Electricity
Market design and policy choices

SPM9541

2013-2014

dr. ir. L. (Laurens) J. de Vries
dr. A. (Aad) F. Correljé
mr. drs. H. (Hamilcar) P.A. Knops
# Table of Contents

1 Introduction  
2 Background  
   2.1 Policy principles  
   2.2 Motives for restructuring  
   2.3 System description: electricity  
   2.4 System description: natural gas  
3 Europe’s energy markets  
   3.1 EU directives for the electricity and gas sectors  
   3.2 The European Union’s Third Energy Package  
   3.3 Dutch electricity and gas legislation  
   3.4 The current situation  
   3.5 Challenges  
      3.5.1 Market integration  
      3.5.2 Competition issues  
      3.5.3 Investment and innovation  
   3.6 The role of market design  
4 Market Design  
   4.1 Introduction  
   4.2 Context of the restructuring process  
   4.3 Market design  
   4.4 The dynamics of restructuring  
   4.5 Hybrid markets and patterns of restructuring in Europe  
   4.6 Analysis  
   4.7 Conclusions  
5 Distribution networks in the Netherlands  
   5.1 Introduction  
   5.2 Network regulation  
   5.3 Distribution network unbundling  
6 Congestion Management  
   6.1 Introduction  
   6.2 Selected congestion management methods  
      6.2.1 Explicit auctions  
      6.2.2 Market splitting  
      6.2.3 Redispachting  
   6.3 Analysis  
      6.3.1 Economic efficiency and income effects  
      6.3.2 Long-term effects  
      6.3.3 Barriers to trade  
      6.3.4 Congestion in a network  
7 Security of supply of electricity  
   7.1 Introduction  
   7.2 Why a capacity mechanism?  
      7.2.1 The narrow investment optimum  
      7.2.2 Market Failure  
      7.2.3 Reserve capacity as a public good  
      7.2.4 Investment cycles  
      7.2.5 Long-term contracts
7.2.6 Market power 94
7.2.7 Trade between electricity systems 95
7.2.8 Risk asymmetry 96
7.3 Performance criteria 97
7.4 Capacity mechanisms 99
  7.4.1 Capacity payments 99
  7.4.2 Strategic reserve 99
  7.4.3 Operating reserves pricing 100
  7.4.4 Capacity requirements 101
  7.4.5 Reliability contracts 103
  7.4.6 Bilateral reliability contracts 105
  7.4.7 Capacity subscriptions 106
  7.4.8 Conclusions 107
7.5 Policy framework 107
  7.5.1 Implementation as a precaution? 107
  7.5.2 Unilateral or regional implementation? 108
  7.5.3 Self-reliance? 109
  7.5.4 Innovativeness 110
  7.5.5 Short-term versus long-term options 110
  7.5.6 Overview of the policy choices 111
  7.5.7 Implementation issues 112
  7.5.8 Assignment of responsibilities 113
7.6 Conclusions 115
Literature 119
Annex: The design of the Dutch gas sector 125
  A.1 Goals and functions 125
  A.2 Production and imports 126
  A.3 Transport network management and system management 130
  A.4 Flexibility in supply and demand 132
1 Introduction

While the plans to liberalize the European energy markets date back to the early 1990s, liberalization continues to draw attention, perhaps even more now than before. The reason twofold. First, while the energy industries may account for only a small part of our Gross Domestic Product (GDP), the products they provide are essential to the functioning of our entire society. Thus the economic and social impact of service disruptions is great. Second, energy markets are based upon physical energy infrastructures with peculiar characteristics, as a result of which energy markets are different from other markets. They require their own, tailored market design. The goal of this class is to learn how electricity and gas markets are designed. As the design of our markets is not completed, the second goal of this class is to learn about the long-term issues that still need to be confronted in these markets.

This class focuses on European markets, specifically the Dutch markets. The physical, economic and juridical aspects of the electricity and gas sectors are described. Based upon this description, an analysis is made of the main market design issues. This class will not deal with related issues, such as the security of supply of primary energy sources other than gas (the Annex provides an introduction to gas market design) and environmental issues such as climate change. These issues are more political in nature, while the design of energy infrastructures is characterized by a high level of technical, economic and juridical complexity.

This reader starts in Chapter 2 with a conceptual framework which is we use for understanding the organization of electricity and gas markets. Chapter 3 continues with a description of current European energy markets. The remainder of this reader focuses on electricity, with the exception of the Annex, in which an analysis is provided of the organization of the Dutch gas sector.

Chapter 4 puts forward a theory of how market design changes over time. The last three chapters provide analyses of practical issues: the regulation of distribution networks in the Netherlands (which is closely in line with theoretical economic prescriptions) in Chapter 5, methods for managing congestion in Chapter 6 and an analysis of investment in electricity generation capacity in Chapter 7. An extensive list of literature is provided for further reading.
2 Background

2.1 Policy principles
Traditionally the public goals for the energy supply industry are reliability, affordability and environmental responsibility. In practice, a trade-off between these goals is inevitable. Reliability, for instance, comes at a cost, which raises the question how much society is willing to pay for a given level of reliability, and whether the optimal level of reliability is higher or lower than the current one. In a liberalized market, the question of how government can obtain such a goal is not trivial.

The essence of liberalization is that competition is introduced where possible, under the assumption that the pressure from competitors will force market parties to become more efficient, which should benefit consumers. Not all parts of the electricity and gas sectors are amenable to competition, however. The networks, among others, are natural monopolies. Three policy goals can be identified for these natural monopolies:

• to obtain optimal quality of service
• to establish optimal tariffs
• to ensure equal access for all.

To keep competitive parties from exploiting control over monopoly functions, the two types of activities need to be ‘unbundled’. How to do this and the extent to which unbundling is necessary are matters of debate, however, which we will discuss in the following chapters. Unbundling is not enough, however, to achieve the above goals. Government has a choice between obtaining the above goals by owning and operating the infrastructure (and other monopoly functions) itself or allowing private ownership and obtaining the goals through regulation.

A final issue is that the Dutch electricity and gas sectors are strongly interconnected to the infrastructures, and therefore to the markets, of neighboring countries. A condition for an efficient market is that market parties face the same conditions, in terms of taxes, subsidies, network access rules and tariffs etcetera, in related parts of the network. This is currently not the case in Europe.

2.2 Motives for restructuring
A large number of countries have embarked on the process of power sector restructuring, each starting from its particular national context (see Jamasb and Pollitt, 2005). A wide variety of motives can be observed for these attempts. Prior to liberalization, the electricity sector was considered a natural monopoly which, in the prevailing
rneo classical approach, justified state intervention (Scherer, 1980; Stiglitz, 1986).

In the US, privately owned utilities where regulated by sector-specific Federal and State agencies. In Europe and its many (former) colonies and Latin America, public ownership became the dominant mode. In the communist world, of course, the state was the owner and operator of virtually the entire economy. In addition to providing electricity at reasonable rates, power sector policy served a number of public interests associated with these services, such as issues of safety, security of supply, acceptable prices for specific types of users, objectives of local and sector development, the supply of jobs, and – more recently – sustainability and environmental protection. (See, for example, Foreman-Peck and Milward, 1994; Correljé et al., 2003).

By the early 1980s, this perspective was replaced by the kind of ‘liberalism’ that was associated with the ideas promoted by the governments of the late Ronald Reagan and Margaret Thatcher, which were based on monetarist and public choice theories. Efficiency, economic reform and political power were sought through a reduction of taxes, “rolling back the state” and by bringing market-driven competition into so-called gold-plated industries. Competition – modeled after the revised economic textbooks – was to be imposed upon public sectors wherever possible (Friedman, 1962; Demsetz, 1968; Helm, 2003; Parker, 2000). Gradually – initially only in a number of Anglo Saxon countries and Chile – privatization and competition were introduced as the basic elements of structural change in the energy sector. Encouraged by the apparent successes of the early efforts, the introduction of competition became a common goal for many markets (see also Chao, H. et al this book).

As the competitive market paradigm became accepted in OECD countries, international organizations such as the World Bank and the IMF started to require developing countries to implement similar market reforms if they wished to be eligible for support. Thus, for many developing countries – as for some EU countries – reform was imposed, rather than a voluntary effort. Another possible reason for reform in developing countries is to attract private capital when the public owners cannot provide sufficient investment.

In addition, other policy objectives often play a role, such as social and economic stability, fuel policy, environmental policy and CO₂ trading. These policies may impact restructuring policy by excluding or enforcing particular fuel mixes and through systems of levies and subsidies (see Glachant and Lévéque, 2005). Moreover, competition policy may influence the success of restructuring. In general terms, it appears that the commitment to, and faith in, competition as a means of maximizing welfare is an important factor. The success of market reforms in other utility sectors may turn out to be an important example to policy-makers.
2.3 System description: electricity

Liberalization has separated the value chain of electricity, which used to be largely integrated. Different actors now control different parts of the electricity system. This has caused a significant increase in the complexity of the sector. Therefore it is useful to discern between, on the one hand, the physical, technical side of the system and the economic, institutional side on the other hand. The technical system consists of the physical chain through which electricity flows, from the power plants in which it is generated, through the transmission and distribution networks (with all their supporting equipment) to the apparatus in which the electricity is consumed, the ‘load’. The institutional layer of the electricity system consists not only of the actors who control the components of the physical system, but also of other parties.

We use Figure 1 to describe the electricity system. In this figure the double-pointed arrows indicate which actors control which parts of the physical system. The arrows with single points indicate the direction of the trade in electricity. The figure is based on the Dutch market, but because the Dutch market is a fairly rigorous implementation of the textbook ideal of a decentralized electricity market (see page 39), we will use it as a reference model.
The supply side of the electricity market consists of generation companies, the owners of the power plants in which electricity is produced. These companies may or may not be part of larger companies. Based upon the prices that the producers offer, the market decides upon the quantity of electricity that each generation company may sell at each moment in time, but the generating companies themselves decide which of their power plants they run. Thus generating companies and consumers together determine the load flow in the electricity network.

**The market: mostly bilateral**

Producers and consumers meet in ‘the’ market, which actually consists of multiple related markets. The largest volume of electricity, about 85% in the Netherlands, is sold in the ‘bilateral’ market. This means that the electricity is sold directly by the generating companies to their customers. These are either large consumers, traders or (often) supply companies who deliver it to small and medium size consumers. Bilateral contracts are confidential, as a result of which there are no good data available regarding their price and duration. According to traders, however, contracts that are longer than a year are rare, however.

**Power exchange**

The APX is the only Dutch power exchange. Its most important activity is the spot market, where electricity is traded for the next day on an hourly basis. This means that the market is cleared separately for every hour of the day. In addition, markets for intra-day trade and future contracts are growing. The need for short-term trade is created by unforeseen changes in demand and supply (generator availability) and the inability to store electricity in a commercially viable way. As a result, the value of electricity is much higher during the peak than during the night. Sellers and buyers trade through the APX spot market: there is only one price for each hour. A power exchange offers a standardized trading platform, which offers low transaction costs and anonymity to its users. The market clearing price, however, is published, which provides a valuable reference function to the bilateral market. Spot prices can be highly volatile, which makes them unpredictable. For this reason, the majority of electricity is traded in the bilateral market. The spot market is highly valuable, however, to producers and consumers who wish to change their production or consumption in the short term. The shares of the APX spot market are owned by TenneT, while the shares of the forward markets are privately owned.

**TSO**

TenneT is the Dutch Transmission System Operator (TSO). It has three tasks: to balance the injections and withdrawals of power in the transmission network, to manage the Dutch electricity transmission network and to manage import capacity. The first task involves management of the energy balance of the transmission network, since energy cannot be stored in the network. The second and third tasks concern network (infrastructure) management: providing enough transmission capacity, operating it (which includes controlling the voltage and the reactive power in the network) and managing
congestion. Thus most of the functions that are not amenable to competition have been placed with TenneT. The only other monopoly function is the management of the distribution networks.

Because electricity cannot be stored in the network, unbalances can cause the electricity system to become unstable and may result in large disruptions of service. Therefore all users of the Dutch electricity system (electricity producers supply companies and consumers) are ‘program responsible’. Program responsibility can only be exerted by recognized ‘program responsible parties’. In the Netherlands there are a few dozen of such parties (large producers, large consumers and supply companies). Most users of the electricity system transfer their program responsibility to one of the recognized program responsible parties. For small consumers this role is filled by supply companies such as RWE/Essent, Vattenfall/Nuon and Eneco, or any of the many other companies that sell to small consumers.

A program responsible party submits, on behalf of all users for which he exerts program responsibility, energy programs to the system operator TenneT. These energy programs are schedules which indicate the planned generation and/or consumption of the program responsible party and/or his customers. All submitted energy programs need to be consistent with each other, so there is a closed energy schedule for the whole of the Netherlands for each moment. It is the responsibility of TenneT to administer this system.

However, it is impossible to forecast precisely how high electricity consumption will be at every moment of the next day. In practice, supply and demand will not be balanced all the time. As system operator, TenneT has the responsibility of maintaining the physical balance nevertheless. To this end, TenneT has operating reserves (power plants that can generate electricity on short notice) with which it can maintain the real-time balance of the Dutch electricity system. TenneT contracts this reserve power in the ‘balancing market’, which it also operates. In this market producers can offer reserve power, but they also pay if they generate less than they indicated in their energy programs. Consumers are forced to buy any power that they consumed in excess of their energy program in the balancing market, typically at prices higher than the spot market. By making consumers and producers financially liable for deviations from their energy programs, they receive an incentive to submit energy programs that are as realistic as possible. This is the second aspect of program responsibility.

Another role of TenneT is that of the manager of the transmission network. This entails, among others, keeping the different parts of the network at the right voltage level and preventing or managing congestion. TenneT owns most of the networks of 110 kV, 150 kV, 220 and 380 kV and operates the rest of these networks. In 2011, TenneT implemented the congestion management method called redispachting for domestic congestion. Congestion on the borders is
handled through market coupling (for the day-ahead market) and explicit auctions for month and year capacity (CASC, 2011).

