750 MW		Output G5 (selected)	1000 MW
Output G3	0 MW	TRANSPORT A→B 250 MW	Output G6	0 MW
Demand in A	1500 MW		Demand in B	2250 MW
Total generation in A	1750 MW		Total generation in B	2000 MW
CONSTRAINED OFF	250 MW		CONSTRAINED ON	250 MW

4.6.2 Economic evaluation

Advantages

The hybrid redispatch model allows solving congestion at the least cost, i.e. it will achieve short-term economic efficiency. Generators in region A who may ramp down at lower costs than the congestion price are happy to reduce their output (which is cheaper than paying the congestion fee) whereas generators whose redispatch costs are higher than the congestion price are better off by just paying the congestion fee.

As with market redispatch, since generators are not compensated for solving congestion in region A, there are no market power problems in the congestion market for constrained off services. Furthermore, contrary to system redispatch, the congestion costs cannot get too much 'out of control' since the costs for purchasing compensatory power in region B will be mostly (if not entirely) offset against the revenues generated by the congestion fee to be paid by generators in region A. The auction in zone A provides a tool for direct (market) valuation of the constrained-off congestion costs.

Furthermore, it is an advantage that generators in A can just bid into the auction at their lost profit, rather than also having to estimate the cost of constrained on power in region B. Smaller generators in A, who may find it more difficult to obtain compensatory power in B, would be more likely to place competitive bids in the auction under the hybrid model.[25] The centralized market in zone B should ensure that constrained-on generation can be purchased at relatively low transaction costs.

As with the market redispatch model, since the congestion fee effectively reduces the market price received by generators in region A, the hybrid model provides efficient incentives for both consumers and generators to locate new plant in places which relieve congestion. However, in the hybrid model the cost to generators in the constrained region is less than in the market redispatch model, since costs are based only on constraining off plant. Hence the hybrid model is more likely to achieve long-term efficiency than the market redispatch model, by avoiding excessive price signals.

Disadvantages

The main disadvantage of the hybrid model is that it is still vulnerable to market power in the constrained on market, since the system operator must buy replacement power. However, because any plant outside of the constrained area can provide constrained on power, market power issues should be less severe than in the constrained off market, where only a handful of plant may be able to offer a service.

From a generator perspective, hybrid redispatch (as market redispatch) is obviously less attractive than system redispatch. In system redispatch, the generator may get some revenues from being redispatched (depending on the difference between the redispatch price paid by the system operator and its marginal costs). With hybrid redispatch, the best a generator in region A can do is minimizing its costs (by placing a bid at its actual marginal cost in the auction for constrained-off power in zone A).

Another 'disadvantage' of the hybrid model (which however takes place with all models except for market splitting) is that the there are no incentives for demand to locate in region A, since prices are uniform throughout the country. Finally, as with market redispatch the effective prices that generators in different parts of the Netherlands receive are changed which may create 'winners' (generators with a majority of plant downstream of constraints) and 'losers' (generators with a majority of plant behind a constraint). However, the price changes will be less than under the market redispatch model.

[25] For example, under market redispatch a single generator in region B may be the cheapest plant to constrain off. But the small plant is worried about the high price of obtaining constrained on power bilaterally in region B that its total bid is very high. This problem does not occur under the hybrid model, since the small generator need only bids its own lost profit.

4.7 Market splitting approach

4.7.1 Description of the model

When congestion occurs in a market splitting system, the area upstream of the constraint (zone A) has a lower price than the area downstream of the constraint (zone B). Specifically, the marginal generator in the upstream market sets the price for the upstream market, and the marginal generator in the downstream market sets the price there. When there is no congestion, all prices are equal. Market splitting is the system applied within the Scandinavian markets (Nordspot) and in Italy. It is also applied to manage cross-border congestion between the Netherlands, France and Belgium.

Economically, for generators in zone A market splitting is broadly equivalent to the hybrid model discussed above. The most expensive generators in the constrained zone are in effect constrained off without compensation, and generators running in the constrained zone receive a price equal to the single market price less the constrained off cost of the marginal plant in the constrained zone. The main differences between the hybrid model and market splitting is that, under market splitting generators outside the constrained zone, i.e. in zone B, will get a price higher than the ‘single market’ price. In the hybrid model, the generators outside of the constrained area will get the single market or average price.[26]

Note that under market splitting, there is no explicit congestion management cost. The hidden congestion management cost for generators is that in the constrained zone they receive a lower price than outside of the constrained zone. Market splitting creates revenues for the system operator, equal to the price difference on either side of the constraint multiplied by the transmission capacity out of the constrained area. These revenues may be used to relieve congestion by investing in more transmission capacity.[27]

In some cases of market splitting customers will also enjoy the lower price in the constrained zone, which should encourage industry to locate in the constrained zone, thereby relieving congestion.[28] However, this need not be the case. In the Italian system consumers pay an average national price, not the price in each zone.

Illustration by means of the reference model

For region A, the market price is set by G2 at 30 €/MWh. The price is lower than the market price for the integrated market (which is 35 €/MWh, see section Appendix IV). For region B, the market price is set by G5 at 35 €/MWh, which equals the market price for the integrated market. In the case of G5 having a capacity of less than 750 MW, G6 would set the marginal price in region B, which would then be 45 €/MWh.

From the example it is clear that the market price in the low price zone is always less than or equal to the market price for the integrated market (without congestion). The market price in the high price zone is always higher than or equal to the market price for an integrated market. All consumers in the low-price zone (A) pay the market price for their region. All generators in the low-price zone (A) only receive the (lower) market price for their region. The same holds for generators and consumers in B. Therefore, consumers in A and generators in B are now better off compared to a single-price market (they pay less

[26] It is also true that if under the hybrid model the revenues collected from generators in zone A are redistributed to generators in zone B, the latter generators will get an effective price higher than the simple single market price. However, even in this case, the price for downstream generators may be less than under market splitting. It makes therefore more sense to redistribute the revenues among consumers in region A instead of generators.

[27] See also section 6.5.

[28] Though it is debatable how responsive industry would be to such price signals, especially if by locating in the zone they relieved congestion and caused price to increase back to the ‘single price’ level.

respectively they receive more). Consumers in B and generators in A are worse off in a situation of market splitting (they pay more respectively receive less). However, given the limitations of the network (due to congestion no more power can be transported from A to B), these results represent the actual economic value of electricity in regions A and B.[29]

Schematically, market splitting can be represented as follows:

AVAILABLE GENERATION IN A		CAPACITIES	AVAILABLE GENERATION IN B	
Output G1	1000 MW	CONSTRAINT: ALLOWED TRANSPORT A→B 250 MW	Output G4	1000 MW
Output G2	1000 MW		Output G5	1000 MW
Output G3	1000 MW		Output G6	1000 MW
Demand in A	1500 MW		Demand in B	2250 MW

MARKET BIDS FOR SUPPLY IN A			MARKET BIDS FOR SUPPLY IN B	
G1	20 €/MWh	FIND MOST EFFICIENT DISPATCH	G4	25 €/MWh
G2	30 €/MWh		G5	35 €/MWh
G3	40 €/MWh		G6	45 €/MWh

REGION A		DISPATCH	REGION B	
G1 (selected)	1000 MW	DIFFERENT MARKET PRICES IN ZONE A AND B	G4 (selected)	1000 MW
G2 (selected)	750 MW		G5 (selected)	1000 MW
G3 (not selected)	0 MW		G6 (not selected)	0 MW
Effective demand in A	1750 MW		Effective demand in B	2000 MW
MARKET PRICE A (G2)	30 €/MWh		MARKET PRICE A (G5)	35 €/MWh

4.7.2 Economic evaluation

Advantages

The advantages of the market splitting model are in essence the same as the market redispatch model and, to a lesser extent, the hybrid redispatch model, the main advantage being that congestion costs cannot get 'out of control' since there are no explicit congestion costs that can be allocated to consumers. Furthermore, since generators are not compensated for solving congestion, there are no market power problems in the congestion market.

Market splitting gives incentives for both consumers and generators to locate in places which relieve congestion, reduces the need for transmission network reinforcements, and increases long-term efficiency.

Transaction costs are relatively low, since trading can be carried out via a centralised pool.

---

[29] Market splitting forms a kind of zonal simplification of 'nodal pricing', which is considered best practice in electricity pricing.

Disadvantages

The key disadvantage is that the introduction of market splitting would change, perhaps significantly, the prices that generators in different parts of the Netherlands receive. This would create 'winners' (generators with a majority of plant downstream of constraints, consumers behind a constraint) and 'losers' (generators with a majority of plant behind a constraint, consumers downstream of a constraint). This redistribution of wealth among generators and consumers could create resistance to implementation from the 'losers'. When market parties have taken long-term positions, a change in the market design may lead to income loss, which may give rise to claims.[30]

Similarly, introducing a new price structure can increase regulatory risk, if it is done without sufficient warning. This is particularly true for generators who have committed to build plant in a congested area based on the expectation of a single price in the Netherlands. Increased regulatory risk could undermine future incentives to invest in the Netherlands, although the force of the latter argument must not be exaggerated given the extremely large amount of new plant announced.

4.8 Other approaches considered

We considered a number of other congestion management mechanisms that for simplicity are not described or evaluated in this report. These models were evaluated during the consultation process, and scored less well than the selected models on all the important criteria.

These models e.g. involve auctioned 'congestion rights' in which *ex ante* transport rights are auctioned among generators. Only generators who obtained such rights are then allowed to transport through the congestion. This model suffers, like market splitting, from the required major change in market design. Furthermore, congestion needs to be accurately assessed in advance, contrary to the models discussed above where a more dynamic approach may be adopted.

Another approach concerns the auction of 'compensation rights', which provides a way to control congestion costs by selling a limited number of compensation rights to generators, which would entitle them to compensation for being constrained off. As congestion costs increased, the volume of compensation rights could be reduced. The main disadvantages of this scheme is that it is very complicated to implement, and offers no real advantages in terms of cost control relative to a system redispatch model where costs can be targeted at generators.

[30] This disadvantage only holds for market splitting as it interferes with the market design. Since system redispatch, market redispatch and hybrid redispatch are *ex post* models, the basic market design is left unchanged (although there may be an indirect effect on the electricity price).

5 Assessment of the options

5.1 Introduction

In the present chapter, the different congestion management options will be assessed according to the assessment criteria presented in chapter 3. The assessment presented is based on an independent analysis of *D-Cision* and *The Brattle Group* themselves, although the input of market participants, submitted during the workshops or in reply to consultation documents, has been taken into account as much as possible.

The assessment will distinguish between the models presented in chapter 4. However, the 'system redispatch model with cost pass-through to generators' is not considered separately, since it scores more or less similar to the basic system redispatch model (except for the final cost allocation).

5.2 Overview of the assessment

The assessment of the models is presented in Table 3. Specific comments on the assessment are given below.