It is possible that the transactions that are planned by the market threaten to exceed the physical capacity of the network. To allocate the capacity of the interconnectors efficiently, month and year contracts for interconnector capacity are auctioned. In addition, day-ahead interconnector capacity is allocated through market coupling (which resembles market splitting). See Chapter 6 for an overview of key congestion management methods. TenneT operates these auctions together with the network operators on the other side of the interconnectors. The revenues are divided between the network operators; TenneT’s revenues are placed in a separate fund, in order not to give TenneT an incentive to create congestion. Within the Netherlands, TenneT has started to apply redispersching in 2011 to alleviate sporadic congestion, mostly between the Maasvlakte (Rotterdam Port) and the rest of the network.

The distribution networks are mostly owned by independent network companies such as Enexis and Alliander, since they were unbundled from the incumbent energy companies in 2010. Large exceptions are Delta and Eneco, which still own distribution networks. All electricity distribution network companies in the Netherlands are currently owned by local governments. Full privatization is prohibited by law, but the government is contemplating privatization of minority shares in order to attract more capital.

In the 1980s, the Dutch electricity sector was organized hierarchically, not only institutionally but also technically. Electricity was generated in large power plants, transported via the transmission network and brought to the final consumers through distribution networks. Only the largest consumers took their electricity directly from the transmission network. (Both the transmission network and the distribution networks are made up of networks of different voltages, linked by transformers. There is a functional difference, however, between networks that are intended to transport electricity over distances and networks that distribute electricity from a central point to consumers.)

The Dutch Electricity Act of 1989 stimulated the application of decentral electricity generation (also called distributed generation or embedded generation) due to the large environmental advantages. This caused a significant growth of decentral power generation in the 1990s, so that an estimated third of all electricity generation capacity now consists of small units. Typical of these units is that they feed in to distribution networks. Consequently the role of these networks is changing: in addition to distributing electricity that is generated elsewhere, they also provide local transport services. Figure 2 shows the structure of the networks in a simplified manner.
2.4 System description: natural gas

Liberalization also separated the production chain of natural gas, which was largely integrated previously. Also in this case liberalization led to a greater complexity of contracts and information flows and to changes in the roles of the different actors in the system. Analogous to the above discussion of the electricity sector, a distinction can be made between the physical, technical side of the gas sector on the one hand and the economic, organizational side on the other hand. See Figure 3. In this section, we will move through this diagram from bottom to top and left to right, as much as possible.

Physically, the gas system consists of on and offshore fields from which gas is produced, storage facilities, transport networks (including pipelines for import and export), local distribution networks and the installations and appliances in which the gas is used (combusted).

With respect to the origin of natural gas in the Netherlands, a distinction can be made between (a) the extraordinary large field in Groningen, (b) other medium-sized on and offshore fields, (c) the small offshore fields, (d) natural gas that is imported through pipelines and (e) gas that is imported by ship in the form of liquefied natural gas (LNG). Different gas fields have different physical characteristics and production costs. Gas from the large Groningen field is relatively low cost, due to the scale of the field, its location near consumers, the natural pressure in the field and the ease with which it can be turned ‘on’ and ‘off’ (its flexibility). This flexibility, however, diminishes as the field is depleted. This means that investments in compressors and adjacent ‘system’ storage are necessary to maintain sufficient pressure. The medium-sized on shore fields are somewhat more expensive than the Groningen field. The costs of producing gas offshore are substantially higher due to the longer distance across which the gas needs to be transported and due

---

**Figure 2: Schematic representation of electricity networks**

- G/g: large/small generator
- C/c: large/small consumer

---

Model: economic and physical layer

Physical layer

Different types of gas fields
to the more difficult working conditions at the North Sea. The so-called small fields are even more costly due to the limited volume of gas that they contain relative to the exploration and production investments and operational costs needed to extract the gas. The supply costs of imported gas are mostly related to the distance over which the gas needs to be shipped: the farther the more expensive. Cost, however, is no indication for the price at which the gas is sold, as will be shown below.

**Transmission and distribution**

The transport networks for gas can be divided in long-distance transport (transmission) networks and distribution networks, like for electricity. In the transmission network gas is transported under high pressure over long distances, while the distribution network serves to conduct gas at a regional level to the end users. This takes place at lower pressure levels.

**Actor layer**

The actor layer is composed of the actors who control the physical components of the system, such as producers, network managers and consumers, as well as other parties, such as traders, brokers and retail companies.

**Production**

The supply side of the gas market - also called *upstream* - is characterized by a diversity of producers. The Groningen field is exploited commercially by a partnership in which the State of the Netherlands (represented by EBN, see below) has a 40% share and the Nederlandse Aardoliemaatschappij (NAM, which means something like the ‘Dutch Petrol Corporation’) has a share of 60%. The actual production of the Groningen gas is performed by the NAM, which is a 50/50 daughter company of Shell and Exxon. The small fields and the ‘other’ fields (a few larger offshore fields) are developed by consortia of several oil and utility companies, among which NAM. All these Dutch fields are developed under concessions which are given by the Dutch Ministry of Economic Affairs to consortia of oil companies that were established specifically for that purpose. In each of these consortia the State of the Netherlands, represented by Energie Beheer Nederland (EBN, ‘Energy Management Netherlands’), participates for 40 or 50% as silent partner, next to the oil companies, as a result of which the Dutch state receives a part of the profits directly.
Due to the high costs, development of the small gas fields (for instance underneath the North Sea) only is profitable when their yield is as constant as possible. Because the government wishes to promote the development of these small fields – in order to maximize the volume of gas that is produced in the Netherlands and to save the Groningen field as long as possible – the so-called small fields policy ('kleineveldenbeleid') was introduced in 1974/1975. This policy makes use of the fact that the output of Groningen can be increased or decreased in a rather flexible manner. GasTerra (the Dutch incumbent gas trading company, see below) gives first priority to the gas from the small fields, next to imports and the regular supply contracts with the larger on and offshore fields and uses the Groningen field to make up the balance with demand. This way the profitable development of the small fields is stimulated.

Imports supplement domestic gas production. The European Union imports natural gas through long distance pipelines from Norway, Russia and Algeria, while pipelines to other countries are planned or under construction. Gasunie, GasTerra’s predecessor, used to import gas from Norway to spare the Groningen field and to keep in touch
with the supply side of the market. Currently, large quantities of gas are imported by a variety of traders without intervention by GasTerra. In addition to imports of (gaseous) natural gas through pipelines, liquefied natural gas (LNG) is transported to other places in the EU with tankers from North Africa, the Middle East and Trinidad and Tobago. In the near future, LNG will be landed in the Netherlands, via the terminal currently under construction in the Rotterdam area.

Slightly more than half of the gas that is produced in the Netherlands is exported. From the beginning of Dutch gas production, Gasunie managed the exports of Dutch gas to Germany, Belgium, France, Switzerland and Italy. This meant that it made all the contracts with the gas wholesale and retail companies from those countries and ensured that there was enough gas for them, through a combination of domestic production and imports. Local gas prices, indexed to local gas oil prices, determined the contract prices; Gasunie received the revenues that remained after subtracting transport costs (the net-back principle). From the beginning of the 1980s Gasunie also supplied the necessary seasonal flexibility to Northwest Europe. This meant that in summer little or no gas was supplied, while the great majority was supplied in winter months. Intra-day flexibility usually was arranged locally.

The Dutch gas sector used to be dominated by Gasunie, an integrated company that managed practically all functions in the midstream part of the gas market, from the delivery of gas from all sorts of sources by the gas producing oil companies to the delivery of the gas to large consumers and retail companies. On the one hand Gasunie performed monopolistic functions such as the management of the national transport network for gas and the import and export connections. On the other hand Gasunie had a commercial task: it sold nearly all domestically produced gas to domestic and foreign distribution companies and Dutch wholesale users. Gasunie is owned by the State of the Netherlands (50%) and of Shell and Exxon (each 25%).

The double role of Gasunie did not appear to be compatible with the current structure of the gas sector, with the increased emphasis on competition, especially considering the public (state) involvement with commercial activities and the participation of private parties (Shell and Exxon) in monopoly activities. Therefore an attempt was made to divide Gasunie in 2002. The proposal was that the monopoly activities, that is, the transport network and the future network and system manager (GTS, see below), would become fully state-owned. The commercial activities, such as trade and delivery of gas, would become the domain of the future trade division of Gasunie (Gasunie T&S, see below), which would be owned by Shell and Exxon. Shell and Exxon would be allowed to divide Gasunie T&S between them and incorporate the parts in the mother companies. (The position of NAM, 50/50 Shell/Exxon and producer of the majority of Dutch gas,
remained unclear.) However, because the parties could not agree about the division of the assets among them, this attempt at ownership unbundling of Gasunie failed. This discussion stranded, however, among others on the question how to determine the value of the assets which would be redistributed between the State and the private companies and how the competitive parts would be regulated. Another stumbling block was how to shape the small fields policy and the long-term depletion policy in a competitive environment.

However, a few years later, full unbundling was achieved. In 2005 a juridically unbundled network manager was created: Gas Transport Services (GTS). GTS continued to be owned by Gasunie, whose shares continued to be owned by Exxon (25%), Shell (25%), EBN (Energiebeheer Nederland, 40%) and the Dutch State (10%). Around that time, the trade leg of Gasunie, Gasunie Trade & Supply (T&S), was made independent and changed its name to GasTerra. After lengthy negotiations, the Dutch government reached an agreement with Shell and Exxon to purchase GTS for €1.8 billion. From 2005 on, Gasunie, with GTS as its main daughter, was fully nationalized, while GasTerra continued with the traditional public-private shareholder structure.

The creation of a separate, juridically independent network manager was made necessary by the Second European Gas Directive 2003/55/EC.1 Since July 2nd, 2004, Gas Transport Services (GTS) is the manager of the national gas transport network. GTS has the tasks that belong to the function of network manager as well as tasks as the system operator. As transmission network manager, GTS takes care of the independent management and development of the Dutch high pressure gas network and the provision of transport services and other related services. GTS provides these services on demand at regulated tariffs. GTS is not allowed to supply gas or services with which it enters competition, except for network balancing, following from its legally mandated tasks. Since July 2nd, 2004, GTS is a limited liability company with its own board of commissioners. GTS is owned by Gasunie, which remains owner of the transport network itself. All shares of Gasunie are held by the Dutch Ministry of Finance.

GTS also has the role of system operator. An important aspect is the task of maintaining the physical balance between supply and demand on the Dutch gas network. Although the margin of tolerance is greater than in the case of electricity, imbalance between supply and demand may cause the gas network to become instable, which may cause large service interruptions. Like for electricity, it is impossible to know precisely how high demand will be the next day, among others due to the strong influence of the temperature upon demand.

1 Pb EU 2003 L 176/37.
It is the task of GTS to correct any differences between scheduled supply and actual demand.

In the Netherlands, gas can only be traded by recognized parties, so-called *shippers*, of which there are a few dozen in the Netherlands. (These shippers may be compared to program-responsible parties for electricity.) On behalf of gas producers and consumers, shippers arrange the transport and the associated contact with GTS. Shippers are required to ensure that their input into the network and their withdrawal of gas from the network are the equal to each other at any time. As part of this, they are required to notify the system operator GTS of their planned supply into the network ('entry') and their withdrawals ('exit', thus 'entry-exit booking'). Through intelligent portfolio management, it is easier for shippers than for individual producers or consumers to remain in balance. So in practice shippers often take care of balancing for the producers and consumers on whose behalf they act. GTS provides incentives to the market parties to maintain their balance as much as possible in a way that is comparable to the balancing system of TenneT. Market parties need to arrange their trades in such a way as to minimize the development of imbalances. If necessary, they can purchase different forms of flexibility from GTS. GTS has a number of options, such as stand-by gas supply contracts with GasTerra, under which the necessary flexibility is provided by the Groningen field, and contracts with other gas suppliers. GTS contracts these flexibility services through a tendering process from multiple suppliers. In addition, there is for instance the possibility to store gas for later use during periods of low consumption. These flexibility services are used to maintain the system balance continuously. The costs are transferred to the shippers, based upon the degree to which they were not in balance. In principle shippers can obtain flexibility services elsewhere, but the options are limited.

It follows that the former Gasunie had an important role in the coordination of the different gas flows which are supplied to the network at different locations and which have different characteristics (such as caloric value). The coordinating role of Gasunie also was a crucial factor in the ‘small fields policy’. This central coordination used to take place out of view of the users – the current shippers – as a technical activity. In the current, more unbundled system, coordination needs to be arranged through commercial transactions, through which shippers need to contract a package of gas, conversion services and transport capacity, such that they meet their obligations towards their consumers as well as the system requirements of GTS. The way in which this system of decentralized coordination will develop is a significant concern. The shippers do not have an interest in optimizing the overall use of the system. The operator, GTS, is better equipped for that purpose, among others because it can take decisions that are not related to specific supply contracts.
A second task of GTS as system operator is to ensure that the gas has a constant quality. The gas from ‘Groningen’ has a lower calorific (combustion) value than the gas from the North Sea fields and the gas that is imported. An important indicator for the quality of gas is the Wobbe index. To make sure that consumers do not experience a change in the quality of their gas, GTS mixes high calorific gas with nitrogen (N₂), an inert gas, so it achieves the lower calorific value that the consumers are used to. GTS also manages other quality aspects, such as the moisture content of the gas.

Thus GTS has most of the important tasks within the gas system that cannot be performed competitively, analogous to the role of TenneT in the electricity system.

The only other clear monopoly task is the management of the distribution networks. These are managed by regional energy companies, many of which also manage electricity distribution networks, such as Eneco, Essent and Nuon.

Like the electricity market, ‘the’ gas market consists of multiple related markets. The majority of all gas that is produced in the Netherlands is sold to GasTerra, which resells it to its customers. Often this customer is a retail company that sells the gas to small end users. Large consumers tend to buy directly from GasTerra or from other suppliers. These bilateral contracts are confidential, as a result of which there are no reliable data concerning their price and duration. Contracts for more than a year ahead are rare, however, according to traders. Over the past few years, however, increasing volumes of gas are being traded on the Title Transfer Facility (TTF), a kind of spot market, or gas hub.