5.3 Assessment by the technical criteria

For most of the technical criteria, the differences between the congestion management models are rather limited. All of them are *applicable* and all of them are *effective* in the sense that they are able to technically solve congestion (by reducing the net flow over the congested link). Similarly, *network safety* can be guaranteed under each model.

The major differences are with the *information supply* criterion. Under the market agent model, market parties may overstate their transport needs (given the *pro rata* constrain-off obligation, the system operator ending up with wrong network information. Under system redispatch, market parties might also be willing to submit wrong estimates, in order to defend higher bids in the constrained-off markets.

5.4 Assessment by the economic criteria

With respect to *economic efficiency*, all models except for the market agent model achieve economic efficiency. In the market agent model, economic efficiency may be achieved eventually if there is a liquid secondary market for trading constrained-off obligations (which is not very likely to develop). The market redispatch model performs slightly worse than the hybrid and system redispatch models since generators in the constrained region are charged with the constrained-on costs as well, which may eventually affect the height of their supply offers.

The models score differently with respect to *gaming possibilities*. Especially the system redispatch model is very vulnerable to gaming (leading to higher costs to the system operator). A specific kind of gaming is also possible in the hybrid and market redispatch models, although this will increase the payment to the system operator for constrained-off generation. However, the consequence is that the congestion fee revenues will decrease as well. Under market redispatch, the system operator will never end up with positive costs. Under hybrid redispatch, in the worst case the constrained-off revenue to the system operator will reduce to the market price times the constrained off power, whereas the constrained-on costs (to be paid by the system operator) will be a bit higher than this. Additionally, there may be market power in the constrained-on market which will show up in a centralized market but may be present in a hidden (and possibly more serious) form under the market agent and market redispatch models as well. A more extended discussion of market power and its mitigation is provided in section 6.2.

Table 3. Assessment of the congestion management options according to the evaluation criteria.

		RELEVANCE	MARKET AGENT	SYSTEM REDISPATCH	HYBRID REDISPATCH	MARKET REDISPATCH	MARKET SPLITTING
Technical criteria	*Applicability*	MAJOR	++	++	++	++	++
	Effectivity	MAJOR	++	++	++	++	++
	Network safety	ESSENTIAL	++	++	++	++	++
	Incentives for information supply	MINOR	−−	+/−	++	++	++
Economic criteria	*Economic efficiency*	ESSENTIAL	−−	++	++	+	++
	Robustness (gaming possibilities)	MAJOR	−	−−	+/−	−	++
	Economic incentives	MAJOR	−	−−	++	++	++
	Transaction costs	MINOR	−	+	+	−	+
Policy criteria	*Proportionality*	MAJOR	+/−	++	++	+/−	−
	Priority access for renewable energy	ESSENTIAL	++	++	++	++	−
	Accuracy of cost assessments	MAJOR	++	−−	+	++	++
	Speed of implementation	ESSENTIAL	++	+	+	+	−−
Regulatory criteria	*Non-discrimination*	MAJOR	++	++	++	++	++
	Simplicity and transparency	MINOR	++	+	+	+	+
	Incentives for investment	MAJOR	+	−−	++	+	++
	Incentives for decommissioning	MINOR	+/−	−−	+	+/−	+

The models rank differently as well with respect to the *economic incentives* given to efficient network use. The consequence of the system redispatch model is effectively that network congestion is ignored by market parties. In the end, the system operator will solve this. However, since the costs are not directly allocated to the parties causing the congestion inefficient dispatch may result. The market agent model scores low as well, since all generators in the congestion zone are equally ‘punished’ for causing congestion, despite some being efficient and others inefficient (i.e. producing at a higher cost than the marginal price for generation). The market and hybrid redispatch models and market splitting do provide better economic incentives, since they effectively lower the market price for generation in the congested region.

With respect to *transaction costs*, the system and hybrid redispatch models as well as market splitting score best. The reason is that constrained-on generation is contracted by the system operator through a centralized market (for which however, generators in the constrained-on area need to supply hourly bids). Under the market redispatch model constrained-off generators must contract constrained-on generation themselves, which implies higher transaction costs. Furthermore, the costs of IT and information systems are more or less comparable under all models.

5.5 Assessment by the policy criteria

On the criterion *proportionality* (applied to slightly congested systems[31]) especially the system and hybrid redispatch model score best since no special efforts are required from market parties apart from submitting bids to the system operator. Under the market agent and market redispatch models, generators in the constrained zone are obliged to contract compensatory generation themselves, which is apparently more ‘burdensome’. Especially the introduction of market splitting can be considered disproportionate for systems with limited congestion.

All models except for market splitting are able to easily cope with the *priority access for renewable energy* criterion:

- *Market agent model:* Under the market agent model, renewable plant can be excluded from being constrained off (the *pro rata* allocation of constrained-off plant being limited to conventional generation only)
- *System, hybrid and market redispatch models:* Under these models renewable plant can be exempted from the obligation to submit bids for constrained-off generation. If the availability of conventional generation is insufficient to solve the congestion, renewable plant can be required to submit bids as well, but will only be called off when all grey plant has been constrained off already.
- *Market splitting:* Under market splitting, implementation of priority access for renewable energy is more difficult to implement. The reason is that market clearing is solely based on submitted price bids. It is difficult to envisage why a bid of renewable generation at a higher supply price than the market price needs to be accepted, unless there are subsidies in place (which however are outside the scope of a congestion management mechanism).[32]

With respect to the *accuracy of cost assessments*, especially the system redispatch model scores badly since this model delivers net costs to the system operator, which highly depend on the

---

[31] See section 3.4.1.

[32] The analysis assumes that under market splitting, the dispatch of all generation is managed by the power exchange (which depends on difference between the clearing price and the bid price). If renewable generators bid at their real costs, they may not be called off at all times. However, if a subsidy scheme is in place which compensates renewable generators for their above market generation costs, they may bid in at lower prices. However, this approach makes the congestion management scheme dependent on the presence of such subsidies.

amount of congestion and the bids placed. The market agent, market redispatch and market splitting models do not involve costs to the system operator and therefore score positively. In the hybrid model the costs for constraining on are wholly or partly offset by the revenues from the constrained off zone. Although it is likely that the system will deliver a net revenue, in the worst case the costs are restricted to the above market costs of constrained-on generation.

The models score differently as well on the *speed of implementation*. For market splitting, the market design needs to be changed, which cannot be done quickly. For the three redispatch models, a market mechanism for constrained-off generation needs to be developed as for constrained on generation in the system and hybrid redispatch models. This takes some time to work out. The market agent model can be implemented quickly since no market place needs to be established.[33]

5.6 Assessment by the regulatory criteria

The final set of criteria concerns the regulatory criteria. With respect to *discrimination* all models score positive since they treat comparable generators in a similar and consistent way. The same holds for the criterion *simplicity and transparency*. Although the models may look a bit complex on first sight, the underlying economics of each model is quite straightforward and not too different from what is common practice in the industry. The market agent model even scores a bit better since no economic mechanism is involved.

The models score significantly differently with respect to *incentives for (generation) investments*. The market agent and market redispatch models score positive on this criterion, since generators in the congested zone pay both the constrained-off and constrained-on costs of redispatch, which will effectively deter investment. Nevertheless these incentives are stronger than the economic optimum which may eventually result in underinvestment. The hybrid redispatch model and market splitting also charge generators in the congested zone, although only for the constrained-off costs. These result in an effective market price which coincides with what is to be expected from economic theory. On the contrary, generators in the congested zone are quite insensitive to the costs associated with congestion under the system redispatch model. This is still true in the system redispatch model with cost pass-through to generators, where the generators in the congestion area effectively still receive a too high surplus given network congestion. Therefore, system redispatch is not expected to give a sufficient locational incentive to achieve long term efficiency.

Finally, the models score differently as well on the *incentives for decommissioning*. As can be expected from the previous criterion, the system redispatch model scores worse (since it may be attractive to keep old plant and let it participate in the congestion management mechanism). The market agent and market redispatch models score average since old plant is not incentivized to remain in operation in the congestion zone on the one hand but charging all congestion costs to these generators may incentivize some efficient plant to stop operation as well.

5.7 Overall assessment

On the four criteria which are considered ‘essential’ (network safety, economic efficiency, priority access for renewable energy and speed of implementation) all models score positive, except for the market agent model (which does not result in an economic efficient allocation of generation unless a liquid bilateral market develops for constrained-off generation) and market splitting (which cannot be quickly implemented and for which compliance with priority access

[33] In this assessment, legal issues possibly hampering implementation have not been taken into account.

for renewable energy may be more challenging). The main choice for a congestion management mechanism therefore seems to be between the system, hybrid and market redispatch models.

The most relevant other criteria (identified as ‘major’) where these three models differ are robustness (gaming possibilities), economic incentives, accuracy of cost assessment and incentives for investments. The system redispatch model performs rather badly on each of these: The costs may easily get out of control, there are significant gaming possibilities and no incentives are given to deter new generation in the congestion zone.

Of the remaining two models, hybrid and market redispatch, the former seems to have a slightly better score, which is mainly due to the economically preferred level of charging generators in the constrained zone and the exercise of market power in the bilateral constrained-on market which is extremely difficult to address.

In conclusion, the hybrid redispatch models seems to offer the best balance given the evaluation criteria presented in chapter 3 and is therefore suggested as the preferred model to manage national congestion in the Netherlands.

6 Issues related to the implementation of congestion management

6.1 Priority access for renewable energy

All congestion management models except for market splitting are able to cope with the priority access for renewable energy constraint:

- *Market agent model:* Under the market agent model, the obligation to reduce generation could be assigned only to non-renewable generators. Renewable generation would then have unrestricted access to the grid. However, it must be guaranteed that renewable generation will not participate in the secondary market where constrained-off obligations are traded.
- *System redispatch, market redispatch and hybrid redispatch models:* Under these models, renewable generators would be exempt from the obligation to bid into the constrained-off auction. Depending on the policy preference they may or may not be charged the congestion fee.
- *Market splitting:* Depending on the exact definition of priority access, under market splitting priority access for renewable energy could be more difficult to implement, since the clearing mechanism of a power exchange might dispatch 'competitively priced' renewable generation but not despatch 'non-competitively priced' renewable generation. The latter involves renewable plant which would not be dispatched in an unconstrained case either. Under some interpretations of priority access for renewable energy, all bids of renewable generation might have to be accepted, irrespective of the bid price. In this case also renewable generation with marginal cost higher than the market price would be called off. A separate (policy) decision needs to be made whether this would be acceptable or not and if yes, how the above market costs of this renewable plant is to be paid for (without raising the market price).[34]

The most transparent (and economically sound) approach to financially include renewable generation in the congestion management model is to let both renewable and conventional energy sources in the congestion area pay the congestion fee. The underlying idea is that the electricity price is lower in the congestion area as a consequence to the excess generation. The appropriate way to compensate for the profits reduction of renewable generators in the congested zone is then to adapt the 'Subsidieregeling Duurzame Energie' (SDE) to award higher subsidies to these renewable generators (so that the net profits of similar renewable generation remain equal anywhere in the Netherlands). However, given the complexity of this approach, simply excluding renewable generators from paying the congestion fee has some merits.