Since 2005, there is a gas exchange comparable to the APX in the Netherlands: the APX Gas. In combination with the TTF, a virtual wholesale hub on the transmission system, in which many pipelines join each other and where multiple suppliers and consumers can trade. On the gas APX spot market, where gas can be traded by the hour, in an automated fashion and with standardized contracts. Like electricity prices, gas prices can be highly volatile. For this reason, a large part of trade will continue to take place in the bilateral market. A spot market is very useful, however, for producers and consumers who wish to change their output or their consumption on short notice. In addition, the gas spot market is gaining an important role in providing a reference price for the bilateral market, like the APX already does for electricity.

Three categories of gas consumers are distinguished, based on their type of consumption. Firstly there are the large consumers with a consumption that exceeds ten million cubic meters per year. Together, they have a market share of about 45% in the Netherlands. Typically, these are large companies in the heavy and process industry and electricity producers. Daily and seasonal consumption
patterns vary by consumer, but many have a relatively flat load profile. Technically, large consumers have always been free to choose their supplier, but in the past Gasunie always was able to underbid its competitors. Secondly there are intermediate consumers with an annual consumption between 0.17 and 10 million cubic meters. Consumption patterns typically are somewhat variable throughout the year and vary, of course, by company. This category contains small and medium-sized businesses and also part of the greenhouse industry. Small consumers – households, service industries, small businesses with an annual consumption of less than 170,000 cubic meter per year – are highly temperature and time sensitive. They constitute about 20% of total consumption. In the past, prices and delivery conditions were generally determined politically, influenced by sectoral industry policy and regional policies. Currently the volume of demand (which determines the necessary capacity) and the consumption pattern (the duration curve) are leading.
3 Europe’s energy markets

3.1 EU directives for the electricity and gas sectors

The liberalization of the European electricity and gas sectors was initiated by the Electricity Directive 96/92/EC\(^2\) (December 1996) and the Gas Directive 98/30/EC\(^3\) (June 1998). In the meantime, they have both been replaced by new legislation. In June, 2003, a new Electricity Directive 2003/54/EC\(^4\), a new Gas Directive 2003/55/EC\(^5\) and an Electricity Regulation (EC) No. 1228/2003\(^6\) were adopted. The new directives required implementation by the member states by July 1\(^{st}\), 2004, at the latest. The central theme of the directives is that the electricity and gas sectors consist of different activities, some of which are considered natural monopolies while others can be provided competitively. The regulation (which does not require implementation) entered into force July 1\(^{st}\), 2004. The regulation mainly concerns cross-border trade in electricity. In 2009, the EU adopted a comprehensive ‘Third Energy Package’, with new electricity and gas directives as well as several regulations.

Electricity prior to 2009

The Electricity Directive considers the transmission and distribution of electricity, that is, the transport of electricity on the high tension network, respectively on the intermediate and low tension networks, as monopoly activities. System operation (maintaining the physical balance between supply and demand) also is a monopoly. Electricity generation, trade and supply are considered competitive activities. For the management of transmission and distribution networks, network managers need to be appointed.\(^7\) A network manager is a legal entity that is responsible for the operation, the maintenance and the development of a network in a certain region and who is responsible for the connections with neighboring networks.\(^8\) One of the challenges is to ensure that the network has enough capacity in the long term to meet the demand for electricity without overbuilding.

---

\(^3\) Pb EC 1998 L 204/1.
\(^4\) Pb EU 2003 L 176/37.
\(^5\) Pb EU 2003 L 176/57.
\(^6\) Pb EU 2003 L 176/1.
\(^7\) Article 8 Electricity Directive 2003/54/EC.
\(^8\) The term ‘system manager’, which is used in the EC Directives, is confusing. In Europe, typically the same party manages the network and is system operator. Therefore we prefer to speak of the functions of network management and system operation. In this text, we will reserve the term system operation for the operation of the entire electricity system or the entire gas system. In case of subsystems, such as a gas storage facility, we will use terms like ‘storage facility management’.
Transmission and distribution network manager need to be ‘unbundled’ from commercial activities in the electricity market to ensure fair competition. A network manager needs to be independent from other activities that are not related to transmission or distribution at least with respect to its legal structure, organization and decision making. If this independence is guaranteed, the network managers may (continue to) be part of a vertically integrated energy company. The European rules contain no requirement to separate ownership of network assets from the vertically integrated company.\(^{11}\)

To prevent discrimination, cross subsidization and unfair competition, electricity companies have separate accounts for all their transmission and distribution activities, their other electricity activities and their activities that are not related to electricity. Revenues from the ownership of transmission or distribution network must be specified in the accounts.\(^{12}\) In addition a network manager must respect the confidentiality of the commercially sensitive data which he receives in the line of operating his networks.\(^{13}\)

Network managers need to provide third parties access to their networks. The system of third party access must be based upon published tariffs that apply to all eligible (‘free’) customers and that are applied objectively, without discrimination between users of the network. The tariffs themselves, or the methodology that underlies the calculation of the tariffs, require prior approval by the (national) regulator and must be published.\(^{14}\)

The generation of electricity is considered a free market activity. Member states may, however, include requirements in the permit that is required for the construction of new generation capacity. Permits must be awarded based upon objective, transparent and non-discriminating criteria.\(^{15}\) If the ‘market’ does not provide sufficient (new) generation capacity, a member state can provide the necessary capacity through a tendering procedure or another procedure that is

\(^9\) Article 2(4) and 2(6), and Articles 9 and 14 of the Electricity Directive 2003/54/EC.
\(^{10}\) Cf. for instance article 11 of the Electricity Directive 2003/54/EC.
\(^{11}\) Articles 10 and 15 Electricity Directive 2003/54/EC.
\(^{12}\) Article 19 Electricity Directive 2003/54/EC.
\(^{13}\) Articles 12 and 16 Electricity Directive 2003/54/EC.
\(^{14}\) Articles 20 and 23 Electricity Directive 2003/54/EC.
\(^{15}\) Article 6 Electricity Directive 2003/54/EC.
at least equivalent with respect to transparency and non-discrimination.16

The supply of electricity to so-called eligible consumers is free. According to Electricity Directive 2003/54/EC, all non-residential electricity consumers in the EU had the right to select their electricity supply company of choice from July 1st, 2004 on, and all consumers from July 1st, 2007. Such eligible consumers have a claim to access to the network.17 In the Netherlands, supply already has been free for all consumers since 2004.

The Electricity Directive explicitly states the possibility for member states to implement ‘public service obligations.18 ‘ Based upon article 86 of the EC Treaty, member states are allowed to assign companies with a public service obligation in the general economic interest. To the extent that that service is essential, it may cause certain European rules not to apply. The Electricity Directive requires one specific public service obligation: the member states must ensure that all residential customers, and if desired small businesses also, have a claim to universal service, which means the right to the supply of electricity of a certain quality for reasonable and transparent prices in their territory.19

Gas prior to 2009
The Gas Directive regulates the transmission, distribution, supply and storage of natural gas. The directive also concerns liquefied natural gas (LNG). Non-discriminatory organization of the ‘production’ (extraction) of natural gas is outside the Gas Directive, but was regulated before in the so-called Hydrocarbon Directive.20 The transport and distribution of natural gas, that is, the transport of the gas through a high pressure pipeline network, respectively through regional and local pipeline networks, are considered to be monopoly activities. The new Directive requires regulation of storage of natural gas and LNG facilities. The supply of natural gas can take place competitively.

For the management of transmission or distribution networks and for storage or LNG facilities, ‘system managers’21 need to be appointed.22

16 Article 7 Electricity Directive 2003/54/EC.
17 Article 21 Electricity Directive 2003/54/EC.
18 Article 3(2) Electricity Directive 2003/54/EC.
19 Article 3(3) Electricity Directive 2003/54/EC.
21 Again, as in the case of electricity, the term ‘system manager’, which is used in the EC Directives, will be avoided here because it is confusing. See also footnote 8.
22 Article 7 Gas Directive 2003/55/EC.
A transmission or distribution system manager is a natural or legal person who carries out the function of transmission or distribution and is responsible for operating, ensuring the maintenance of, and, if necessary, developing the distribution system in a given area and, where applicable, its interconnections with other systems, and for ensuring the long-term ability of the system to meet reasonable demands for the transmission or distribution of gas. A storage system manager (‘operator’) is a natural or legal person who carries out the function of storage and is responsible for operating a storage facility. An LNG system manager (‘operator’) is a natural or legal person who carries out the function of liquefaction of natural gas, or the importation, offloading, and re-gasification of LNG and is responsible for operating a LNG facility.

Transmission and distribution network managers need to be ‘unbundled’ (separated) from the commercial activities. A transmission or distribution system manager of a vertically integrated company must be independent at least in terms of its legal form, organization and decision making from other activities not relating to transmission. If this independence is assured, the network managers may (continue to) be part of a vertically integrated energy company. The European rules contain no obligation to separate the ownership of the assets of the transport or distribution system from the vertically integrated company.

To prevent discrimination, cross-subsidization and distortion of competition, natural gas companies need to keep separate accounts for each of their transmission, distribution, LNG and storage activities. Revenues from the ownership of transmission and distribution networks must be specified separately. In addition, a manager of a transmission or distribution network or of a storage or LNG facility must respect the confidential nature of the commercially sensitive data which he receives as part of his activities.

Managers of transmission and distribution networks and LNG facilities need to provide third parties access to their facilities and services. Third party access needs to be based upon published tariffs that apply without discrimination to all eligible customers. The tariffs themselves, or the methodology underlying their calculation must be approved prior to their entry into force by the (national) regulatory authority. Additionally, the tariffs (or the methodology) must be published prior to their entry into force.

23 Article 2 sub 4 and 6, and Articles 8 and 12 Gas Directive 2003/55/EC.
24 Article 2, sub 10, and Article 8 Gas Directive 2003/55/EC.
25 Article 2, sub 12, and Article 8 Gas Directive 2003/55/EC.
26 Article 9 and 13 Gas Directive 2003/55/EC.
27 Article 17 Gas Directive 2003/55/EC.
28 Articles 10 and 14 Gas Directive 2003/55/EC.
29 Articles 18 and 25 Gas Directive 2003/55/EC.
For the access to storage facilities, member states may choose between negotiated and regulated third party access. In the case of regulated access, eligible customers have a right to access to the storage facilities on the basis of published tariffs and terms, while in the case of negotiated access eligible customers can negotiate about access to the storage facilities, while the storage system manager needs to publish his main commercial conditions (indications of his tariffs and terms).30

The supply of gas to eligible customers is free. Based upon the recent Gas Directive 2003/55/EC, from July 1st, 2004 all non-residential customers must be ‘free’, and from July 1st, 2007 all customers. Such eligible customers are eligible for access to the transmission and distribution networks and to storage and LNG facilities.31 In the Netherlands all customers already are free.

Contrary to the case of electricity (see above), the Gas Directive does not impose a public service obligation. The Gas Directive does explicitly state the possibility for member states to implement a public service obligation.32

3.2 The European Union’s Third Energy Package

In April, 2009, the EU adopted a new set of directives and regulations in the energy field with the goal of improving the functioning of the market, strengthening consumer rights and improving security of supply, especially in view of potential disruptions to the supply of gas. It is called the Third Package in reference to the two earlier rounds of energy legislation: the original electricity and gas directives of 1996 and 1998 and the second directives of 2003. The Third Package consists of:


30 Article 19 Gas Directive 2003/55/EC.
31 Article 23 Gas Directive 2003/55/EC.
32 Article 3(2) Gas Directive 2003/55/EC.
The Third Package was developed partly in response to the findings of the EU’s Sector Inquiry (EU, 2006), which found a lack of market integration and an abundance of market power. The findings of a special ‘rapporteur’ supported the idea that further regulation was needed to create effective competition. The Third Package represents an attempt to improve wholesale markets through better regulation, the unbundling of transmission from generation, trade and retail and better integration of national European markets. The latter is considered to serve the goals of economic efficiency, reducing market power and security of supply.

A key goal of the Third Package was to remove the strategic advantage that vertically integrated firms have from combining network ownership with production and trade activities. The original intention appears to have been to require all TSOs to become fully (ownership) unbundled from competitive activities. This would mean that the vertically integrated firms would need to sell their transmission networks and new, independent TSOs would need to be established. However, already at the first presentation of the legislative proposals, the EC offered the concept of Independent System Operators (ISO) as an alternative. The concept of ISO is derived from the USA, where several ISOs exist. An ISO is an independent company that takes care of the operational aspects of network management, such as load balancing, voltage control and congestion management, without being the owner of the network. The name is somewhat of a misnomer, because like a TSO an ISO takes care of both the energy balance (system operation) and transmission line operation. In the course of the debate, a third alternative, the Independent Transmission Operator (ITO), was launched. This, too, is a misnomer, because its role is the same as a TSO but it is not independent. To the contrary, this option, put forward by the industry, entails more or less a continuation of the status quo in which transmission ownership and management may be integrated with commercial activities, only with stronger regulation. Among others, investment and asset valuation will be regulated and the network operator will need to have a compliance program and compliance manager to prevent discrimination in network access. The outcome of the debate is that member states may choose which of the three options they implement.

In order to improve and harmonize the regulation of energy by national regulators, an Agency for the Coordination of Energy Regulators (ACER) was created. ACER has mainly a monitoring and

---

advisory role; its impact will depend on the degree in which the EC will adopt its recommendations. It replaces ERGEG (the European Regulators Group for Electricity and Gas) as the cooperation platform for national energy regulators. ACER will regulate ENTSO (see below) and will also have decisive authority over cross-border network capacity in cases where the involved countries cannot agree.

The directives also require the establishment of national regulators and stipulate the functions that this regulator must perform at a minimum. A number of provisions have been made to ensure the full autonomy and independence of the national regulatory agencies, also with respect to their budgets. The directives also have as an objective to expand and harmonize the powers of the national regulators.

Another instrument in the integration and harmonization of European energy markets is the creation of ENTSO: the European Network of Transmission System Operators. ENTSO is to create European standards for network codes and monitor their implementation, plan network development, monitor generation adequacy, coordinate research and make recommendations for binding EU regulation.

With respect to consumer protection, an important requirement is that consumers must be informed regularly of their energy consumption. This can only be done with the aid of ‘smart meters’, digital energy meters that measure energy consumption frequently and can be read out electronically, so effectively the EU is requiring the installation of smart meters at all consumers. Other aspects of consumer protection are time limits for switching from provider and for bills, the requirement to create agencies for conflict resolution between energy companies and consumers, and protection of vulnerable consumers against being disconnected.

3.3 Dutch electricity and gas legislation

The European Directive 96/92/EC and the Gas Directive 98/30/EC caused a thorough adjustment of the Dutch legislation of electricity and gas. In the following, an overview is presented of the development of legislation in the Netherlands from the beginning of liberalization.

Electricity

Before liberalization, the 1989 Electricity Act was in force in the Netherlands. This act was based upon central planning and operation in the electricity supply industry, with a central role for the Samenwerkende Elektriciteits-produktiebedrijven (SEP, ‘Cooperating Electricity Production Companies).

---

34 Staatsblad 1989, 535.
To implement the Electricity Directive 96/92/EC, in 1998 that act was replaced by a new Electricity Act in the Netherlands.\textsuperscript{35} The 1998 Electricity Act was, originally, structured in a fairly transparent manner. Activities which became free economic activities under this act, such as the production of electricity and its supply to eligible customers, were no longer regulated by the 1998 Electricity Act. The act limited itself to activities which remained a monopoly, such as network management and supply to captive customers. Network management would need to be regulated by the newly created DTe (Dienst uitvoering en Toezicht Energie, Dutch Office of Energy Regulation, now called Energiekamer, or Energy Office of the Dutch Competition Authority).

The 1998 Electricity Act has been changed many times in its short existence. The first change already dates back to 1999, when the most important tariff structures and conditions for the network managers were included in the law.\textsuperscript{36} In 2000 the Gas Act followed,\textsuperscript{37} in which DTe also was provided some tasks with respect to gas. The Gas Act regulated a ‘free’ activity: the supply of electricity to eligible small customers was submitted to a license requirement.

In the meantime, the issue of stranded costs\textsuperscript{38} required a more extensive legislative solution. The term ‘stranded costs’, also called ‘sunk costs’, refers to costs that were incurred under the old, regulated monopoly, but that were expected not to be recoverable in a competitive market. Examples of such stranded costs are district heating projects and certain investments for environmental reasons. The solution for this stranded cost problem was provided by the 2000 Transition Act for the Electricity Production Sector (‘Overgangswet elektriciteitsproductiesector, OEPS’).\textsuperscript{39}

The Transition Act also caused ownership of TenneT to be transferred to the Dutch State.\textsuperscript{40} In addition, the Transition Act laid the basis for a market-oriented allocation method for cross-border transmission capacity, namely auctions.

\textsuperscript{35} Staatsblad 1998, 427.
\textsuperscript{36} Staatsblad 1999, 260.
\textsuperscript{37} Staatsblad 2000, 305.
\textsuperscript{38} For more information about this subject see: Knops 2000.
\textsuperscript{39} Staatsblad 2000, 607.
\textsuperscript{40} Kamerstukken II, 1999-2000, 27 250, nrs. 1-2, p. 1).
Because one aspect of the stranded costs arrangement was not approved by 'Brussels', the Transition Act for the Electricity Production Sector had to be changed on that point. This correction act subsequently was used to attach many changes to the 1998 Electricity Act. During the same time, a law was enacted to include a new system to stimulate improvement of the environmental performance of the generation of electricity in the 1998 Electricity Act (MEP, ‘stimulering Milieukwaliteit Elektriciteitsproductie’). Later, the 1998 Electricity Act was also changed in order to introduce a system of certificates of origin for sustainably produced electricity.

In the summer of 2004 two laws were enacted. Proposal 29303 concerned the opening of the last part of the electricity and gas markets on July 1, 2004: from that date on all consumers, so including small consumers, were free to choose their energy supplier. Proposal 29372 was nicknamed the ‘I&I-act’, for ‘Implementation and Intervention Act’. ‘Implementation’ refers to the necessary implementation of the Electricity Directive, while ‘intervention’ is the label for all other adjustments to the 1998 Electricity Act and to the Gas Act that were proposed. The next chapter will discuss several elements of the I&I Act.

Gas

Before liberalization, there was no Gas Act in the Netherlands. The institutional design of the Dutch gas sector was arranged by the ‘de Pous’ white paper and through laws pertaining to the exploration and production of oil and gas in the Netherlands. The Gas Act (from 2000) implemented the Gas Directive 98/30/EC. This act, too, has been changed multiple times, the first time by the Transition Act for the Electricity Production Sector (Oeps), which provided for more authority for the director of the DTe, the Dutch Office of Energy Regulation.

Other changes were caused by the Mining Act (Mijnbouwwet), an act which changed the Transition Act for the Electricity Production Sector (Oeps) and the act that implemented the ‘guarantees of origin’ of electricity.

---

41 The European Commission did not approve the proposed levy on the transmission tariff with which the stranded costs would be compensated.
42 Staatsblad 2003, 316.
43 Staatsblad 2003, 235.
44 Staatsblad 2003, 493.
45 Staatsblad 2004, 328.
48 Staatsblad 2000, 305.
In the summer of 2004, two acts were passed. Proposal 29303 concerned opening the last part of the electricity and gas market on July 1st, 2004: since that date, every consumer in the Netherlands is allowed to choose his provider of electricity and gas. Proposal 29372 provides for implementation of the new Gas Directive and also contains other changes to the Gas Act (see above).

### 3.4 The current situation

A set of overall policy goals for the electricity industry can be observed that are fairly stable over time and between countries. These goals may be described as are the ‘triple A’ goals: availability, affordability and acceptability, or in other terms, security of supply, low end user prices, and social and environmental acceptability. Social acceptability refers to issues such as access to minimum energy services for the poor, while the environmental issues are mostly climate change, local air pollution and environmental harm from the production of primary fuels. The availability of primary fuels is usually considered an aspect of security of supply. Ten years after the first EU Electricity Directive (96/92/EC), achievement of these goals remains a challenge. Affordability is not always called that; in the Sector Inquiry, for instance, the EU gives competition as a main goal (EU, 2006). However, in the next sentence, the goals of competition are explained as higher growth rates and increased competitiveness, which only would result if end users paid less on average. The preamble of the original electricity directive also reflects the assumption that improvements in the efficiency of the electricity industry will contribute to the competitiveness of the European economy.

Different starting points, different speeds of liberalization and different market design choices have led to a highly fragmented European electricity market. The consequence is a low degree of transparency, limited competition and significant variations in electricity prices. Competition policy has not been effective in a number of countries, in which the incumbents continue to dominate the market. Also because the excess generating capacity that was present at the start of liberalization has disappeared and because gas prices have risen, electricity prices are rising in many parts of Europe (EC, 2006). Investment in generating capacity was nearly absent during the first years after liberalization, but has been increasing substantially since the beginning of the 2010s. However, it is unclear how many of all the paper plans will be realized, and whether they will be sufficient and arrive on time to maintain the current high level of reliability.

Several technological developments are taking place that have the potential to change the fundamental characteristics of the electricity supply industry, but the full impact of which is as yet unclear.

- The technology to generate electricity in very small units is becoming commercial. A significant advantage is that it would
allow more efficient use of waste heat, which could significantly reduce fossil fuel consumption and the associated environmental impacts. However, a prerequisite of the widespread use of distributed generation, as it is called, is the presence of a primary energy infrastructure such as the natural gas network. This is not the case everywhere. Alternatively, use could be made of renewable energy sources such as solar and wind energy. If distributed generation becomes popular, it may require changes to the electricity networks, both physically and in the way they are operated, priced and regulated.

- As the use of wind and solar energy increases, their fluctuating nature poses increasing challenges. A solution is to increase international transmission capacity drastically, so the fluctuations can be absorbed over a larger area.
- A technology for storing electric energy cheaply would resolve this issue, as well as the issue of peak generation capacity. Many years of research have not yet led to an economically successful technology. However, the high price spikes in liberalized markets have raised the potential rewards for this kind of technology, so research continues.
- The use of ICT in networks and metering opens doors to more innovative applications and contracts.

The European Directives for gas and electricity leave much room for implementation by the member states. This is evidenced by the existence of three institutional models in the region around the Netherlands. The electricity and gas markets in England and Wales were (partly) the source of inspiration for the design of the Dutch markets. Similarities are a relatively high number of competing electricity generation companies and a transmission network manager (National Grid Company) who is juridically separated from commercial market activities. The same pattern can be found in the gas sector, where multiple national and foreign gas producers have access to the independently managed National Transmission System, which is operated by Transco. Recently, National Grid and Transco were merged into a single company. Another similarity is the fuel mix, in which coal and gas play a dominant role, while both countries also are important producers of natural gas, as opposed to the other member states of the EU. Great Britain has a relatively isolated market.50 The electricity network is connected to the mainland through interconnectors with France and the Netherlands. TenneT would like to develop an interconnector with England. The British gas network is connected to the mainland by means of an interconnector to Oostende.

---

50 The England and Wales market will be connected to that of Scotland.
Between 1990 and March, 2001, England and Wales had a mandatory 'power pool', a sort of spot market in which all produced electricity had to be offered and which, therefore, was the only place where retail companies could purchase electricity. This centralized model has lost favor, at least in Europe, and was replaced by NETA ('new electricity trading arrangements'), which gives market parties more contract freedom. The former pool provided producers capacity payments for non-active but available generation capacity. This system, which was intended to secure a sufficient volume of available generation capacity, was manipulated, which was one of the main reasons for changing to NETA. The effect of the manipulation of the pool was that the prices remained high and that excessive investment in generation capacity took place. After the pool and the associated capacity payments were abolished, the existing excess capacity caused the electricity price in the England and Wales market to plummet.

Natural gas is produced in more than 100 offshore fields around Great Britain. Producers supply their gas at one of seven terminals to one of 90 trading companies. These so-called shippers supply the gas to different consumers, or to each other, with the use of Transco's integrated transport network. Production of gas in the British offshore fields has peaked, however, as a result of which the UK is becoming significantly dependent on gas imports. For this reason, the Netherlands-UK gas interconnector 'BBL' (Balgzand-Bacton Link) was built in 2006.

With respect to unbundling, the situation in England and Wales is as follows. Network managers are required to be juridically separated from commercial activities and must also have a separate location and identity. It is allowed, however, that a network manager is part of a larger energy firm in which other legal persons are involved in competitive activities (AER, 2004).

Historically, France and Belgium have an electricity and gas sector that is even more centralized than used to be the case in the Netherlands. Since shortly after the Second World War, Electricité de France (EdF) and Gaz de France (GdF), both state-owned companies, take care of the production and transport of electricity, respectively the import and transport of natural gas. In Belgium there also used to be two national monopolies, Electrabel and Distrigaz, which are privately owned companies. These companies have been restructured in order to meet the requirements of the Electricity Directive and the Gas Directive, but effectively retain a near complete monopoly over the electricity and gas supply in both countries. Electrabel was purchased by Suez and is now part of GDF Suez, while Distrigaz has been purchased by the Italian company ENI.

---

51 In 2008, GdF merged with Suez to form GDF Suez.
In France, the trade unions provided formidable opposition to attempts to restructure the state energy companies. There is also a substantive reason for the electricity monopoly: a majority of electricity is generated in nuclear power plants. There are large scale advantages with respect to purchasing fuel, waste management and safety. Due to the lack of competition, the French spot market for electricity is not representative for the electricity prices in that country. Belgium does not have a power exchange. Electrabel also is a player in the Dutch market. As owner of the former EPON generation company it controls about 20% of the Dutch generation capacity. As a result, Electrabel has a market share of about 60% of electricity generation in Belgium and the Netherlands together, which may explain why this company is a strong proponent of merging the Belgium and the Dutch electricity markets.

The French transmission network is only unbundled to the extent that the Electricity Directive requires: until now, the network is separated administratively, which means that the network manager (RTE) is a part of the EdF and (as yet) not a separate legal person. In Belgium, the network manager, Elia, has been separated juridically from market activities. Elia is owned partly by the state and is partly privately owned, with Electrabel (now part of GDF/Suez) having a majority share. The distribution network managers in France are only unbundled administratively from competitive activities. Most are publicly owned. The Belgian distribution network managers are juridically unbundled and also largely in public hands (AER 2004).

Traditionally, Germany has a different market structure than Belgium and France. Mostly private companies used to have regional monopolies within which they took care of generation and imports, transmission and usually also a part of distribution. In addition, municipal utility companies had an important role in retail, as well as in distribution and in generation. Initially, Germany took a somewhat passive stance with respect to liberalization. Formally, the market was opened 100% in 1998, but the development of effective competition was not stimulated. In particular it was left to the companies themselves to establish the network tariffs. As a result, they have the possibility to deter new entrants to the market (for generation and supply). The companies can use their income from the networks to cross subsidize the electricity price. As a result, effective competition is absent and the market is difficult to access for foreign companies (AER, 2004). With respect to unbundling, in Germany only administrative unbundling was required. In 2005, Germany created an energy regulator, as one of the last countries in Europe. The Bundesnetzagentur has done much to improve competition in Germany. The unbundling of E.On (under pressure from the EU) has

---

52 A consequence of Electricity Directive 2003/54/EC is that RTE will need to be separated juridically from EdF.
further contributed to the accessibility of the German electricity market to new market entrants.

3.5 Challenges

The following sections will present a brief overview of the main challenges, in the view of the authors, to the European electricity supply industry. They are grouped into three categories: implementation issues, issues related to market structure and competition policy, and investment issues.