6.2 Mitigation of market power

Given the experience elsewhere, the assessment of gaming options is of special importance in the assessment of congestion management models.

Market power in this context means that generators are able to sustainably increase their profits (and therefore the costs of congestion) by submitting bids in the congestion management scheme which do not reflect their costs. Market power may materialise both in the constrained-off market and in the constrained-on market. The different models will be assessed on this below.

---

[34] A solution might be to assign a low bid price to renewable generation and financially compensate the renewable generators for the difference between their generation cost and the market clearing price by a subsidy scheme. See also footnote 32.

6.2.1 Market power under the market agent model

The issue of market power is not a concern in the market agent model (at least for consumers), since no payments to the system operator are involved. However, the model may exhibit market power in the bilateral market, since generators may need to contract bilaterally for constrained on power. Some smaller plant in particular may become victim to other generators who are only willing to take over the redispatch obligation at an unreasonable high price. It is difficult to address this kind of market power either ex ante or ex post.

6.2.2 Market power under the system redispatch model

As mentioned, one of the main disadvantages of system redispatch is that it is vulnerable to gaming or market power (section 2.3.3 describes the way in which market power is exercised). Such market power will have the effect of offering too high prices for constrained-off and constrained-on power to the system operator (so that the system operator pays more than the marginal costs in region B for constrained-on power and received too little for constrained-off power in region A).

There are various ways to prevent or mitigate market power in congestion markets. These include:

- *Capping of constrained off price:* The idea here is to place a floor on the amount that the system operator will pay generators for being constrained off. For example, the system operator could estimate the avoided costs of various plant, and refuse to pay plant more than this amount for being constrained off. Similarly, the system operator could audit the costs of plant.
- *Use of long-term contracts:* The system operator could enter into long-term (e.g. one year) contracts with generators for constrained on and constrained off services. The underlying idea is that competition to provide services on a long-term basis is likely to give a more competitive outcome than short-term purchases in the constraint market. A drawback is that this will take out generation capacity out of the market, which may slightly increase the average market price, though the overall effect should be a reduction in costs for consumers relative to a situation without long-term contracts, since the savings on constraint costs should offset the price effect.[35] In the event that there is only one generator that can provide the constraint services, the competition authority should have powers to compel the generator to sign a service contract with the system operator if the terms seem 'reasonable'.

Despite these suggestions, it must be noted that market power is very difficult to mitigate completely. In its July 2007 report for TenneT on connection policy,[36] The Brattle Group noted that while market power mitigation in constraint markets is possible, it can be burdensome for the parties involved and cause more problems than it solves. For example, it could be difficult to establish a 'fair' constrained off price, especially when being constrained off involves one-off shut down and start up costs. Such costs can make even a relatively small reduction in power output very expensive on a per MWh basis. For this reason the potential for market power, and the costs involved in mitigating it, are a major disadvantage of a system redispatch scheme.

[35] In a short-term constrained on market, only generation which has not yet been dispatched will compete. The reason is that all generation cheaper than the spot market price, is already called off. In case of longer term contracts, it is possible that some generation exists which has lower marginal cost than the market price but is nevertheless not dispatched in the market since it is reserved by the system operation as constrained on power.

[36] *Loc. cit.* footnote 10.

6.2.3 Market power under the market redispatch model

Under the market redispatch model, no significant market power problems exist since generators are not compensated for solving congestion. However, in the example of a generator having a significant volume of both constrained-off generation and non-constrained-off generation in the congestion area (region A), or collusion between generators, there might be an incentive to bid in constraint costs at 0 €/MWh. In a competitive market the generator would not do this, since if there was another generator that could solve constraints more cheaply, the generator would rather pay this lower marginal constraint cost than pay his own higher constraint costs. But without sufficient competition, the generator may prefer to risk paying slightly more than the marginal cost for being constrained off, if the remainder of his plant benefit by paying a zero constraint charge. In the worst case there would be no revenues for the system operator.[37] But there would be no costs for consumers here as well, since generators bear all the constraint costs.

It is also not clear if such a situation would be a stable equilibrium. There would be an incentive for other generators to locate in the constrained area since constraint costs are zero for non-constrained off plant in region A, and this would increase competition between generators in region A, making a revenue for the system operator more likely.

6.2.4 Market power under the hybrid redispatch model

In the hybrid model, market power may exist in the constrained-on market. However, given that any generator in region B can participate in this market, the constrained on market is expected to be much more competitive than the constrained off market, where only one or two generators may be active.

Since the system operator procures constrained on power, this relieves some of the bilateral market power problems that could occur under the market redispatch model. Smaller generators in A, who may find it more difficult to obtain compensatory power in B, would be more likely to place competitive bids in the auction under the hybrid model.

A specific form of market power may show up under the hybrid model, which could reduce the revenues from plant in region A (which the system operator uses to fund constrained on costs) to zero. For example, assume a generator G in A with a capacity of 1000 MW, of which 250 MW must be constrained off. Furthermore, assume the generator surplus equals 5 €/MWh.[38] Now, when G bids in at its marginal costs, it will pay 250 MW times the market price minus its marginal costs (i.e. 250 $\times$ 5 €/MWh) plus 750 $\times$ 5 €/MWh as a congestion fee for the unconstrained output, which totals €5000. However, by lowering its bid in the constrained-off market to 0 €/MWh, the total costs for G will reduce to 250 $\times$ 5 €/MWh or €1250. The consequence is that the system operator will not collect any congestion revenues from which to fund constrained on costs.

First it must be noted, that in this example, all constrained-off costs are now absorbed by G. Therefore, the hybrid redispatch model still performs much better than the system redispatch model, where especially market power in the constrained-off zone cannot be mitigated easily.

[37] For example, suppose there are 10 200 MW plants in region A, all owned by different generators, and one of the 200 MW must be constrained off. Each generator would bid their true costs of constraints, since if there is a cheaper generator available each generator would rather pay the lower cost. Now imagine that all of the plants are owned by the same generator. The generator would simply bid 0 €/MWh, even if the actual cost of being constrained off was 10 €/MWh. In this way, the generator loses 10 €/MWh on the 200 MW of plant that is constrained off, but avoids paying anything for the remaining 800 MW of plant that is still running.

[38] This generator is similar to G2 in the reference model of section of 0.

Under the hybrid redispatch model, any market power problems only exist in the constrained-on market. In the worst-case all constrained-off costs are taken up by the generators in congested region and the system operator needs to pay for the extra-marginal costs of the constrained-on plant only.

The effect can possibly be mitigated by applying a minimum price for constraining off (e.g. a percentage of the APX price or an administrative charge), which is then also used for defining the level of the congestion fee. An argument for this is that, given the congestion, the marginal costs for generation in the congestion area must be lower than the market price – otherwise the plant would not be running and there would be no congestion. Another option is to apply a charge to generators in region A to make up any shortfall in costs from finding constrained on payments.

Alternatively, running longer-term auctions (as the monthly auction suggested) may decrease the likelihood of strategic underbidding of generator surplus, since this will only have a positive effect in those hours when G will have residual output when constrained off, which may not be the case for each hour in a longer time period. Finally, a different auction design could be applied, for instance by adopting a Vickery implementation, where the clearing price is set at the least rejected costs (and where, additionally, bids of the marginally accepted generator could be excluded).

A specific improvement in the design of the hybrid model could further mitigate this effect. We call this ‘variant B’, while the original scheme described above is variant A. Instead of requesting bids for the generator surplus and accepting the lowest bids (i.e. the highest generation cost) in the congestion area, the system operator could ask the generators to bid according to their marginal cost. When they are called off, they would be charged their bid price (‘pay-as bid’ auction), with the system operator taking over their supply obligation – exactly as in system redispatch. All other generators could then be charged the difference between the market price (i.e. the APX price) and this lowest accepted bid of constrained-off plant.[39]

The variant differs from the model as described in section 4.6 in the sense that the constrained-off generator does not lose their profits or surplus. By this, a distinction exists between constrained-off generators in the congestion area (who are allowed to keep their generation surplus with respect to the original market price) and generators in the congestion area who are not constrained off (who are charged this surplus by means of the congestion fee).

The effect of this approach is twofold. Firstly, the congestion income to the system operator is slightly lower since it must pay for the profit of the constrained-off generator. Secondly, generators have an incentive not to bid at 0 €/MWh. If they do bid zero, the generators who are constrained off will lose their entire generation profit. Under Variant A presented above, by bidding at zero the generator only risks losing the *difference* between his costs of being constrained off and the cost of the cheapest constrained off plant. When they bid at a lower price they may collect their generation surplus if constrained off. Generators like G will offer non-zero bid prices as well since they will balance their costs of the congestion fee to be paid by running generation with the revenues from the generation surplus for constrained-off plant (which they are allowed to keep under this scheme). This will effectively result in bid prices somewhere between G’s marginal costs and the market price.

In sum, which variant of the hybrid model is best depends on market power. Variant A collects more money in a competitive scenario, since the system operator does not compensate for the

[39] See Appendix V for a description of this variant.

lost profits of constrained off plant. But it is more vulnerable to market power if plants in region A are owned by a few generators. Variant B collects less money in a competitive scenario, since the system operator compensates plant for the lost profits of constrained off plant. But Variant B could work better where plants in region A are owned by a few generators, since generators have a greater incentive to offer non-zero bids so that they recover profits on constrained off plant. Note that if plant is owned by only one or two generators, it is likely that generators will bid zero or in such a way as the system operator has no net revenue.

6.3 Role of interconnections

The question whether interconnections should or should not participate in a congestion management scheme is a complex one.[40] There is however a clear difference between the market agent approach and system redispatch on the one hand and market redispatch and hybrid redispatch on the other.

- *Market agent and system redispatch models:* Under these models the choice to include interconnections does not have major consequences. If they do not participate, they are just excluded from the pool of possibly constrained-off generators (just as renewable generators). The price generators receive in the constrained-off zone is however not affected by the congestion management mechanism.
- *Market redispatch and hybrid redispatch models:* Under the market and hybrid redispatch models, one may exempt interconnections from providing constrained-off bids. The major question however seems to be whether interconnections are charged the congestion fee. On the one hand, it does not seem fair to (virtually) reduce the market price for generation in the congestion area by applying a congestion fee without charging this to imports. It could result in the weird situation of reducing the effective value of local generated electricity but keeping the electricity price of imported electricity unchanged (letting import unrestrictedly feed in into the congested area thereby price discriminating between imports and local generation). On the other hand, changing the market arrangements with foreign system operators may be a painstaking process, which will endanger quick implementation of a national system congestion management.

  In the case interconnections are treated non-discriminatorily with generators in the congestion area, they might be called off as well to reduce imports (which evidently has an impact on the flow through the link). If market redispatch is adopted, parties importing through the interconnection are under the obligation to purchase constrained-on generation instead which may be complex for foreign market participants. If interconnections are not called off (e.g. when they bid in at a high price), market and hybrid redispatch reduce to a financial settlement only (i.e. payment of the congestion fee for each MWh imported).[41]

Finally, a system of market splitting only makes sense if interconnections are included. Otherwise suboptimal outcomes may result. However, in practice market splitting will probably

---

[40] The discussion focuses on the management of congestion in the national grid excluding the interconnections. When international tie lines are included in the (national) congestion management scheme, the '*Congestion management guidelines*' apply (included in the annex of Regulation 1228/2003).

[41] A different issue is how easy it is to charge the congestion fee to imports into a congestion area. It is likely that a legal basis must be created for such a charge. However, since the fee is related to the effective market price level in the congestion area (which is virtually reduced by the congestion fee under the market and hybrid redispatch models), the charge is not a transport charge but implies a market price correction.

only be implemented in the Netherlands within the wider scope of the regional Northwest European Market.

6.4 Information requirements

Obviously, a network operator needs sufficient and reliable network information in order to identify potential network congestion in advance. Only when congestion is identified, any mitigating action can be initiated. As a consequence, two information requirements can be distinguished, namely:

- Information related to identifying congestion in advance (or estimating the likelihood congestion to occur), and
- Information needed to manage the congestion, once it has been identified. The latter includes both monitoring actual flows and providing incentives (or penalties) to parties who deviate from their announced transport (especially when this aggravates congestion).

These information requirements apply to all models, irrespective of the congestion management model implemented. In practice, several approaches are conceivable, like:

- *Collective T-prognosis:* This approach builds on the present practice of a rough estimation of transmission requirements, based on information provided by the network users on a voluntary basis, which are then transformed into a system-wide transport prognosis (‘T-prognosis’).
- *Individual T-prognoses:* In this approach, transmission estimates are provided to the system operator on an individual basis (like the ‘C-programs’ in the Westland congestion management system).
- *T-programs:* Connected generators and consumers nominate network use in a transport program (‘T-program’). Contrary to T-prognoses, the T-programs are binding. Changes in the program are not allowed unless approved by the network operator and deviations are penalized.
- *E-programs:* Energy programs (‘E-programs’) concern commercial transactions and are locational unspecific since it encompasses the trading position in the entire Dutch electricity market. Presently, program responsible parties daily submit their E-programs to TenneT.
- *Differentiated E-programs:* In case of (expected) congestion, it may be considered to differentiate between transactions inside and outside the congested area. For each region a separate E-program needs to be submitted. Any cross-program flows now involve transactions over the congested link.

Each approach has its own characteristics, which are summarized in Table 4.

Table 4. Applicability of five information approaches for congestion management.

	Determination of congestion	Management of congestion
Collective T-prognosis	Presently applied for connections <2 MW. It is based on historic data and rather inaccurate[42]	Due to its non-binding nature not suited for financial settlement
Individual T-prognoses	Presently applied for connections ≥2 MW. The accuracy is limited (due to lack of incentives)	Due to its non-binding nature not suited for financial settlement
T-programs	Sufficient for assessing congestion, since it contains information on the location of generation	Sufficient for adding financial incentives to market parties to stick to their program (flows aggravating congestion can be identified)
Energy programs	Not sufficient for assessing congestion, since it does not contain information on the location of generation	Not sufficient to identify deviations from the plans aggravating congestion
Differentiated energy program	Sufficient for assessing congestion if the congestion area is location independent	Sufficient for adding financial incentives for market parties to stick to their program (flows aggravating congestion can be identified)

Both for assessing network congestion and for managing the congestion, the existing legislation and regulation seem to offer a more comprehensive framework than presently applied. Especially articles 5.1.1.2 to 5.1.1.9 of the Netcode, prescribing the submission of transport prognoses (T-prognoses) do not seem to be yet adhered to. We recommend that the network operators discuss a more extensive implementation of these articles with market parties.

6.5 Use of the auction revenues

A specific question is what should be done with the auction revenues resulting from the market redispatch model (as with market splitting) and likely from the hybrid redispatch model as well. Several options exist:

1. The revenues may be netted with the costs. Since the hybrid model the costs incurred during some hours, may be offset with the net revenues for other hours.
2. The revenues may be saved by the system operator to compensate for congestion costs made elsewhere. If different congestion mechanisms are applied in different areas, the revenues from one area can be used to pay for the costs in other areas.
3. The revenues may be used for network reinforcements relieving congestion. To this end, the costs might be absorbed by the regulatory system governing network investments.
4. The revenues may be allocated to generators outside the congestion area who (given remuneration according to the average market price for the integrated area) receive too little for their power. By adopting this approach, a welfare transfer takes place between generators in the congestion area to generators outside the congestion area. However, an economic argument can be made that not all revenues should be assigned to the generators

[42] For generators with a connection capacity larger than 2 MW, an obligation exists for submitting individual T-prognoses on an hourly and day-ahead basis (see §5.1.1.1 of the Netcode). For generators with a connection capacity smaller than 2 MW this obligation is replaced by an obligation for the programme-responsible party to submit (aggregated) T-prognoses (see §5.1.1.2 of the Netcode).

(i.e. excepting the revenues relating to the value of the congested link times the price difference).

5. Finally the revenues may be socialized through the network tariffs. Given that consumers in the congested area have apparently paid too much for their electricity (they paid the system marginal price for the integrated market which is higher than the marginal cost of generation in the congested region), it could make sense to socialize (at least a part of) the revenues in the tariffs of the consumers in this area. Alternatively, if consumer price differentiation is not desirable, the revenues may be socialized in the network costs of all consumers.

Apart from option 1, which makes sense anyway, a choice for one of the alternatives is mainly related to a policy (or political) preference.

6.6 Relation to the imbalance market

Congestion management has at least a threefold relation to the imbalance system:

Firstly, since congestion management aims to redispatch generators in order to reduce the electricity flow out of a congestion area, any *changes in the generator dispatch* afterward impacting these flows need to be discouraged (in order to prevent transports reaching the transportation limits again). In practice, this implies that (energy or transport) program changes which affect the flows across congested links should be prohibited. The latter can be implemented by splitting the energy program in accordance with the congestion and separately invoicing imbalances (deviations from the energy program) in the congestion area with high imbalance fees. If the network security is not affected, trade within a congestion zone may be allowed whereas trade between the congestion zone and the outside area could be penalized.

Secondly, the distinction between generators inside and outside a congestion area also plays a role when the system operator would like to dispatch generators for *upward regulation*. The system operator needs to take into account the location of the generator to be dispatched from the imbalance system’s bidding ladder. Effectively, this could lead to separate prices for upward regulation in both regions (since an imbalance due to a lack of generation in the congestion area might be solved at a lower cost than an imbalance outside the congestion area).

Finally, the *bids in the (centralized) constrained-on market* under the system and hybrid redispatch models, will partly reflect those in the imbalance market, where all generators are obliged submit bids for non-used generation. However, whereas for the imbalance market only those generators are relevant which are able to respond quickly (within 15 minutes), for congestion management also generators who need more time to respond as well as block bids (covering supply for a longer time period) should be allowed. This will not only make the market more liquid, this type of generation will probably be cheaper as well. Moreover, it may be advisable to make bidding into this centralized constrained-on market mandatory in order to further increase the liquidity of the market.

6.7 Payments within the congestion management model

Finally, a few words need to be spent on the financial implementation of the congestion management model, especially with respect to the payments to and from generators in the congestion area.

- *Market agent model*: The implementation of a market agent model is straightforward since no specific payments are involved.

- *System redispatch model:* Implementation of a system redispatch model is similar to the imbalance market. Accepted bids of constrained off and constrained on generators are settled bilaterally between the system operator and the respective generators. The total costs can be socialized in the transport tariff (system redispatch model) or charged to all generators in the constrained-off zone (system redispatch model with cost pass-through to generators). The latter will lead to flat surcharge on the G-tariff in the congested zone.
- *Market redispatch models:* Implementation of the market and hybrid redispatch models additionally demands charging the congestion fee to non constrained generators in the congestion zone. The most straightforward approach is to include the congestion fee in the transport fee (as a G-charge) for the congestion zone. The hourly congestion fees can be added for a month or a year and charged to the generators in retrospect.
- *Hybrid redispatch model:* The payments to the system operator for constrained-off generation may be settled bilaterally between the respective generators and the system operator. If this payment is corrected for the congestion fee, there is no need to differentiate the G-charge for constrained-off and non-constrained generators, so that a flat G-tariff can be used for all generators in the congestion area.[43]
- *Market splitting:* The implementation of market splitting is quite intrusive in the present market design and will not be discussed here.

The payments for constrained-on generation can be made directly between the parties involved, i.e. the system operator (in the system redispatch and the hybrid redispatch models) or the constrained-off generators (in the market redispatch model) on the one hand and the constrained-on generators on the other.

[43] The constrained-off generator has to pay the market price to the system operator but is exempt from paying the congestion fee. The former payment can however be spit in 1. the market price minus the congestion fee, and 2. an amount equal to the congestion fee. The former payment can be dealt with bilaterally. The second payment is equal to the amount due from other (non constrained-off) generators and can be invoiced through the surcharge on the G-tariff, which is then equal for all generators, so that a flat fee for all generators in the congestion area can be applied.

7 Conclusions and recommendations

Given the preference of Parliament that generators pay most or all of congestion costs, the locational incentives to give to generators (which will reduce future congestion), and the expected problems with gaming in the constraint market, the 'best' congestion management model for application in the Netherlands seems to be the **hybrid redispatch model**. It is therefore suggested to implement this model for solving congestion in the transmission network in the Netherlands.

Additionally, some further choices need to be made on several remaining issues like:

- the precise way in which the priority access for renewable energy is implemented in the redispatch system;
- whether (and possibly when) interconnections are included in (national) congestion management;
- which information requirements are given to market participants in order to be able to assess the congestion beforehand and monitor any deviations from approved flows;
- how to allocate the congestion revenues or costs to network users, and
- the precise impact of implementing congestion management on the design of the imbalance market.

All congestion management models analysed may suffer from possible abuse of market power. The system redispatch model is especially vulnerable to market power in the constrained-off market. All redispatch models have to face market power in the constrained-on market. Additionally, market and hybrid redispatch may suffer from 'optimization' strategies of generators owning much generation in the congested area. Several options to mitigate this market power have been discussed in the report.

Furthermore and given that the present study only involves a qualitative assessment, it may be considered worthwhile to perform a quantitative (regulatory) impact assessment of the hybrid model, as compared to system redispatch. Such an assessment would provide more insight into how much (generator) prices would change, which generators would benefit and lose and how consumers prices in the different areas would be affected.

Finally, since implementation of the hybrid redispatch model may cause delays in commissioning new generation in congested areas (which is a desirable locational incentive), the co-ordination problem between generation investments and network reinforcements may need to be addressed again.