3.5.1 Market integration

What are the obstacles to the integration of electricity markets in the EU? They can be divided into physical, economic and regulatory obstacles. Physically, congestion of interconnectors limits trade between member states. A major goal of the EU is therefore to expand interconnector capacity (cf. EC, 2005; EC, 2006). Economically, differences in the availability of primary energy (hydropower, gas network, deep sea ports for the supply of coal) and historical differences in the choice of primary fuel cause inherent price differences between countries that will not disappear easily. Differences in market regulation, such as the regulation of networks and the degree of unbundling, also cause price differences. Differences in other aspects of regulation, for instance taxes, environmental levies, and the regulation of the gas market, further detract from a level playing field for competition. Consequently, the development of European electricity markets is slow, both formally, with respect to the implementation of the directives, and materially, in terms of competitiveness and the integration of national markets.

As top-down efforts to integrate European markets were not yielding enough progress, a shift was made to regional integration. From the beginning, the Nordic markets were most strongly integrated. The need to integrate the western continental markets was also high, though. This need was demonstrated by the large blackout of 2003, when a lack of coordination between TSOs caused the UCTE grid to fall apart into three asynchronous grids and electricity service to millions of people was disrupted.

3.5.2 Competition issues

The EU Sector Inquiry (EC, 2006) presents four main concerns relating to the competitiveness of the market: market concentration, vertical foreclosure, lack of transparency and price issues. Lack of

---

53 In 2009, the Union for the Coordination of Transmission of Electricity was superseded by ENTSO-E, the European Network of Transmission System Operators for Electricity. The UCTE covered continental Europe, whereas ENTSO-E also includes Nordel and the British islands. The function of the UCTE and of ENTSO-E is to facilitate cooperation between TSOs.
transparency is a matter of harmonization and implementing requirements and facilities for data reporting.

According to the EU’s Sector Inquiry (2006), market concentration is one of the primary challenges to creating a competitive electricity market. Currently, there is significant local market power in most of the EU member states. Due to the fact that most electricity companies used to have a regional monopoly, ownership of generation plants is geographically clustered. Creating effective competition is therefore a challenge.

The EU’s limited ability to deal directly with local market power (e.g. through forced divestiture) may be an important reason why it promotes increases in cross-border capacity, the idea being that competition within member states would be increased by providing access to foreign companies (cf. EU, 2005). However, this is not necessarily an effective solution, as it takes a long time to build new transmission capacity and cross-border mergers may quickly undo the effect of more capacity. It also is not necessarily an economically efficient solution, as the cost of building network capacity may be much higher than the cost of breaking up an incumbent. Interconnector capacity expansion therefore appears a second-best way of increasing the competitiveness of European markets. In principle, interconnector investment is only rational when it is a cheaper option of providing certain customers with electricity than building generating capacity locally, or where the benefits of increased network stability and security of supply outweigh the costs.

Another issue of competition that is discussed in the Sector Inquiry is vertical integration. This reduces the liquidity of the wholesale market, which makes it more difficult for new parties to enter the market. Along the production chain of primary fuel markets, generation, networks and end user supply (retail), different options for vertical integration exist. Vertical integration between network companies and generation or retail are discouraged by the European requirement to make network operators at least legally independent. Vertical integration between generation and retail is common and is increasing. This is a self-reinforcing trend: as vertical integration increases, the liquidity of the wholesale market decreases and the pressure increases for the remaining companies to integrate vertically too. Moreover, integration of generation and retail significantly reduces price risk for both parties.

Similarly, there is an incentive (not mentioned in the Sector Inquiry) for electricity generation to integrate with gas production, transport and/or trade, in order to reduce the fuel price risk for electricity generation and secure the sales of gas. At the upstream end, there is a need for large, integrated gas companies that have sufficient negotiating power for dealing with the likes of Gazprom. Long-term contracts are usually also needed for handling the investment risk associated with the construction of large infrastructure facilities such
as LNG terminals and pipelines. At the downstream end and in the electricity market, this trend conflicts with the objective of creating liquid, competitive markets. In the electricity market, expectations are still much higher with respect to the potential for developing effective competition. However, when purchasing gas from the state-owned monopolies of the gas-selling countries, large European energy companies may provide benefits, also to end users, through their negotiating power. A prime example of this conflict between policy goals is the E.On/Ruhrgas merger, which was blocked by the German competition authority because it would create too much market power, but overruled by the government because it wanted to strengthen its position on the international gas market.

**Price issues**

The reason that prices are considered an issue in the Sector Inquiry is that they have been rising consistently in much of Europe during the past two years, as is indicated by Figure 4 (EC, 2006). In countries where gas is the marginal fuel, there is a fairly close correlation between gas and electricity prices. The same correlation cannot be found between coal and electricity prices in countries that mostly rely on coal, raising concerns about the exercise of market power in the Sector Inquiry (EC, 2006).

![Electricity prices on the rise all over Europe](image)

**Figure 4: Rising electricity prices were a reason for the Sector Inquiry**

Source: Sector Inquiry (EU, 2006), based upon data from Argus Media, Platts, and NordPool

3.5.3 **Investment and innovation**

An unsolved question is how to provide network operators with incentives for optimal investment in their networks. In fact, even the question what would be an optimal network is unanswered. Network development is characterized by a strong path dependence, high capital costs and long lead times. There are significant uncertainties about the actual demand for network services by the time new
capacity has been realized, which in turn contribute to investment risk. And even if an optimal expansion policy could be determined, it is a question how to provide a network manager with incentives to provide it, without risking over or under investment. In practice, transmission networks are often subject to a form of cost-plus regulation. Being able to pass the costs along to the consumer, this provides the network manager with an incentive to invest more than necessary. Therefore large expansion projects are usually reviewed by the regulator, who makes a cost-benefit analysis. While this does not lead to an optimal network, it does provide an acceptable outcome – given that we do not know the optimum anyway – and it is the way that it always used to be done.

The share of renewable energy sources for electricity has been hovering around 14% for the past several years. Clearly, the policy goal of 21% in 2010 will not be reached (Observ’ER, 2005). There are three aspects to this issue. First, apparently the current support schemes for renewables (which currently are not competitive) need to be strengthened. Second, there is a need for increased coordination between national support schemes, to remove double counting of renewable energy when it is traded between countries. Third, as the volume of wind energy increases – to date the most successful renewable energy technology – the increasing impact upon the electricity network needs to be managed.

The contribution of renewable energy sources is insufficient, at least in the next decade or two, for meeting the EU’s goals with respect to reducing carbon-dioxide emissions. The Kyoto Protocol has a limited impact upon the electricity sector, as the reduction targets are not very strict, but the EU’s longer-term goals will require other measures, such as sequestering carbon-dioxide or increasing the use of nuclear power. These measures require significant investments by power generators. However, any power plants that are built after today will operate entirely or almost entirely after the end of the Kyoto ratification period in 2012. It is unclear whether there will be a successor treaty to the Kyoto Protocol and if not, what climate change policy will look like. As a result, the benefits of investing in CO₂ abatement measures are uncertain. At the same time, investing in fossil-fuel plants entails a risk of high future CO₂ costs, either in the form of costly emission credits or taxes, or in the form of having to implement retrofit abatement measures.

Last but not least, one of the big questions is whether markets provide adequate investment incentives for investment in generation. The current high level of reliability can only be maintained if there is some excess generating capacity nearly all the time, for available

---

54 Even if formally the TSO is subject to a form of incentive regulation, such as revenue-cap regulation, network expansion projects often are exempt from this regulation.
generating capacity must be at least equal to demand. Who will provide this capacity? Experience to date does not provide us with conclusive evidence whether competitive markets provide adequate investment incentives. At the time that most electricity markets were restructured and competition was introduced, these markets had ample reserve margins, if not downright overcapacity. It took a number of years for demand to catch up and only recently has the need for new investment developed in many markets. The power crisis in California was partly due to insufficient investment, but the market design was so flawed that this case provides no evidence of a general shortcoming of competitive markets.

### 3.6 The role of market design

The aim of this course is to assess how these issues are affected by the design of the market, this is, the organizational structure, the regulation and the allocation of responsibilities in the electricity industry. However, not all of the above issues are affected by market design. Harmonization and transparency are matters of implementing and enforcing the rules, whereas the question of market design is what the rules should be. Although these issues are politically related, conceptually they are a separate matter. Moreover, in this reader we step back from the current conditions in the European electricity markets and consider the fundamental choices. Therefore we will leave these issues aside. Similarly, we will consider competition policy a different issue, unrelated except to the extent that the design of the market may affect the ability to exercise market power. Finally, the question of how to stimulate the development of renewable resources is largely separate from the issue of market design. In principle, support mechanisms such as feed-in tariffs or renewables quota obligations can be introduced independently from the market design. Therefore we will discuss renewables policy only where market design directly affects it.
4 Market Design

4.1 Introduction

Many electricity markets around the world have been or are in the process of being restructured with the purpose of introducing or expanding competition, but few have actually reached a state that could be described as commensurate with the economics textbook ideal of a liberalized, competitive market (cf. Joskow, 1996; Stoft, 2002). Perhaps the electricity markets in the U.K., Argentina, Texas, New Zealand, Chili and Alberta have come closest to this ideal, at least in terms of market design. However, they also have many idiosyncrasies, vestiges of their pre-liberalized state or results of political compromises, such as publicly-owned competitive companies, preferred treatment of incumbents, inadequate network regulation, price regulations or limited competition. Most electricity markets are still farther removed from the textbook ideal and are somewhere in between their former pre-liberalized state and retail competition (Joskow, 2006; Newbery, 2005a,b; Victor, 2006; Rudnick et al., 2005).

If a function can be performed adequately by competitive firms, there is no need for public ownership. Nevertheless, public ownership remains common in restructured power markets, often so the state can retain a strategic position. For instance, in many developing countries the single buyer remains a public entity. If the single buyer also owns some of the generating capacity, this leads to a situation in which it competes with independent power producers (Dehdashti, 2004). In many OECD countries, the state, provinces or municipalities maintain their shares in power production and networks.

Another phenomenon is that markets have been restructured with all the features of a competitive wholesale market, but continue to be dominated by a single party – often the former monopolist. Examples are Electricité de France, Electrabel in Belgium, or CFE in Mexico. Whereas ENEL, in Italy, was forced to divest a substantial amount of its generating capacity a few years ago, currently there appears to be a tendency in Europe to create ‘national champions’ – large power companies that are expected to protect national interests in an increasingly international power market (Thomas, 2003; Haas et al 2006).

---

The main focus of this chapter is: How can the current divergence in the design and structure of liberalized electricity markets be explained and what are the prospects for convergence and agreement upon a single market design?

A promising approach for exploring the situation in these hybrid electricity markets is provided by Institutional Economics, a sub-discipline of economic science that deals with the evolution of institutional arrangements in markets (see, for example, North, 1990; Williamson, 1998; Glachant and Finon, 2000; Victor et al., 2006). In this theory, the behavior of market actors, as regards pricing, production, resource allocation, investment, strategies of horizontal and vertical integration and so on, is assumed to be influenced by a market-specific body of rules and conventions: the institutional arrangements. Of course, the traditional economic perspective also relates the conduct of actors and their performance to structural characteristics of markets, like the degree of concentration, the shape of production functions, regulation, etc. Yet, these characteristics are considered as a given, exogenous to a market.

What distinguishes the approach of institutional economics is that it considers this body of rules as being an endogenous part of the market. Institutional arrangements are shaped by a path-dependent interaction between political, economic and physical factors, driving deliberate interests, choices and strategies of the policy makers, the firms in the industry, (groups of) consumers and other parties involved, like international organizations and NGOs. This implies that an explanation of why certain types of hybrid markets have developed – and their future – should include the interaction between these factors and the resulting behavior of the actors in a specific market. Main elements in the institutional economy of power system restructuring are the situation at the start of the restructuring process, policy decisions regarding market design, related policies such as competition policy, fuel policy and environmental policy, and a number of exogenous factors that constrain the decision space, like the scale of systems and the availability of primary energy resources.

4.2 Context of the restructuring process

The physical situation in a country provides a set of relatively ‘hard’ constraints. A crucial element is the presence of indigenous energy resources, such as hydropower, coal, natural gas and oil, or – at the other end – dependence upon other countries for energy resources. Market size and degree of isolation matter too. Small isolated systems, in small countries or on islands, like Iceland, Malta or in the Caribbean, cannot efficiently support multiple competing generating companies as a consequence of the relatively large minimum efficient scale of the several types of generation units. This also impedes the use of specific technologies or fuels, such as large scale coal plants. The geographic distribution of demand also plays a role: in thinly populated areas or small, remote concentrations of electricity demand.
it may be difficult to create competition in supply (see also Weinmann et al, 2004).

A second set of constraints are macro-economic characteristics such as the level of economic development, the rate of demand growth and the availability of investment capital. These factors influence the acceptability of changes in tariffs or prices to different categories of users, the need for investment and financing options for system expansion and/or rehabilitation. Three obvious categories of countries are, firstly, developing countries with a relatively stagnant economy; secondly, countries on the path of economic development and industrialization, and, thirdly, the OECD countries.

The third category of constraints derives from the institutional and socio-political environment of a power system. North (1990), Williamson (1998), Glachant and Finon (2000) and Finon (2003) explain how informal institutions such as culture, traditions and values affect the development of formal institutions, such as property rights, legislation, regulation and the role of the (federal) state in the economy. De Vries and Correljé (2008) discussed how formal institutions can be divided into general institutions, such as the polity, the judiciary and the bureaucracy, and sector-specific institutions, such as sector legislation and regulation and jurisprudence, which are the main tools of market design. Arguing that the freedom of action for those who are in control of the reform process and the need to coordinate different aspects of the reform process are essential to the success of a restructuring process, Glachant and Finon (2000) consider the power of the central government a key factor with respect to the success of market reform.