Appendix I: Congestion in Great Britain

I.1 Introduction

The electricity market in Great Britain provides a particular relevant example for the Netherlands, because it is experiencing congestion problems now that the Netherlands expects in the near future. The transmission System Operator National Grid Electricity Transmission (NGET) has forecast constraint costs of £262 million for the year 2009/10, which is 10% higher than the current forecast for 2008/09 and almost four times higher than actual constraint costs in 2007/08.[44] As a result of these rising costs, in February 2009 the energy regulator, Ofgem, asked NGET to undertake a review of constraint costs.[45] The review will include options for reducing constraint costs, and a review of the way the costs of constraints are shared between consumers and generators.

The main location for constraints is the border between Scotland and England. Scotland is an area rich in wind resources, but this power must be transmitted to consumers in the south of Great Britain. Since the Scottish and English transmission grids were originally separate systems, the links between the two are not as strong as they might have been had there been an integrated system from the beginning. As a result, there are frequently constraints between Scotland and England. When the Scottish market was separate from the England & Wales market i.e. prior to April 2005, the constraint was managed through the use of explicit auctions of the interconnector capacity and so the effects of the constraint were only manifest in the different wholesale prices in Scotland and England & Wales. Since the introduction of the market ("BETTA"), there is a single wholesale energy price for Great Britain, which encourages plants in Scotland to generate at high levels. This often creates constraints that have to be resolved by the system operator. In the short to medium term, the expansion of the network to deal with the constraints will actually exacerbate the problems as existing circuits will need to be taken out of service to enable the reinforcements to be completed.

The need to increase the level of renewable generation in Great Britain is leading to proposals to develop a plethora of wind farms that are often located in areas where the grid is particularly weak. At present, the system operator manages this situation by delaying the connection dates for new plants until it has time to complete the grid reinforcements (both local and deep) it deems necessary. There are concerns that this system provides little encouragement to the system operator to accelerate its investments to enable early connections. Partly as a result, there are several proposals to allow generators to connect to the system before there is sufficient 'downstream' transmission capacity available i.e. before the deep reinforcement is completely. However, moves in this direction are likely to increase constraint costs unless they are accompanied by other changes to the trading arrangements. For example, NGET believes that 450 MW of Scottish renewable generation would like to advance their connection dates, which would increase congestion costs by £40 million per year.

There are a number of specific features of the situation in Great Britain which are relevant to the Netherlands, which we discuss below.

I.2 Basic congestion management

The principal means of dealing with congestion is by redispatching plant close to real time, in the case of Great Britain this is via the Balancing Mechanism. This is a rolling "market" where the system operator is the only counterparty and each session opens one hour before the start of the half-hourly balancing period to which it relates. The system operator uses the Balancing

---

[44] Ofgem letter to NGET, February 17th 2009 'Managing Constraints on the GB Transmission System'.

[45] *Ibid*.

Mechanism for all types of balancing activities e.g. reserve, reactive power, frequency response, not just congestion management. For each half hour, market participants (mostly generators but large demand side sites can also participate) can submit offers to be constrained on and bids to constrained down/off. Note that there is no requirement for market participants to submit bids and offers. If the system operator accepts their bids or offers, then they are paid the prices that they submitted.

The system operator can also into constraint contracts with particular generators where there is likely to be a long lasting constraint. The terms of these contracts are negotiated bilaterally, and so are not publicly available, but the idea is to agree a pricing structure for the bids and offers that the generator will submit to the Balancing Mechanism. In return, the generator receives either an availability fee (for agreeing to make submissions) or some other form of guaranteed revenue. Whilst such contracts are not common, they have occurred although the System Operator is often fearful that the generator has locational market power and hence may be able to extract unduly advantageous terms.

I.3 Information requirements

Generators are required to submit Physical Notifications to the system operator on a plant by plant basis – this is equivalent to TenneT's proposed T-programme. At least one hour before the start of the relevant half-hourly balancing period plants must submit a Final Physical Notification (FPN). It is on the basis of the FPNs that the system operator plans constraint management. Generators and customers also have to submit to the system operator their planned production or consumption on an aggregated basis – it is against these contractual nominations that imbalances are measured. For generators, the sum of the FPNs over their plant should equal their declared aggregated production.

Note that the FPNs are non-binding – generators can in theory deviate from their FPNs on a plant-by-plant basis, without financial penalty as long as their aggregate production remains as declared. However, generators have a license obligation to make 'accurate' FPNs. Therefore, continued deviation from FPNs could result in the system operator referring the generator to the regulator, which would investigate a possible breach of license condition. So far this system of 'ex post' regulation of FPNs seems to have worked – we are unaware of any gaming or complaints that have arisen as a result of deviations from FPNs.

I.4 Connection policy

As discussed above, in the past generators in Great Britain had wait until there was sufficient downstream transmission capacity before they could connect. This policy has recently changed as part of the 'interim connect and manage' proposals, under which generators would have the right to connect from a set number of years after they have signed a connection agreement, regardless of whether the wider works were complete. This means that the situation is becoming more like the Netherlands, where generators can connect even before sufficient transmission capacity is available.

I.5 Cost allocation

Constraint costs are recovered from all market participants (supply and demand), in proportion to their metered volumes via the Balancing Services Use of System (BSUoS) charges. In principle, the bids and offers accepted for constraint management are not supposed to influence the imbalance prices paid by parties who are out of balance i.e. whose metered volumes do not match their contracted volumes. In practice, imbalance prices are sometimes "polluted" by constraint costs because of the difficulty of developing rules that always manage to distinguish constraint actions from those taken for general system balancing reasons. For example, a plant that is constrained down may solve a constraint *and* bring the system back into energy balance.

As part of its constraints review, Ofgem has asked NGET to consider if the allocation of constraint costs is "equitable and efficient", the implication being that generators that cause constraints could bear more of the costs.

I.6 System operator incentives

In the four years after the privatisation of the electricity transmission system in 1990 transmission constraint costs, which the system operator passed on to customers, doubled in real terms. Ofgem realised that inefficient balancing actions were increasing constraint costs significantly, but the system operator had no direct incentive to reduce constraint costs. The system operator had some control over cumulative constraint costs through its capital expenditure, maintenance expenses and contracts for ancillary services with generators. However, NGC could make money by saving on capital expenditure and maintenance even if this inefficiently increased constraint costs.

The regulator introduced an incentive scheme that gave NGC rewards for reductions in uplift costs, and potential penalties for increases. As Table 5 shows, the scheme appears to have succeeded in reducing incentivised balancing and constraint costs, which fell by more than 20% over the first three years that the full version of the scheme was in place. Payments to NGC (£31 million) represent less than 30% of the savings achieved (£103 million), implying that the scheme benefited all parties concerned. It is difficult to be certain, however, what proportion of this fall in costs should be attributed to the effect of the incentives since constraint costs would likely have declined in any event, due to the closure of a number of small, old power stations connected to the distribution network, which accounted for a significant proportion of the system operator costs in the early 1990's.

Table 5. NGC's Incentivised Cost Savings (£ million, 1999 prices).

Cost category	*1996/97*	*1997/98*	*1998/99*
Costs incurred	503	472	400
Incentive payments	10	11	10
Costs recovered	513	482	410

Source: NGC Incentive Schemes from April 2000, Transmission Services Uplift and Reactive Power Uplift, A Decision Document, Offer *(25/02/2002)*

NGET continues to be incentivised to control the costs that it incurs as the system operator and its incentive scheme covers all costs that it controls – both external (e.g. in resolving transmission constraints, controlling frequency and voltage and contracting for reserve) and internal (i.e. the staff and systems it uses to fulfil its system operator function). Cost elements that are considered to be outside NGET's control are not incentivised but simply passed directly through to market participants. Figure 8 below illustrates NGET's incentive scheme for 2008/09. NGET keeps 25% of the savings if outturn costs are less than £529 million, but NGET pays 25% of costs above £544 million. A 'collar' and a 'cap' apply a floor and a ceiling to the gains and losses that NGET can make under the incentive scheme – in 2008/09 the gains and losses are capped at £15 million.

Figure 8. NGET's incentive scheme for system balancing and constraint management 2008/09.[46]

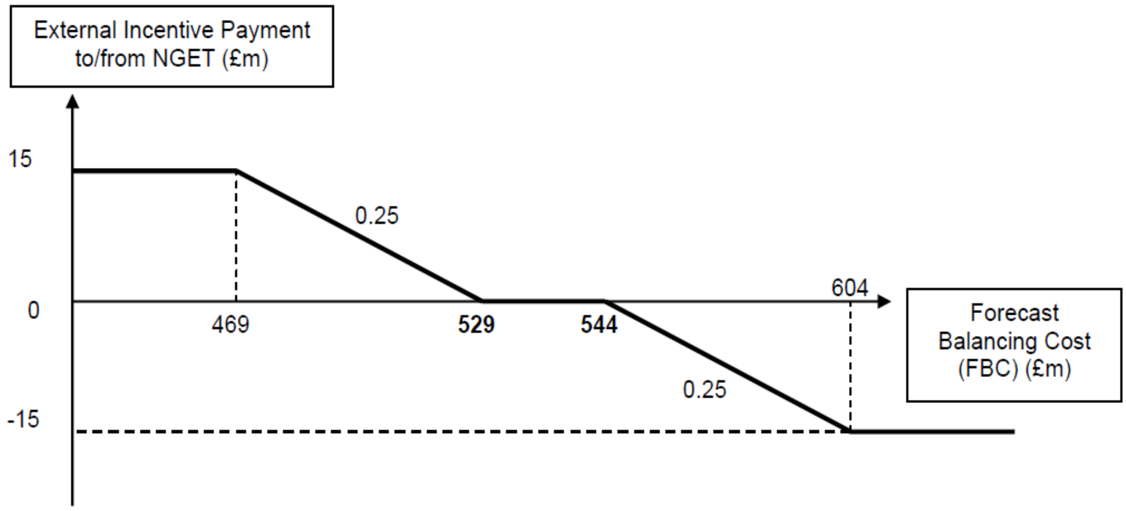


Table 6 below shows how constraint costs and system operator incentive payments have developed since the introduction of the BETTA. It is perhaps not surprising that the system operator and Ofgem had difficulty in forecasting the likely level of constraint costs for the first year when Scotland was included in the trading arrangements (2005/06) and, in recognition of this Ofgem made an ex post adjustment to the incentive scheme target. However, the large discrepancy in 2006/07[47] is more difficult to understand and led to a market power investigation, see below. One result of these cost under-estimates is that the system operator has made a loss in each year that there has been an incentive scheme.

Table 6. Constraint costs and incentives payments under BETTA (£m).