Table 1 provides an overview of the physical, macro-economic and institutional constraints. Together they determine to a large extent the solution space that is available to governments who wish to restructure their power sectors. Within the context of these constraints, governments need to find a balance between their own multiple objectives and those of the energy sector and of consumers.
### Table 1: Context of the restructuring process

<table>
<thead>
<tr>
<th>Factor</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical factors</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Natural endowment with energy sources</strong></td>
<td>Presence or absence of primary energy sources drives the choice of primary fuels, the technical and economic characteristics of the sector and drives interests and policies.</td>
</tr>
<tr>
<td><strong>Physical size of the market</strong></td>
<td>Due to scale effects, small markets are likely to be more concentrated. Larger markets may constitute a number of separate subsystems, with their own economic and institutional structure.</td>
</tr>
<tr>
<td><strong>Geographic distribution of demand in relation to network capacity</strong></td>
<td>Relatively dispersed demand and/or limited network capacity increase the likelihood of network congestion, which results in market fragmentation and limits competition.</td>
</tr>
<tr>
<td><strong>Economic factors</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Level of economic development and growth</strong></td>
<td>Influences demand growth, the potential for investments and institutional stability.</td>
</tr>
<tr>
<td><strong>Growth rate of demand</strong></td>
<td>Capacity investment lead times are long. With a high growth rate, large volumes of capacity must be under construction. Market signals or regulation must be effective. Stable demand, on the other hand, limits the ‘room’ for new market entrants.</td>
</tr>
<tr>
<td><strong>Financing options</strong></td>
<td>Especially in developing or transition countries with a weaker economy financing options may be limited.</td>
</tr>
<tr>
<td><strong>Institutional factors</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Ideology</strong></td>
<td>General acceptability of and commitment to particular policies and institutions.</td>
</tr>
<tr>
<td><strong>Institutional stability and rule of law</strong></td>
<td>Facilitates investment and external funding; stabilizes and provides coherence in policies; helps align policy, regulation and the legal framework.</td>
</tr>
<tr>
<td><strong>Degree of institutional centralization and homogeneity</strong></td>
<td>The power of the central government influences the coherence of policies and their support in terms of regional, sectoral and social dispersion.</td>
</tr>
<tr>
<td><strong>Influence of stakeholders</strong></td>
<td>Strong stakeholders may be able to influence the reforms in their own interest.</td>
</tr>
</tbody>
</table>

4.3 Market design

The preceding section provided an overview of the motives for power sector restructuring. Now the analysis will turn to the options for achieving these goals, the market design variables. Basic choices are the pace of market opening, the degree in which competition is introduced, the market model, decisions to restructure horizontally or vertically and decisions to privatize (Glachant and Finon, 2000; Newbery, 2005a; Littlechild, 2006). An important means of achieving competition is by redesigning the market to make it attractive enough for new entrants to whittle away the market power of the incumbents. In designing markets, however, a trade-off is to be made between competition and the investment climate.

A relatively light form of restructuring of publicly owned industries is their corporatization. State-owned enterprises normally operate under the responsibility and control of a particular department with “soft” budget constraints. Placing the enterprise at arm’s length from the responsible public body, while imposing a hard budget constraint, is a first step towards enhancing the efficiency of the electricity industry and making it less responsive to political and interest group capture.

A next step, introducing some competition without structurally addressing the state monopoly, is to allow independent power producers (IPPs). They sell to the state-owned monopolist (called the single buyer), who often continues to generate electricity itself (Heller and Victor, 2004). While this allows the state monopoly to dominate the system, the IPPs may provide benchmarks for performance and may increase their influence over time.

The following step is to remove the single buyer and create a wholesale market in which there is wholesale competition between a number of generators, supplying large customers and distribution companies with a retail franchise. The final step is the introduction of competition at the level of retail customers. To an increasing extent, these categories of market design allow for competition (Hunt and Shuttleworth, 1996). It is possible to open up the market to full retail competition at once, but it is also possible to gradually introduce competition by moving through these different models. The pace of restructuring is therefore a variable. Retail competition was often regarded as the desired end state of restructuring (it is, for instance, required by the EU in Directive 2003/54/EC). Newbery (2002) also pointed out advantages of retaining the retail franchise, such as retail companies being able to enter into long-term contracts for generation.

With respect to the wholesale market model, there is a fundamental choice between integrated markets and decentralized markets (Hunt, 2002). Integrated markets are preferred in the USA. In these markets the system operator operates a mandatory pool, in which the physical and economic aspects of electricity trade are strongly connected. In
decentralized markets, the preferred model in Europe, the system operator only has a technical function and supply and demand meet elsewhere, either bilaterally or in voluntary power exchanges.

**Privatization**

In the case of restructuring publicly owned utilities, privatization becomes an issue, as it is generally not considered necessary nor desirable for government to be involved in competitive activities. Nevertheless, it is not uncommon for states to have their publicly-owned companies operating in competitive electricity markets, perhaps due to the strong public goals with respect to this sector. The choice of approach has important consequences for the resulting vertical and horizontal structure of the electricity sector and for the behavior of the firms. Main variables are, firstly, the degree of privatization of the industry, ranging from a government monopoly via the admission of new privately financed firms into the sector to full privatization. Secondly, there is the issue of what segments may be privatized; generation, wholesale and retail trade, the networks? In some cases, like in the Netherlands, it is argued that the networks, as critical infrastructures, should remain in public ownership (see Künneke and Fens, 2007). The third element involves the character of the new private owners; will they be anonymous shareholders, national or foreign financing institutions, foreign electricity companies? A final aspect involves the timing of privatization, relative to other elements of a restructuring.

**Breaking up incumbents**

If the goal of restructuring is to introduce competition, it may be necessary to break up existing companies into smaller ones or to force the incumbent to divest some of its assets. Indeed, currently a debate is taking place within the EU about the question as to what extent large cross-national firms should be allowed, as a means to enhance security of supply and investment potential, or whether these former national champions should be divided up in a number of potentially competing firms. Even if the market design is perfect, the potential benefits of restructuring are not likely to be obtained if the market continues to be dominated by the former monopolist. Fringe competition may provide some benefits, but due to economies of scale and the long life cycle of key assets such as power plants, it is not likely to lead to a level playing field for competition very quickly. Therefore horizontal unbundling is a key element of restructuring if there are not already multiple potential competitors present.

**Position of networks**

The way the networks are regulated will fundamentally affect the network owners’ investment policies and therefore impact the adequacy with which demand growth and shifts in the pattern of supply and demand are met. Depending on the choice for a model, the issue of the position of the transmission and distribution networks’ monopoly becomes important. The regulation of network tariffs, the provision of regulated or negotiated third party access and the degree of unbundling of the networks from competitive generation, trading and retail activities strongly influence the overall effectiveness of competition in the market:
• The quality of access regulation and the level and structure of tariffs affects the competitiveness of the supply of power and the development of trade.

• The degree of unbundling also plays an important role in this respect, as it keeps incumbent owners of networks from obstructing access for new entrants and avoids cross-subsidization of competitive by non-competitive activities.

• Incentive regulation of transmission and distribution networks directly impacts the level of transport costs as a component of overall supply cost.

In addition to access to transport, there are certain essential system operation functions that need to be provided. The main functions are scheduling and dispatch of transmission and distribution, balancing (in case of decentralized markets), congestion management and ancillary services (such as black-start capacity and voltage control). These functions can be designed in multiple ways, but because of their relation with network management they are often provided by the transmission network manager (Knops, 2003).

Trade between different electricity markets often is an issue. The connection with neighboring electricity systems may have a significant impact upon the competitiveness of the market and the incentives to market parties. At the same time, it exposes the industries of connected countries to competitive forces, a phenomenon that may or may not be appreciated.

In many markets, wholesale and/or retail prices are regulated. In this respect a balance needs to struck between the interests of consumer groups, who often embody specific political power, and the incentives to the industry for providing sufficient investment to cover future demand. It is still debated whether competitive electricity markets without price restrictions provide sufficient investment incentives, or whether consumers would be better off with some kind of capacity mechanism. If there are price restrictions, theoretically a capacity mechanism is necessary to compensate generating companies for the foregone revenues, otherwise they will under invest (Hogan, 2005, Stoft, 2002). A capacity mechanism may also be needed, or beneficial, for other reasons (De Vries, 2004).

An overarching issue is the role and position of the regulatory function. At what level should the regulator be placed? Local, provincial, national of even supranational regulatory bodies may exist. The choice has consequences for the relation between the regulator and the regulated industry and its independence from other parts of the public realm. It also affects the degree of detail, specificity and generality in which regulatory problems can be solved. A second aspect is the balance between sector-specific regulation and the application of general competition law. Both approaches have an *ex ante* component, addressing structural sector characteristics, prescribing behavior and evaluating plans for mergers and
acquisitions, and an *ex post* component for the monitoring and mitigation of abuse of market power. Consequently, the position of the regulator(s) and competition authorities within the policy making arena and *vis-à-vis* interest groups has consequences for the transition towards a competitive market and the design of that market and, ultimately, for the allocative and dynamic efficiency of the sector and the market outcomes in terms of efficiency and welfare.

The organization of the electricity supply industry prior to the restructuring process affects a number of these decisions by providing default choices. The number of generating companies at the beginning of the restructuring process, the degree to which they are integrated with network companies, the ownership structure of these companies, generating and network capacity, the capacity of network links with neighboring systems, whether there already is a regulator (e.g. for the regulation of private utilities) determine the starting conditions of the restructuring process. These conditions are relevant, as will be seen later in this chapter, because the restructuring process is subject to path dependence. Table 2 provides an overview of the market design variables.

### 4.4 The dynamics of restructuring

The solution space of governments with respect to restructuring – the range within which governments can choose the market design variables that were discussed in Section 4.3 – is restricted by the situation at the outset of the restructuring process and by the ‘hard’ constraints that were described in Section 4.2. Within the context of the different constraints, governments need to find a balance between multiple objectives with respect to the energy sector. To a large extent the specific balance between the policy objectives is a function of a country’s socio-political context, involving ideology and the representation of interest groups such as specific categories of energy consumers, producers of indigenous energy resources and their staff, the citizens at large and components of the central and local administrations (see also Heller *et al.*, 2004).

Strategic decisions by the firms, finally, determine the space within which they make their operational decisions, which lead to the market outcomes (Williamson, 1998). Based upon the performance of the market with respect to the government’s policy goals – and responding to public perception – government may adjust its policies (Willman *et al.*, 2003; Correljé, 2005).
### Table 2: Market design variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Degree of market opening</strong></td>
<td>Corporatization of a state monopoly, single buyer, wholesale market and competition in the retail market allow for an increasing degree of competition, but involve increasing transaction costs and requirements as regards the economic and institutional structure.</td>
</tr>
<tr>
<td><strong>Pace of market opening</strong></td>
<td>Being a leader or a follower vis-à-vis neighboring countries/states, or higher (EU, USA Federal) policy.</td>
</tr>
<tr>
<td><strong>Integrated versus decentralized market</strong></td>
<td>Integrated markets with mandatory pools, reduce transaction costs, but combine both economic and physical control over the system in the hands of a single party, potentially facilitating governance.</td>
</tr>
<tr>
<td><strong>Public versus private ownership</strong></td>
<td>Public ownership provides a means for direct control but entails policy captivity may impede effective regulation and limits financial resources. Private ownership requires political and institutional stability and regulatory commitment.</td>
</tr>
<tr>
<td><strong>Competition policy and horizontal unbundling</strong></td>
<td>Influences the competitiveness in market segments, trading off economics of scale and scope.</td>
</tr>
<tr>
<td><strong>Network unbundling</strong></td>
<td>Unbundling influences the independence of network managers and their interest in providing equal conditions for all network users.</td>
</tr>
<tr>
<td><strong>Network regulation of network tariffs and access conditions</strong></td>
<td>Influence the conditions for competition and the pressure upon network managers to work efficiently.</td>
</tr>
<tr>
<td><strong>Congestion management method</strong></td>
<td>Affects trade opportunities between regions.</td>
</tr>
<tr>
<td><strong>Arrangements with neighboring networks and interconnector congestion management</strong></td>
<td>Market integration may enhance the competitiveness but may also cause higher prices in the exporting country.</td>
</tr>
<tr>
<td><strong>Balancing mechanism (in decentralized markets)</strong></td>
<td>Balancing mechanisms affect cost and revenues of type's generation, especially for intermittent sources and influence entry.</td>
</tr>
<tr>
<td><strong>Wholesale and end user price regulation</strong></td>
<td>Protects consumers, at the expense of investment stimuli to the industry.</td>
</tr>
<tr>
<td><strong>Capacity mechanism</strong></td>
<td>Different types of capacity mechanisms exist in order to stimulate investment in capacity.</td>
</tr>
<tr>
<td><strong>Position of regulator</strong></td>
<td><em>Ex ante or ex post regulation.</em></td>
</tr>
</tbody>
</table>

Source: Correljé and De Vries (2008)

Combining the above elements leads to the conceptual framework that is represented in Figure 5. This figure summarizes the relations between the different factors that influence the development of electricity markets. The market design process takes place within an institutional context: the process is influenced by informal institutions
such as culture and values and constrained by formal institutions such as international treaties and the constitution. This context influences the choice and relative weight of the policy goals, as well as the general effectiveness of policy-making and implementation, as a function of the ‘power’ and legitimacy of the government (Williamson, 1998; North, 1990). The institutional context also manifests itself through decisions of the past, as the institutional context is relatively static.

Figure 5: Conceptual framework

4.5 Hybrid markets and patterns of restructuring in Europe

With the exception of the U.K., the Netherlands and Scandinavian countries, the restructuring of electricity markets in Europe is mainly driven by EU Policy, meaning that there is a limited drive for market-based reform in these countries. The European Commission (EC) started with a rather loosely defined model of sector reform – with clear Anglo-Saxon principles – but has imposed increasingly strict requirements with respect to market design (CEC, 1996 and CEC, 2003). The intention of the EU was to increase economic efficiency through the introduction of competition and thereby reduce end-user prices. With respect to the goal of sustainability, restructuring is also intended to facilitate the application of economic instruments, most notably the CO₂ emissions trading system. The goal of security of supply, the third general policy goal, is expected to be served by restructuring, as the unbundling of the industry and the introduction of competition where possible should lead to optimal investment in each link of the value chain. Moreover, well-functioning markets...
should promote diversity, which provides resilience in case of disturbances.