Year	Constraint costs			Incentive
	Forecast	Actual	Difference	payments
2005/06	36	84	48	-4.0
2006/07	57	108	51	N.A.*
2007/08	82	70	-12	-1.1
2008/09	124	194	70	-15.0
2009/10	307			

* There was no incentive scheme in 2005/06 - NGET
rejected Ofgem's proposals and Ofgem took no further action

I.7 Market power

In April 2008 Ofgem launched a competition act investigation into the behaviour of two Scottish generators, Scottish Power (SP) and Scottish and Southern Electric (SSE). The allegation was that these two generators had withheld plant from the forward wholesale market, so that they could use the same plant to supply NGET with power to manage constraints at excessive prices. On 19 January 2009 Ofgem decided to halt the investigation, since it judged that the likelihood

46 National Grid, The Statement of the Use of System Charging Methodology Effective from 1 April 2008, Issue 4, p.65.

47 The system operator attributes the predicted large under-estimate in 2008/09 to unexpectedly high wholesale electricity prices, due to very high coal and gas prices.

of finding the firms guilty was low, and it would be more effective to address the concerns by consulting on additional powers and bringing forward changes in the market rules to deal with the problem.[48]

This case illustrates two points. First, that market power is not simply a theoretical problem in the market for managing constraints. Second, that dealing with such problems using ex post competition law powers appears to be difficult.

I.8 Current capacity right products

The basic capacity right product is known as transmission entry capacity ("TEC"). A generator's TEC defines its maximum export capacity to the grid, and is basis on which its transmission charges are calculated. The TEC can be less than the maximum capacity of the generator's connection assets (the connection entry capacity or "CEC"). Once connected, a generator can keep its TEC for as long as it desires, subject only to an obligation to pay its transmission charges. TEC is a financially firm product: once released the system operator can only retrieve it by buying it back. Generators can hand back some or all of their TEC at any time[49] and their obligation to pay transmission charges is adjusted accordingly from the next year onwards. Once TEC has been surrendered, there is no guarantee that it will be made available again. In addition to this basic product, the system operator offers a number of other capacity products, subject to their release not causing or exacerbating any constraints, which provide generators with greater flexibility whilst still allowing the system operator to manage constraints.

Proposals relating to long-term capacity options

There are two competing proposals for the introduction of defined long-term capacity rights. Both ideas are intended to provide greater certainty regarding where transmission investment is required by committing generators to pay for capacity rights for longer periods than one year. From the perspective of congestion management, the proposal to introduce long-term auctions of capacity rights, akin to those already held for gas entry rights, is of more interest.[50]

Using auctions to assign capacity would ensure that it was allocated to the users who valued it most. Auctions would also provide useful advance signals about where investment is required and would potentially allow users to make a choice between fewer but cheaper access rights and more but expensive access rights. Moreover, by varying the volume of rights made available for the auction between years, the system operator could use the auctions as a means of controlling temporary congestion problems.

The proposals envisage that there would be reserve prices for the auction that are based on current charges for entry capacity, to ensure that users cannot get a "free ride" on the system in areas where there is no competition for capacity. Successful bidders would have to pay for the capacity they acquired at their bid price for duration that they have bought the rights. In order to prevent hoarding, users would only be able to bid for access rights up to the capacity of their local connection.

---

[48] Ofgem Press Release 19th January 2009, Ofgem closes competition act 1998 case against Scottish Power and Scottish Southern and Electric.

[49] They only have to provide five days' notice.

[50] The other proposal retains the concept of transmission charges based on long run marginal costs but requires (post-commissioning) generators to specify the number of years for which they require rights, for which they will have an obligation to pay transmission charges. Consequently, it does not provide the system operator with the possibility to vary the volume of rights released for each year as a way of controlling congestion management.

I.9 Conclusions

Great Britain has a congestion problem which is getting worse due largely to the growth in renewables, and compounded by alleged market power abuse in the constraint market. Financial incentives on their own do not enable the system operator to control congestion costs through redispatch.

The regulator is currently looking at ways of changing the rules governing congestion management, in an effort to reduce costs. Possibilities include both limiting the volume of constraints for which generators are compensated, and limiting the prices paid for constrained on and off plant.

Appendix II: Congestion in other large EU electricity markets

II.1 Italy

Zonal market

The principal method of dealing with congestion in Italy is by splitting the market into six separate price zones, the boundaries of which coincide with the main transmission constraints. As with other market splitting mechanisms if the transmission between zones is constrained each zonal market sets a different equilibrium price applied to all offers in that zone. The price paid by demand is always set at the national level, and is equal to the weighted average of the zonal prices.

Must-run units

In order to mitigate the risk of market power abuse, The Italian system operator, Terna, creates every year a list of thermoelectric plants considered essential for the despatching. These plants are forced to offer at a price equal to zero on the day-ahead or MGP market, as they would have too much market power in the constraint market. These plants are paid at the prevailing day-ahead price, but they can receive a further mark-up if they can prove that the market price is not sufficient to recover their costs. This mark-up is calculated as the difference (if positive) between the day-ahead price and the estimated cost-recovery value defined by the regulator for each plant. Terna recovers these extra costs with a specific charge applied to all consumers.

Connection policy

In Italy the system operator has the legal obligation to connect any authorised generator to the grid. There is not any formal mechanism allowing the system operator to postpone the connection in order to avoid congestion.

Intra-zonal congestion management

The principal means of dealing with intra-zone congestion is by redispatching plant. Some qualified plants have to bid on the despatching market (mercato Servizi Dispacciamento – MSD) to be constrained on and constrained off, and if called for these services they are paid their bid price. The criteria for the qualification are defined by Terna and aim to include all plants that have the technical capability to supply reserve services. Since a modification of the regulation in 2006[51], plants are only paid for the electricity and not for the reserved capacity.

Information requirements

After the day-ahead market and again after the MA (Adjustment Market), the Market Operator provides the system operator with the plan of the production for the following day, on a unit by unit basis. The Market Operator has to submit the schedule of the MA at the latest with the opening of the constraint and balancing market, which is before 2.30pm of the day before the production.

All plants that fulfil determined technical requirements have to participate in the constraint and balancing market, and Terna calculates the net position of the operators at the end of each market session. Unbalances higher than 3% of the planned capacity are fined through lower levels of prices to sell and higher levels of prices to buy, compared to the day-ahead price.

Cost allocation

Constraint costs are recovered through an “uplift” charge calculated quarterly (since 2007) by Terna and applied to the consumers. This charge covers all costs of the constraint and balancing

[51] Delibera 165/06.

market which are not recovered by the imbalance charge paid by unbalancing operators. This allows Terna to recover ex-post all costs and reduces the incentives to minimise the constraint and balancing market transactions.

Insufficient transparency

The main problem of the actual structure of the MSD is the insufficient information on the causes of Terna's costs. Indeed, the lack of monitoring on Terna's utilisation of different resources and the lack of transparency on the economic value of each resource, prevent a clear indication of the nature of the congestion/balancing problems. Costs could be due to systematic network congestion, low flexibility of the generation mix or continuous imbalances by the supply and demand units; understand the source of these costs could help improve the system's efficiency[52]. The nature of the market and the lack of transparency make surveillance of the balancing and constraint market difficult.

Recent and future reforms

The increasing and volatile trend of the "uplift" charge during the last years has convinced the relevant authorities to introduce reforms. Congestion costs are very high – while in is hard to get reliable estimates, one trade journal report estimated costs as over €1.3 billion per year.[53] Possible changes indicated by the recent discussion are the followings:

- Incentives for Terna to develop specific parts of the network. A recent legislation introduced incentives for Terna to develop the network to solve capacity constraint, by allowing a higher WACC on those investments compared to other investments of network expansion (9.9% instead of 8.9%). However the legislation leaves to Terna the decision of the priority of the investments.
- Creation of a separate market for solving congestion. This would increase transparency by treating congestion costs separately from other balancing costs. This would provide an implicit incentive for the system operator to reduce constraint costs, perhaps by investing more in the network. While these reforms would not solve market power, they would facilitate the detection of abusive practices easier.

II.2 France

The main congestion management mechanism in France is redispatching plants. Plants have to offer the residual capacity not engaged in the day-ahead market to the RTE, through a specific adjustment market. Here traders use a system of upward and downward offers, and indicate the technical and financial conditions under which the RTE may change their generation or consumption programmes. It is open to French generators, to major interruptible consumers and to foreign operators, while RTE is the only counterparty to the adjustment offers on this market. Market participants are paid for their bid price.

The network constraint is not a major problem in France. The total cost of the system operator (RTE) for the domestic congestion management was €47.0 mln in 2007 and only €34.3 mln in 2008[54]. In 2008 92.3% of this cost was due to congestions arising in only two areas: West and South East.

[52] Ref. Ricerche e consulenze per l'economia e la finanza, Quaderno n. 47 Giugno 2008, "Le criticità del mercato elettrico attuale e gli effetti delle congestioni di rete".

[53] Article Quotidiano Energia 18/03/2009 '*Mercato elettrico, ora l'MSE tira le fila*'. The article estimates that about half of the system operators €2.7 billion despatching costs are due to congestion.

[54] RTE, "Bilan du Méchanisme d'Ajustement", January 2008 and January 2009.

At first glance its seems surprising that, given the dominant position of the incumbent in the French electricity market, congestion costs are not higher. The most likely explanation of the relatively low costs is the French connection policy. Generators in France have to queue before being connected to the grid. In particular they have to wait until there is sufficient transmission capacity in the area where they want to connect before they receive the authorisation to proceed. This is most likely the reason congestion costs in France are relatively low.

II.3 Spain

In common with the countries above, Spain manages its congestion using a system of redispatch and counter-trading. The Spanish system operator cannot refuse connection to generators on the grounds of insufficient transmission capacity, since the relevant law states that such constraints should be dealt with via redispatch.[55]

The Spanish constraint market is interesting for two reasons. First, it is one of the few EU constraint markets we are aware of that applied administrated prices to constrained on prices, presumably in a bid to contain market power. Pre-July 2005 constrained on prices were set as the greater of the average of the extra-marginal offers (rejected offers) in the Spanish electricity pool, or 15% above the day-ahead price. This provides one example of how the MO can set pricing rules to mitigate market power in the constraint market.

The second reason Spain is interesting is because, since it reformed the constraint market rules in July 2005, it has experienced many instances of market power abuse. More unusually, Spain has been successful in prosecuting the firms involved and fining them. For instance, on 26th of April 2008 the CNC imposed a €1.5 million fine on Gas Natural for abuse of dominant position in the electric energy market under technical restrictions (the constraint market). Because of transmission constraints, GN's plant was required to supply a given demand – other plants could not serve all the load because of transmission constraints. GN had offered energy at the daily market at artificially high prices so that the plant would not be scheduled to run, creating a 'constraint', which GN was in a unique position to solve.

Similarly, the former Tribunal for Defense of Competition and the CNC had previously sanctioned other electric companies for abuse of dominant position within the electric energy market under situations of technical restrictions.[56]

II.4 Germany

In common with the other countries studies, German system operators apply a redispatch/counter-trading scheme. Bids and offers are supposed to be cost reflective. Participants in the workshops indicated to us that market power was controlled by auditing of generators, to check if bids were in fact cost reflective. We have heard of similar schemes (though not in connection with congestion management specifically) in other markets. However, further inquires to parties active in the German market could not confirm the existence of such auditing practices. The German system operators do not publish data on congestion costs, and so it is not possible to benchmark the German costs against other systems where redispatch is used.

Connection policy in Germany was changed in only 2007. Before this date, power plants could only be connected to the transmission grid if the network capacity was in place. If grid reinforcement works were necessary, these would have to take place before the station's

55 See Article 52 of Royal Decree 1955/2000.

56 See Resolution 601/05 Iberdrola Castellón, of 8th of March 2007; Resolution 602/05 Viesgo Generacion, of 28th of December 2006 and Resolution 624/07 Iberdrola of 14th of February 2008.

connection to the grid. One can imagine that, in the presence of vertical integration, the network companies had strong incentives to delay investments in the required deep reinforcements, to delay or prevent entry of plant that would compete with their generating affiliate. Since the 2007 law, power plants should have been connected to the grid immediately, even if the required reinforcements are not yet in place.

Appendix III: Congestion costs in other EU countries

In Table 7 we report the cost of congestion in Great Britain, Spain, Italy and France. So as to make the numbers more comparable to the Netherlands, in Table 8 we have adjusted the congestion costs so that they are proportional to the size of the Dutch electricity market. For example we multiply French congestion costs by 0.247, because France consumes about four times as much electricity as the Netherlands.

Table 7. Costs of congestion in Great Britain, France, Italy and Spain

	GB [A]	France [B]	Italy [C]	Spain [D]
2006	118.8			
2007	77	47		337
2008	261.8	34.3	1300	434
2009	288.2			

[A]: Managing constraints on the GB Transmission System, 17 February 2009, letter from Ofgem to National Grid. Converted to Euros at 1.1 Euros/Pound. Figures for
[B]: RTE, "Bilan du Méchanisme d'Ajustement", January 2008 and January 2009.
[C]: Article Quotidiano Energia 18/03/2009 'Mercato elettrico, ora l'MSE tira le fila'. The article estimates that about half of the TSOs €2.7 billion despatching costs are due to congestion.
[D]: Derived from data published by Red Electrica

Table 8. Costs of congestion in Great Britain, France, Italy and Spain scaled to the size of the Dutch electricity market

		UK [A]	France [B]	Italy [C]	Spain [D]	Netherlands [E]
Electricity consumption 2006, 1000 toe	[1]	29,474	36,968	26,509	21,484	9,117
Ratio to Dutch consumption	[2]	30.9%	24.7%	34.4%	42.4%	
Pro-rata congestion costs, € mln						
2006	[3]	37				
2007	[4]	24	12		143	
2008	[5]	81	8	447	184	
2009	[6]	89				

Notes:
[1]: Eurostat
[2]: [E][1]/[1]
[3]-[6]: The unadjusted cost multiplied by [2].

The purpose of this exercise is not to forecast the cost of constraints in the Netherlands. Dutch constraint costs will depend on many details including network topology, fuel costs and which plants ultimately connect. However, looking at the (scaled) constraint costs of other countries provides a useful reality check on TenneT's approximate forecasts of Dutch congestion costs. In this case, international experience indicates that TenneT's estimates of congestion costs of the order of €100 million per year from 2012 onward are consistent with other countries that have regular congestion. The costs of Italy are clearly exceptional, though the data should be treated with caution since there is considerably uncertainty about the true nature of congestion costs in the Italian market.

Appendix IV: Effect of transmission capacity on prices and generating costs

In the situation where there is no transport capacity between A and B, market A and market B will clear separately, resulting in the following outcome:

REGION A		**No transport capacity**	**REGION B**	
Effective demand	1500 MW	*effective transport:* 0 MW	Effective demand	2250 MW
Marginal generator	G2		Marginal generator	G6
Market price	**30 €/MWh**		**Market price**	**45 €/MWh**
Generation costs in A	€ 35.000	Total for A and B: **€ 106.250**	Generation costs in B	€ 71.250

As can be observed, region B is more expensive than region A. When transport capacity would be available, some expensive generation in B could be exchanged for cheaper generation in A (which would then need transport from A to B).

In the situation of an integrated market, we assume infinite transport capacity between A and B. Market A and B will now clear with a single price, resulting in the following outcome:

REGION A		**Infinite transport capacity**	**REGION B**	
Effective demand	2000 MW	*effective transport:* *no limitations*	Effective demand	1750 MW
Marginal generator	G5		Marginal generator	G5
Market price	**35 €/MWh**		**Market price**	**35 €/MWh**
Generation costs in A	€ 50.000	Total for A and B: **€ 101.250**	Generation costs in B	€ 51.250

In this situation, first G1 is dispatched to its full capacity (1000 MW), begin the cheapest plant; then G4 (1000 MW), then G2 (1000 MW) and finally G5 for the remaining 750 MW. G5 is therefore the marginal plant, setting the market price at 35 €/MWh for both regions.

The total generation costs are apparently lower than in the previous case, since more expensive generation in B (G6 and G5) has been exchanged for cheaper generation in A (G2 and G3).

Appendix V: Description of an alternative hybrid redispatch model

V.1 Description of the model

In this appendix, a variant is given of the hybrid redispatch model ('variant B'). The difference with the model presented in the main text ('variant A', see section 4.6) is that the generation surplus of the generators in the constrained-off market is now left with these generators. This provides an incentive not to bid at 0 €/MWh for large generators in the constrained zone (see section 6.2.4 for a discussion).

The model works as follows: Generators in the congestion area (region A) bid into a congestion auction market where the bids represent their bid for the system operator to take over delivery (i.e. representing their marginal costs). When congestion occurs, the system operator requests the generators representing the highest bids first (i.e. the most expensive plant) to reduce their output and pay their bid price to the system operator ('pay as bid'). The difference between the APX-price and the lowest bid accepted by the system operator for a specific hour sets the congestion price for that hour. For each MW of remaining generation delivered to the grid (which has not been reduced by the system operator), the congestion price must be paid (by generators in A).

V.2 Illustration by means of the reference model

When placing the bids in the congestion auction, each generator in region A estimates its costs for reducing generation. These bids represent the marginal costs of each generator. G1 will therefore bid in at 20 €/MWh and G2 at 10 MWh. G3 does not place a bid in this example.

After collecting the bids on a bidding ladder, the system operator will conclude that G2 represents the most expensive plant and therefore provides the 'cheapest' output reduction. The system operator therefore assigns the constrained-off capacity to the amount of 250 MW to G2, who will consequently reduce its output from 1000 MW to 750 MW.

In region B, the system operator also organizes a market (which must be an hourly market). When all generators in B are obliged to submit bids to the system operator for their non-used generation capacity, the system operator will be able to call the generators with the lowest marginal costs first. In the example, G4 already runs at its maximum output. G5 may bid its remaining 250 MW at a marginal costs of 35 €/MWh and G6 will bid 1000 MW at 45 €/MWh. The system operator will now contract 250 MW with G5 (the cheapest plant in B) to compensate for the reduction of 250 MW in zone A (which was needed to solve the congestion).

For solving a congestion of 250 MW, G2 will therefore be called to ramp down and G5 will be called to ramp up. The costs for G2 are 250 MW × 30 €/MWh (its bid price) = € 7500, which are being paid by G2 to the system operator. For this, the system operator takes over the supply obligation from G2 to customers in region B. The 'loss' for G2 is however its surplus of 5 €/MWh.

Given that the market price is 35 €/MWh (which is the clearing price of the APX) and that the lowest accepted congestion bid was 30 €/MWh, the congestion price is set at (35 minus 30 =) 5 €/MWh for this hour. This congestion fee is paid by all remaining generation. Therefore G1 and G3 (who were not selected) as well as G2 (for the 750 MW at which it will continue to run) must now pay a congestion fee to the system operator of 1000 MW × 5 €/MWh = € 5000 for G1, 750 MW × 5 €/MWh = € 3750 for G2 and 0 MW × 5 €/MWh = € 0 for G3, which amounts to € 8750.

For constrained on power in zone B, the system operator contracts 250 MW at 35 €/MWh with G5, which (in the example) equals the market price. The total costs for this are € 8750. Now the total balance for the system operator is the following:

Region A:	Income for taking over supply obligation from G2 (250 MW at marginal cost)	+€ 7500
	Income due to congestion fee (1750 MW at congestion price)	+€ 8750
Region B:	Costs of constrained-on generation (250 MW at extra-marginal price in B)	–€ 8750
Net congestion revenue for system operator		**+€ 7500**

Just as in variant A of the hybrid redispatch model, application of the hybrid redispatch model lowers the net income of running generators in region A by 5 €/MWh. However, the constrained-off generators are allowed to keep their surplus (since they only pay 30 €/MWh for being constrained-off but are credited with the supply to customers in B at the market price). For this reason, the net congestion revenue for the system operator is a bit lower (€ 7500 instead of € 8750 for the hybrid model presented in section 4.6).

V.3 Assessment of this model

The present variant of the hybrid redispatch model ('variant B') has overall the same qualities as variant A. The major difference is that the constrained-off generators are allowed to keep their surplus, which will decrease the likelihood of strategic 0 €/MWh bids.

A disadvantage is that the height of the congestion fee to be paid by the generators in the congestion area is now related to the APX price. In variant A, where the congestion fee is only derived from the bids of the generators, the APX price is only being paid by the constrained-off generator to the system operator (although the height of the bids in the congestion auction will be related to the APX price as well). In both schemes, the height of the APX price for a given hour will influence the height of the congestion fee. When the congestion auction is held on a longer-term basis (e.g. monthly), the resulting congestion price may end up rather high or even negative for specific hours. For the hours when the APX price peaks, especially variant B may induce large congestion payments, which are undesirable. Application of this model therefore implies the congestion auction to be held hourly.

Appendix VI: Specific comments by market parties

In this appendix, some arguments brought forward by market participants will be discussed.

I. Incentives for network companies are lacking

An argument brought forward by (predominantly) generators is the importance of providing sufficient investment incentives for network companies. Since congestion results from a lack of available transmission capacity whereas the legislation requires network companies to invest sufficiently, it has been suggested to charge the costs of congestion management to the network companies.

In principle, this argument makes sense. It is the responsibility of network companies to sufficiently invest in network capacity. Nevertheless, it is difficult to use the same 'bill' (consisting of the costs of congestion management) for incentivizing both generators (for efficient network use and locational incentives since too much generation 'causes' congestion) and network companies (for timely network reinforcements since insufficient network capacity 'causes' congestion). In the present study, the focus lies on incentives for generators. The main arguments are the following:

1. The existence of the coordination problem is not new. Given the unbundling of generation and transmission, the timing issue of network investments yields a dilemma which has not yet been solved. Within the Dutch regulatory context, it may be even a bit worse, since network companies who invest in transmission lines may be considered 'inefficient' (as a consequence of the regulatory scheme) when the lines are not utilized (which occurs when generators eventually do not build plant near the transmission line). On the other hand, once a generator has decided to invest, a network company will almost by definition be too late since network reinforcements generally demand more time than generators.
2. Secondly, congestion revenues do not seem to be a very good indication of the need for network investments. The amount of congestion revenues does not necessarily indicate that investments in new lines are a good idea since the expected duration of the congestion and the costs of the network reinforcements need to be balanced against the social benefits of higher transfer capacity. Therefore, although the argument may be appealing at first glance, it is quite difficult to give correct investment incentives to network operators through a congestion management system.
3. Finally, since a balance must be found between incentives for network companies and incentives for generators, one could argue that in the Netherlands, there is in general a lack of incentives related to the siting of generators. On the other hand, the present regulatory approach toward network operators already includes many constraints and incentives related to network investments.