Recent reports by the EC show that these objectives have not been reached (CEC, 2006a, 2007b). The Sector Inquiry provides a thorough analysis, supported by a wealth of empirical data, that shows that European electricity markets are far from integrated and in most cases also far from competitive (CEC, 2006b, 2007c). A main problem is that the restructuring effort focused on market design but ignored competition issues and horizontal integration in most member states. Consequently, many member states’ electricity markets are dominated by a very small number of companies. Other important shortcomings are:

- Insufficient unbundling of the transmission networks from generating companies. The EC now strives for ownership unbundling or, if that is not feasible, the creation of independent system operators (ISOs).
- A lack of effective regulation (in particular in case of incompletely unbundled networks) and too strong a focus by regulators on their national markets instead of on the development of the EU internal electricity market.
- A continuing lack of sufficient transparency. The availability of information varies significantly between member states.
- Inadequate capacity of the infrastructure between member states.
- Insufficient network security standards.
- Insufficient signals for investment in generating capacity.

Many of the ‘hard’ measures that are prescribed by the EC have been implemented only pro forma, which has not led to an effective single market or even a set of competitive national or regional markets. Member States strategically interpret and implement the directives and guidelines of the EU, their main concern being their own (future) position within the emerging EU energy market. Countries differ with respect to their commitment to, and faith in, competition, which appears to have an important impact upon the extent to which effective competition is created.

An important explanatory factor for the market situation in European countries appears to be the initial ownership structure of the power sector. Apparently, the existence of national monopolies, like in France, Portugal, Belgium and Greece, makes it difficult for a government to create competition. The only exception is the U.K. that split up its Central Electricity Generating Board. The general solution is ‘fringe competition’, the facilitation of ‘competitive’ entry alongside the monopoly. However, the continued alliance between the dominant public monopoly and the state has as a consequence that the development of new institutional arrangements and the regulatory body proceed only slowly. Price regulation needs to be maintained against the market power of the monopoly, but this often harms the interests of (potential) new entrants (Glachant and Finon, 2005).
If the government is confronted with a number of vertically integrated firms, that are publicly owned, whether by the state or by provinces and/or municipalities, the situation appears relatively straightforward. Unbundling, the introduction of access rules and tariffs and wholesale and retail competition can be established by political decision making. Main causes of resistance are conflicts of interest between public bodies at different levels and interest groups clamoring for consumer protection, environmental issues and the unions. Eventually, this state of affairs may progress and lead towards privatization. When the government needs to confront a number of private firms the situation is more complex. In the EU it can be doubted whether effective unbundling is legally possible, as it may be considered expropriation.

States that only reluctantly accepted liberalization may try to protect the interest of their ‘national’ firms through influencing its quasi-independent regulator, assuming that they will serve the country better than new entrants. This argument runs two ways. If new entrants are private firms, they are thought to be driven by profits only and ignorant of the public character of electricity supply. If the new entrants are foreign public firms, the situation is also suspect, as these firms are thought to give preference to their home market in case of supply problems.

European markets have witnessed the emergence of a number of large multinational conglomerates, the so-called seven brothers (see Thomas, 2003; Green, 2006; Haas et al., 2006) that extend their activities by taking over energy companies in their home country and abroad. These companies may be considered as (potential) competitors in the future pan-European market. Their purchasing power in the international fuel markets is often brought forward as an important asset with respect of security of supply issues, as it provides a counterbalance to the market power of Gazprom and the Middle Eastern gas suppliers. On the other hand, the EC recently announced plans to take on the dominant position of companies such as Electricité de France, E.On, RWE and ENI (EC, 2006; EC, 2007).

In response, governments of countries who are not home to one of the seven brothers feel an even stronger need to protect their power market against ‘foreign domination’, either by establishing their own mini-champions or by keeping the market relatively closed to take-over’s, for instance by a minimalist implementation of the Directives or by not privatizing their companies. So while these governments may sound supportive of the notion of introducing competition and may not oppose the corresponding changes to the market design, they may oppose attempts to really restructure the market. This situation causes inconsistencies, on/off policies and a slow development of an effective market. It also leads to competition between the EU member states’ “national” companies, which has led to the reintroduction of a great deal of economic nationalism in the
restructuring process, often veiled in terms of environmental protection, security of supply measures or public service obligations.

4.6 Analysis

In this section, we will review how the empirical evidence that was presented in Section 4.5 fits with the conceptual framework that was proposed in Section 4.4. While in general, the empirical observations appear to match our expectations, it is now possible to fill in the blanks in the conceptual framework.

Clearly restructuring processes are driven by a wide variety of objectives. Much of the literature describes projects that were meant to evaluate and support policies of particular organizations, such as the World Bank and regional organizations such as the EU, APEC (2000) and industrial associations, which is why these studies tend to focus on sets of relatively comparable countries with similar objectives. Without such a restricted focus, a rich variation in motives for restructuring and in the backgrounds of such processes emerges.

So whereas the introduction of competition is often presented as a logical and inescapable consequence of the superiority of ‘the market’ as economic coordination device, it can be observed that a range of different motives exist in different countries and at different times. Shifts in drivers and process dynamics lead governments and market parties to adjust their strategies (see Green, 2006). Hence, when evaluating the progress in restructuring and the emergence of hybrid markets, it is necessary to distinguish the initial drivers behind the process from later adjustments to them. Indeed, Rudnick et al (2005) and Sioshansi and Pfaffenberger (2006), respectively, refer to “second stage” and the “reform of the reforms”.

The impact of related policies, such as environmental policies, upon market design appears to be limited. These policies are typically achieved through other means, such as standards or taxes. The main exception is security of supply of primary fuels, which may run counter to the drive to create a market with many small competitors, as the perception may exist that the market power of large firms is needed to secure good fuel import contracts.

In summary, with respect to the motives for restructuring, empirical analysis shows that governments may either or not:

- support a neo-liberal program in order to enhance the efficiency of public sector management in the broadest sense by privatization and restructuring;
- seek to enhance the efficiency of their power systems specifically, to achieve lower supply prices, to improve quality of supply and to diversify their fuels;
- seek private investment in their electricity industry, to facilitate an expansion of their supply potential and therefore accept some degree of restructuring, involving most likely higher prices to (groups of) consumers;

Empirical observations vs. the framework

Wide variety of objectives

“Reform of the reforms”

Impact limited

Summary
• seek to *achieve market integration* with (groups of) adjoining countries and accept some degree of reduction in their autonomy over ‘public interest’ sectors, in exchange for advantages for other economic sectors;

• support integration with neighboring markets, for the purpose of *exploiting (potential) advantages in exporting or importing power*, or by engaging in the electricity industry of associated countries;

• Restructure merely to *satisfy requirements of other authorities*, such as the EU, the FERC, the IMF or the World Bank.

---

**Level of economic development**

With respect to *exogenous factors* that were proposed in Section 4.2, two stand out, apart from the cultural and ideological commitment to competition as a means of maximizing welfare. The first is the level of economic development, in which a high income per capita generally correlates with a lower rate of growth in electricity demand and the presence of a relatively stable legal and policy framework. Lower income countries often experience higher *rates* of growth in their electricity demand, but also experience higher risks of economic and institutional instability. Consequently, investment risk is higher and they face greater difficulty in attracting sufficient investment capital. Indeed, a competitive energy-only market significantly raises investment risks. To reduce this risk, the introduction of competition can be limited to wholesale competition or be accompanied with a capacity mechanism (De Vries, 2007).

---

**Indigenous energy sources**

A second important exogenous factor is the availability of indigenous primary energy resources, like hydropower, coal, petroleum or gas. Hydropower fundamentally changes the dynamics of the market, as peak capacity is cheap, but the total amount of energy stored in the reservoirs needs to be spread out over the dry season. Coal exploitation often is associated with employment and social issues. Petroleum production generally yields high revenues to the state, thus reducing financial pressures to restructure, while supporting particular political interest coalitions in maintaining the status quo. The presence of (un)tapped natural gas reserves may support pressures to restructure the power industry, as it promises low (capital) cost and an efficient expansion of generation capacity. Here geopolitics and the organization of the gas industry are important variables. Countries that are endowed with sufficient indigenous energy sources show a stronger tendency to pursue effective competition with a number of smaller players. In contrast, in Europe, resistance to privatization and horizontal unbundling is sometimes linked to concerns about fuel import dependency, in particular with respect to natural gas (Correljé and Van der Linde, 2003).

---

**Path dependence**

Due to path dependence, history matters. Decisions made in the past affect future options. Therefore the situation at the start of the restructuring process has a significant impact upon the restructuring process and its outcome. Most important is whether at the outset of restructuring there is a national monopoly in the power sector or a number of regional firms. Glachant and Finon (2000) consider the
presence of an incumbent monopoly the ‘Achilles’ heel’ of restructuring as it will be difficult for a government to split it up in order to create competition, especially if the company is privately owned. A likely outcome is the facilitation of ‘competitive’ entry alongside the incumbent. Empirically, this situation is observed in developing countries, where IPPs were encouraged, and in some EU countries. If the alliance between the (public) monopoly and the state is continued, new institutional arrangements and a regulatory body develop only slowly. To mitigate the market power of the monopoly, price regulation needs to be maintained, but this often harms the interests of (potential) new entrants. Also with respect to vertical unbundling, ownership is an important factor. If incumbent, vertically integrated firms are publicly owned, unbundling, the introduction of access rules and tariffs and wholesale and retail competition are more or less easily established through the political process. In addition to opposition by the incumbents, main sources of resistance to restructuring may be public bodies at different levels and with different interests, consumer protection interest groups, environmental advocacy groups and, last but not least, trade unions.

When a government wishes to restructure private firms, the situation is much more complex. Unbundling often involves a struggle over compensation payments for ‘stranded cost’. Moreover, the incumbent firms will try to maintain their oligopolistic position by obstruction of the market or by mergers, or they may be taken over by foreign firms. On the other hand, if the state only halfheartedly pursued liberalization in the first place, it may try to protect the interests of the ‘national’ firms, for instance because they are presumed to serve the country better than foreign parties, the more so if these firms have other public tasks. This situation causes inconsistencies, on/off policies and a very slow development of an effective market.

Taking all these issues into consideration, what is the prospect for agreement upon a single market design? To the extent that exogenous factors or fundamental differences in policy objectives are the causes of differences in the market design, they will likely have a lasting effect and multiple market designs are likely to continue to coexist. Examples are markets in which concerns for security of supply, particular physical conditions or high investment risk have prompted government intervention. On the other hand, when variations in market design are due to differences in opinion regarding the means, in a context of similar policy goals and similar exogenous conditions, one would expect that mutual learning would eventually lead to convergence of the market design. North (1990), however, warns that this may be a slow process. The bounded rationality of the actors contributes to path dependence in which preexisting patterns are replicated with only modest alterations. This may explain, for instance, the preference for integrated markets (with mandatory pools) in the USA versus the decentralized markets of Europe, as well as the presence of capacity markets in the USA and their absence in Europe.
### 4.7 Conclusions

This chapter has presented a framework for explaining the many differences in the approach to electricity market restructuring around the world. A number of policy objectives exist with regard to power market restructuring; the particular institutional context has a strong influence upon which policy objectives are chosen and how they are prioritized. Implementation of these objectives takes place through two main policy areas: the market design process and competition policy (including the policy with respect to horizontal unbundling of dominant incumbent firms). This process is constrained by external factors such as the economic climate and the physical situation in a country. In addition, the situation at the outset of the restructuring process influences the process itself. Due to path dependence, this influence may be long-lasting.

Feedback exists in the sense that governments observe the performance of the electricity market and respond by adjusting their policies. This feedback is limited by time lags, incomplete data, bounded rationality and the fact that much of the feedback represents the perspectives of lobbyists. The limitations of the information feedback loop give rise to path dependence: given uncertainty about the best direction for policy, decisions are favored that generate the least resistance, which means those that depart the least from the status quo. A second cause of path dependence may be that the government is not able to overcome resistance to change by vested interests.

This framework has helped to explain the diversity of restructuring processes and market designs that can be observed. The prospects for future convergence of these different market designs hinge upon a number of factors.

- The policy goals and their relative weights vary. A key factor is whether the restructuring of a country’s power sector is ideologically and politically motivated or whether it is more or less forced upon a country by a higher authority, the latter often leading to a halfhearted and unstable approach.
- Due to path dependence, the situation at the start of a process of restructuring significantly influences the potential for competition, especially the degree of concentration and whether or not the industry already is privatized.
- Exogenous factors, such as physical scale, endowment with natural resources and macro-economic conditions, may have a dominant influence upon the design of the market. Countries that are endowed with sufficient indigenous energy sources tend to be more inclined towards pursuing effective competition with a number of smaller players than countries who are severely dependent upon imports of fossil fuels.

When exogenous factors or fundamental differences in policy objectives are the causes of differences in the market design, there is
little prospect for convergence of market designs. Otherwise, if the circumstances and policy goals are similar, mutual learning should lead to convergence, but this process is limited to similar markets and will be slow due to the imperfections of the policy feedback loop.
5  Distribution networks in the Netherlands

5.1  Introduction

In this chapter, the focus will be on the position of the managers of the distribution networks for electricity and gas in the Netherlands. Besides the distribution network managers, the managers of the national high tension network for electricity, TenneT, and the manager of the national gas transport network, GTS, have their own very different position, caused by the fact that they are not only network managers, but also system operators. (See Sections 2.2 and 2.4). Several years ago, the choice was made that TenneT should become 100% state-owned. In the Fall of 2004, the same was decided for GTS. The discussion in this chapter therefore does not concern the positions of TenneT and GTS, but only the managers of the distribution networks. The position of GTS and its mandate will be discussed as part of the analysis of the gas sector in Chapter 0. The many particular functions of Gasunie and its public-private character merit separate treatment.

5.2  Network regulation

Management of both electricity and gas networks can be considered from two perspectives. First is the perspective of the individual consumers: can he obtain access to the network and under which conditions? How content is this consumer? Second is the perspective of the overall regulation of the network manager. At issue are economic regulation and also the regulation of network quality. How content are ‘we’ about the performance of the network?