An incentive is only required when a firm has a choice of options. In the case of the Dutch transmission network, TenneT has little choice but to invest to solve congestion. It is obvious where congestion is or will be, and equally obvious if TenneT is addressing the problem or not. Given regulatory pressure, it is not a realistic option for TenneT to ignore prospective transmission constraints once they are identified. Even absent any financial incentives, TenneT has already made plans to solve the congestion issues identified. Congestion in the Dutch transmission network therefore does not seem to arise because of a lack of financial incentives for TenneT, but rather because of the timing and regulatory issues we identified above.

II. Congestion management should not be introduced before its costs are clear

It is argued (mostly by generators) that as long as a detailed assessment of the costs associated with congestion management is lacking, congestion management should not be implemented.

To some extent, this is a fair point, especially since it is likely that generators will partly pay the congestion costs. Nevertheless, a few arguments can be brought forward against postponing the implementation of congestion management until the exact costs are known. The major one is that a system of congestion management is required regardless of the costs, and that our analysis makes a qualitative assessment of the difference in costs involved. For example, system redispatch scores badly on some parts of the evaluation because the costs, due to market power, could be very high.

In the longer term the introduction of congestion management itself may relieve congestion, at least when the system gives locational incentives (as the hybrid and market redispatch models do). Given that future congestion costs are paid by generators in the constrained area, additional generation will receive a financial incentive to move elsewhere, or the commissioning might be delayed. Both effect will help to prevent (or at least reduce) further future congestion, which is good thing in itself. Even in the absence of an accurate estimate, the sooner locational incentives are given, the less future congestion will occur so the lower will be the congestion costs.

III. Socializing congestion costs is most in line with the present regulatory approach

Since congestion results from a lack of investments by network companies and network costs are socialized through the tariffs (mainly to consumers), it is argued that congestion costs need to be socialized as well (instead of being partly charged to generators).

Although it is correct that a network operator may better optimize for operational measures (as system redispatch) or investments for solving congestion when it will bear all these costs, a case has been made already for not charging too much congestion costs to network operators (see section **Fout! Verwijzingsbron niet gevonden.**). All costs not absorbed by the network operator itself must however be allocated to specific groups of network users or be included in the network tariffs (which are paid mostly by consumers).

The present network tariffs are basically derived from Ramsey pricing. However, given the need to provide locational incentives, other tariff approaches deserve implementation (at least for the costs related to congestion management). Since charging generators will provide locational incentives (which may reduce future congestion and therefore limit future congestion costs), from an economic perspective it makes sense to allocate at least the constrained-off costs to generators in the congested area.

IV. The introduction of congestion management is not fair

The introduction of congestion management is not considered fair. The argument is that the generators' profits in a congestion zone will go down by the introduction of a hybrid or market based redispatch scheme since the generators end up paying the constrained-off costs or a congestion fee. It is argued that the introduction of such a scheme could not be foreseen making it unreasonable to implement it today.

On the one hand, the argument does not make sense. In any market, prices (and profits) will go down when new producers enter the market. This is simply a matter-of-fact consequence of competition. Economic theory predicts that as long as short-term prices (based on the variable costs) are above long-term prices (based on average costs), a market will attract new entrants. In this respect, the electricity market is not different from other commodity markets. A major difference however is that the power industry has large sunk costs. Once a generator is sited, moving to another market area is not possible (unless at extremely high cost).

On the other hand, generators may argue that although this is true for the Dutch market as a whole, the fact that network congestion would isolate specific congestion zones with an effectively lower price level could not be foreseen. This argument is more complex to analyse. Indeed, in the past it was assumed that the Dutch grid would behave as a 'copper plate'. Nevertheless, in many other countries it has already been clear for many years that there are transport limitations to any electricity transmission network. The basic question is then whether it was reasonable to suppose that society would pay for managing this scarcity (e.g. by a system redispatch mechanism where the costs are socialized or by leaving the cost with the network operator) or whether it could be expected that the bill would eventually end up with the generators, who make the siting decisions for their new plant.

V. Implementing a congestion management system increases regulatory risk

A related argument concerns the increased regulatory risk for generators, which may deter further investments. Given the 'new' congestion management scheme which incurs additional costs to generators, new investors may be deterred. After all, who knows whether any additional market design changes will be implemented?

This criticism seems fair – implementing a congestion management mechanism that changes generators' revenues *ex post* will increase regulatory risk. The more relevant question is whether the increase in regulatory risk is proportional to the size of the problem involved. We argue that congestion has been identified as a major issue, that could cost several million euros per year, and that a policy to address congestion and ideally reduce its costs is required.

There do not seem to be many alternatives to implementing market-based congestion management mechanisms which provide locational signals for new plant to site outside a congestion area. If such market-based mechanisms are not applied, new investments will only lead to more congestion, which then needs to be solved by the system operator at society's cost. Only when the expected profits for new plant in a congestion area are lowered (which is effectively done by a market-based mechanism), a locational signal is sent to generators to find a more attractive location for their generator (i.e. near demand or at a location where demand can be supplied without passing congested lines).

We also note that, given the obligation to supply sufficient network capacity, any congestion is of a temporary nature (although it may take several years to construct new lines). However, the chicken-egg dilemma shows up here again: Should the announcement of new plant in a congestion area be postponed until the transmission lines have been reinforced (leaving the risk with the network operators, who will not know whether new generation will connect or not), or should a generator bear this risk and invest (taking congestion for granted, which will reduce its profits during the first years). Apparently, there is no easy solution to this dilemma. It is therefore advised to reconsider the relation between the present connection policy and transmission expansion planning.

VI. The Netherlands is too small for locational signals

Another argument forwarded against the introduction of a congestion management system is that the Netherlands is considered too small for locational signals. This argument begs the question: too small compared to what? For example, Singapore (which is considerably smaller than the Netherlands) has a complex system of locational signals for generators via its *Locational Marginal Pricing System*.

A related question might be whether locational signals are likely to be efficient or effective. In its 2007 report for TenneT *The Brattle Group* recommended against introducing a locational signal via a G-charge.[57] The main reasons for this were that a G-charge in the Netherlands could deter investment, could create winners and losers and could be unstable and therefore ineffective. However, the report did recommend an alternative locational signal, which was to make generators wait until there was sufficient transmission capacity that they could be connected without causing congestion.

Since 2007, several factors have changed. First the decision was taken to connect all generators without waiting for network reinforcements, removing the locational signal that would be generated by having to wait for a connection. Second the volume of generation waiting to connect has continued to grow. Finally, the Dutch parliament has expressed a wish for congestion costs to be targeted at generators. If congestion costs are targeted at generators, it makes sense to target generators causing congestion.

The main aim of a congestion management method is to manage congestion and control congestion costs for consumers. The locational signal is a by-product of the congestion management system, rather than a goal in itself. Some of *The Brattle Group*'s original concerns still apply. A congestion management may create winners and losers, which could encourage opposition. Again, given the size of the congestion issue and the need to control costs for consumers the introduction of a congestion management system seems inescapable. The locational signal could discourage investment in the Netherlands. But given the length of the connection queue this is currently less of a concern than it was.

Similarly, in 2007 *The Brattle Group* recommended the Ministry of Economic Affairs not to introduce a system of locational signals via G-charges, for similar reasons to those given in the TenneT report.[58] Again, it does not make sense to introduce a system of locational signals for its own sake. However, given the need to manage congestion, it makes sense to target the congestion costs in a way that will minimise future congestion.

Given the current connection policy and the resulting congestion, the locational signal which results from the congestion mechanism will at least result in a fair allocation of cost. Since the locational signal is a by-product of the congestion management mechanism, it is less of a concern if it is not 100 % effective. Rather we recommend a congestion management system that will result in efficient dispatch, avoid market power as far as possible and allocate costs to the people causing them – which moreover gives a locational signal.

VII. It is not fair that generators pay for congestion caused by reduced demand

At some locations, e.g. in Zeeland, congestion is partly the consequence of a significant demand reduction (apparently related to the present economic crisis). As a consequence, the excess generation (total generation minus consumption) in the region increases, which causes congestion on the lines transporting the power from Zeeland to the Randstad. It is not perceived to be fair that generators pay for this load-driven residual network connection.

The arguments mentioned under item IV. above apply here as well. On the one hand, it is an intrinsic market risk that demand may turn out to be lower than expected. On the other hand, it is not clear whether the occurrence of regional congestion causing a local mismatch between generation and supply could be reasonably foreseen by the generators.

[57] The Brattle Group, *op.cit.* footnote 10.

[58] The Brattle Group, Locational signals in North-West Europe, November 2007.

VIII. Implementation of a market or hybrid redispatch model is time consuming

A further argument is that the introduction of market or hybrid redispatch model may be more time consuming than a system redispatch model.

Technically, this argument is not true. In case the constrained-off auction of a hybrid redispatch model would be organized as an hourly auction, the procedure would be completely similar to the procedure under a system-redispatch model. The same IT systems are needed and the same information requirements apply. The only difference between a hybrid model and a system redispatch model would be the financial settlement.[59] Technically speaking, the implementation time of all redispatch models is therefore more or less similar.

Although the bids and procedures will be highly similar under all models (strategic bidding not taken into account), there may be some longer-term effects on the contractual position of generators (related to the economic incentives given, which are therefore desirable effects). It may indeed take some time to adapt to the new market situation.

IX. Pass-through of congestion costs by regional network operators is subject to regulatory incentives

The network companies have raised objections against the pass-through of congestion costs through the network tariffs. Since several generators have connections with networks operated by regional network operators, the congestion costs may be charged by TenneT indirectly, i.e. through the regional network operators. The latter are however subject to an incentive regulation scheme. Therefore, the passed-on congestion costs to charge to the generators may be differently adjusted for the regional network operators (although the sum will remain unchanged). The consequence is that some of the regional network companies may have to absorb some additional costs, whereas others gain some additional profits.

Therefore, when a policy decision is being made that the congestion costs are charged to generators, it makes sense to leave these costs outside the regulatory incentive regime, since these costs cannot reasonably be influenced by the regional network companies.

59 Under system redispatch, the system operator pays the constrained-off and constrained-on generators. Under hybrid redispatch (variant A), all generators in the congestion area pay the system operator whereas the system operator pays the constrained-on generators. Under market redispatch, only the non constrained-off generators pay the system operator.