The central issue with respect to the management of electricity and gas networks is the observation that they form a natural monopoly: economically, it is most efficient that there is only one network. Competition between different electricity and gas networks within a single region can therefore not be expected. An exception is possibly the transmission of gas, in which in principle it is possible that multiple networks compete with each other (as is the case with the Zebra pipeline, owned by Essent and Delta, which competes with the network of GTS). At the level of the distribution networks, however, competition between gas and electricity networks is certainly not possible. For electricity, the Dutch Electricity Act of 1998 establishes that only one network manager may be active per region. 56 For gas, this has not been established explicitly, but in practice there is only one gas distribution company per region. For each of these

distribution networks a network manager needs to have been appointed.\textsuperscript{57}

Because the users of the electricity and gas networks do not have the ability to provide incentives to the network managers for maintaining the quality of service through their choice of provider, the position of consumers will need to be protected through regulation of the network managers. Electricity and gas network managers therefore are subject to regulation, based upon the 1998 Dutch Electricity Act, respectively the Dutch Gas Act, in most cases by the Energy Regulator.

For the purpose of unbundling monopoly activities, such as network management, from competitive activities, the Dutch Electricity Act of 1998 and the Gas Act require network managers to be legally separate from producers, traders or suppliers.\textsuperscript{58}

Generally speaking, there are three relevant aspects with respect to the regulation of network managers:

- the technical conditions for users of the networks,
- the price of the services, and
- the quality of the delivered services.

Both price and quality can be considered from the point of view of the individual user and for the entire network of a certain network manager. (In the latter case, the issues are total revenues and average quality.)

The services which network managers deliver need to include at least a connection to the electricity or gas network, transport of electricity or gas to that connection point and system services. The delivery of the gas molecules or electric energy is assumed to be taken care of by the supplier (retail company, for small consumers). The measurement (‘metering’) of electricity production or consumption, or of gas consumption, is considered a ‘free’ activity, i.e. not regulated, so that any company can do this.

Technical conditions

Electricity

The management of electricity networks is a highly technical activity. Generation facilities, network components and also the equipment of consumers need to meet all sorts of technical requirements. In addition, there is a need for clear rules about the way the different actors (producers, network managers and consumers) interact. There are two basic issues at stake with respect to these technical requirements. The first is which of the technical conditions for the use

\textsuperscript{57} Art. 10 Electricity Act (1998) and art. 2 Gas Act.

\textsuperscript{58} Art. Electricity Act (1998) and art. 3 Gas Act.
of networks are determined centrally (i.e., for all networks). The second is who determines those conditions. It appears inevitable that at least part of the technical conditions are established centrally and hold for everyone, for all network users (including network managers) influence the way the entire system functions, and the general performance of the system benefits from a certain uniformity of the rules.

The technical conditions for the use of the electricity system are established at three levels. First, the Dutch Electricity Act of 1998 contains certain basic rules. Second, the (technical) conditions for electricity are contained in the technical codes, such as the Network Code (NetCode), de System Code (SysteemCode) and the Metering Code (MeetCode). These codes are created upon the initiative of the (joint) network managers and with the involvement of representative organizations of market parties, after which they are established officially by the director of the Office of Energy Regulation of the Competition Authority.\(^{59}\) The I&I Act provided the Minister of Economic Affairs the authority to establish rules which the network managers and the energy regulator need to follow when developing the technical conditions for network access.\(^{60}\) Everything that is not regulated by the law or the technical codes may be arranged through agreements between the network manager and the individual user of the system.

Gas

For gas, there did not use to be a system with technical codes like for electricity. This was related to the original choice not to have regulated but negotiated third party access to the network in the Netherlands. Only the (technical) conditions which gas network managers used for small consumers were subject to supervision by the Minister of Economic Affairs.\(^{61}\) The director of the regulator established the technical requirements for installations.\(^{62}\) For the other conditions the network managers were required to publish an indication. These indicative conditions were required to meet the guidelines of the regulator.\(^{63}\)

The I&I Act brought significant changes to the gas sector. Now, technical codes will also need to be developed for gas. The joint network managers will take the initiative and, after representative organizations of market parties have been involved, the codes will be established formally by the director of the energy regulator.\(^{64}\) As in the case of electricity, the Minister of Economic Affairs will have the

---

\(^{59}\) Art. 31 and 36 of the Dutch Electricity Act (Elektriciteitswet) of 1998.

\(^{60}\) Art. 26b of the Dutch Electricity Act (Elektriciteitswet) of 1998.

\(^{61}\) Art. 82 (old) Gas Act.

\(^{62}\) Art. 11 (old) Gas Act.

\(^{63}\) Art. 12 (old) and 13 (old) of the Gas Act.

\(^{64}\) Art. 12b through 12f of the Gas Act.
authority to create rules which the network managers and the energy regulator need to obey. This article (12 of the Gas Act) has not yet entered into force, however.65

(Economic) regulation of the network managers

There are three goals for the economic regulation of the electricity and gas network managers. Network managers need to be able to collect sufficient revenues to be able to perform their functions properly. Individual consumers need to be protected against excessive tariffs. Third, the tariff system should stimulate the efficient use of the networks and the network services.

Previously: cost plus

The traditional way of economic regulation is called cost plus regulation: the tariffs of the energy companies were based upon the incurred costs plus a reasonable rate of return. A disadvantage of this system of regulation is that energy companies have little incentive to pursue efficiency (lower costs). Another disadvantage is that this system of regulation requires the regulator to have an understanding of the costs of the network manager and the purposes for which these costs were incurred. In the Netherlands, the establishment of the tariffs under the Electricity Act of 1989 actually resembled more a gentlemen’s agreement: as long as the tariffs appeared reasonable, the Minister approved them.

Electricity: now price cap regulation

For these reasons, the Dutch legislature chose for a different system to regulate electricity networks after liberalization, which should provide better incentives for efficiency and which should require less detailed involvement of the regulator in the financial decisions of network managers. This is the system of price cap regulation. In this system the costs are no longer the basis for the tariffs. Instead, the tariffs are limited to a maximum. In the Netherlands, the tariffs for the next year are derived from the tariffs in the current year by allowing them to increase with inflation (cpi, the consumer price index), while on the other hand they are subject to an efficiency reduction (x).66 This ‘cpi – x’ system should provide an incentive to network managers to reduce their costs (and thereby increase efficiency). In addition, the regulator would require less specific information from the network managers: essentially, he would only need to determine the value of x.

Output control

The method that is applied in the Netherlands can be characterized as output control: the total revenues of the network manager are determined by his total output (the total of services that he delivers in a given year), multiplied by the tariffs for those services. It is left to the network managers to decide how they can provide that output as

66 Art. 41 through 41b of the Dutch Electricity Act (Elektriciteitswet) of 1998.
efficiently as possible within the revenue cap. The more efficient they are, the more revenues remain for profit.

**The efficiency factor**

The idea behind the efficiency factor $x$ is that the network managers have room for improving the efficiency of their activities. A number of options exist with respect to the implementation of an efficiency reduction, such as the choice between a uniform efficiency factor versus a different factor for each network manager. A uniform efficiency factor is fair if the tariffs of the different network managers are at a similar level of efficiency; if this is not (yet) the case, efficiency factors that are differentiated per network manager may direct the different network managers to the same level of efficiency.

The first regulation period (1999-2004) of the ‘cpi – x’ system of regulation of the electricity networks was somewhat difficult. The regulator DTe had conceived of an innovative system for determining the efficiency factor $x$ for the different network managers. According to the judge (the College van Beroep voor het bedrijfsleven, an appeal court for businesses) this system was not supported by the law, as a result of which DTe had to change its practices. The legislator would have had a uniform efficiency factor in mind, whereas the DTe applied a different factor to each network manager. In the meantime the Electricity Act has been adjusted so DTe now is allowed to apply different efficiency factors to different network managers.67

Later, the efficiency factor $x$ was deduced from the efficiency improvement of the other network managers: *yardstick competition*. In a certain sense this caused network managers to compete indirectly with each other with respect to efficiency. Yardstick competition imitates a competitive market by allowing the most efficient network manager to determine the ‘price’ (i.e. the network tariffs) which will also apply to the other network managers (see also Ajodhia *et al.* 2003).

A particular feature of price cap regulation in the Netherlands is that it is applied to both operational expenditures and to investment (Ajodhia *et al.* 2003). By including investment under the price cap, investment decisions come under pressure, because a reduction of replacement and expansion investment leads to higher returns. Therefore the concerns are justified that the current system of economic regulation may affect the quality of the networks in the long term. For this reason, DTe introduced a system of quality regulation, in which the quality of the network services becomes an integral part of the regulation system.

67 See regarding the issue of the efficiency reduction $x$: Knops 2002 and Knops 2003.
Yardstick competition

Yardstick competition appears to be an interesting possibility for creating some form of ‘competition’ between monopolists. However, three problems may present themselves when implementing yardstick competition (Weyman-Jones 1995):

- **commitment** of the regulator: a regulator must accept the possibility that high profits are made or that companies go bankrupt; it may be doubtful that the regulator can resist the political pressure to intervene;
- **collusion**: if the income of one company depends upon the performance of the other companies, there is a danger that the companies make deals among each other with respect to their performance level;
- **comparison**: the assumption that underlies yardstick competition is that the different companies are comparable; however, it is very difficult to devise a good and workable yardstick with which all the relevant companies can be compared.

**Tariff structure:**

Until now, the focus was upon the level of the tariffs and the resulting income of the network managers. Besides the level of the tariffs, their structure is relevant. In which way are the costs of network management presented to the users of the system? For instance, does a user pay for transport per unit of energy (kWh, m³), or (partly) for his peak use (kW, m³/h)? The Dutch Electricity Act of 1998 contains the basic structure of the tariffs for connection, transport and system services. The tariff structures are further developed in the Tariff Code (*TarievenCode*), which is established by the director of the energy regulator. The I&I Act added that the network managers and the director of the energy regulator need to consider the rules which the Minister of Economic Affairs has established.⁶⁸

**Gas: previously only transport for the benefit of small consumers was regulated**

Prior to the I&I Act, for gas only the tariffs for transport of gas that was destined for small consumers were regulated in a manner similar to that for electricity. The Minister of Economic Affairs established these tariffs annually according to the ‘cpi – x’ formula. For the remainder of the transport tariffs, the network managers were required to published indicative tariffs, which should have taken into account the ‘guidelines’ which the energy regulator had established.

The I&I Act introduced a new system. Also for gas the tariff structures will be established in a kind of tariff code. Article 12a of the Gas Act prescribes that the joint network managers make a proposal to the director of the energy regulator with respect to the structure of the tariffs for the transport of gas which they will charge to network users (including imports, transit and exports of gas), the tariffs for ancillary services that are related to transport and for the tasks of the manager of the national gas transport network.

Subsequently the director of the energy regulator establishes the tariff structures.\footnote{For LNG installations and for access to storage facilities there is a somewhat different regulation, see Art. 13 of the Gas Act, respectively Art. 18 through 18g of the Gas Act.}

Not only the tariff structures are regulated now, but also the tariffs themselves (for transport and ancillary services) will be established by the government, as was already the case for electricity. The I&I Act provided the director of the DTe with the authority to establish (maximum) tariffs for the transport of gas that is intended for consumers (as opposed to export), as well as tariffs for ancillary services.\footnote{Art. 81c of the Gas Act.} ‘Ancillary services’ include the connection service.\footnote{Kamerstukken I, 2003-2004, 29 372, C, p. 16.} The tariffs are determined with the ‘cpi – x + q’ formula. The ‘q’ in this formula is a quality term; its meaning will be explained below.

**Quality regulation**

As was mentioned above, a system of price cap regulation, in which the capital expenditures are included under the price cap, contains a risk that the quality of service is affected. To prevent the quality of the networks from deteriorating to a socially undesired level, DTe proposed to implement a system of quality regulation (DTe, 2002). In this system, the focus is on the reliability of supply as the main quality parameter.

On the proposal of DTe, a quality term ‘q’ was added to the tariffs that are established with the ‘cpi – x’ formula, the network tariffs for electricity and the tariffs for the transport of gas. The formula became ‘cpi – x + q’.

With ‘quality regulation’, an increase in quality (better reliability of service) is rewarded with higher revenues for the network manager and a reduction of quality is ‘punished’ with a reduction of revenues. The underlying idea is that quality has a certain ‘price’: society is assumed to be willing to pay a certain price for quality improvements; conversely, a lower level of quality would cost society a certain amount.

If the regulator is able to determine the correct ‘price’ of (a change in) quality, the network manager will be able to decide which level of quality is ‘efficient’ for him. This will be the level at which the extra revenues of a quality improvement equal the extra costs which he needs to make to obtain this level.
But what happens if the level of the quality incentive is not determined correctly? If it is too low, the quality level will become lower than optimal; if it is too high, we as consumers will pay more than the resulting level of quality is worth to us. This possibility certainly exists, witness the results from studies of the social costs of electricity service interruptions which may vary by a factor 10. (These social costs are a determining factor for the quality incentive, because the quality incentive needs to reflect the social costs of service interruptions in order to guide the quality of electricity supply to the social optimum.)

The proposed system of quality regulation focuses on the average quality of service of a network manager. Individual consumers may experience a much worse or better level of quality. So the proposed system of quality regulation provides little protection for the individual consumers. This issue could be resolved by letting network managers pay penalties to affected users in the case of severe service disruptions. (Afterwards, this would need to be taken into account when determining the quality incentive.) Another option is establish minimum standards, in addition to the quality regulation of the average performance, to prevent that locally the quality level will decrease too much.

The question is to what extent network managers are liable for damages resulting from supply disruptions. If network managers are liable for (part of) the damages, they experience (part of) the social costs of supply disruptions also without quality regulation. Currently, the energy companies avoid being liable for service disruptions as much as possible. The Minister of Economic Affairs, however, would like to expand their liability.

With the I&I Act and the Minister’s wish to increase the energy company’s liability, there is a threat of accumulation of rules and regulations: the general quality incentive ‘q’, individual penalties for significant outages and possible damage compensation to consumers. The network manager experiences the sum of these three incentives. In order to arrive at the desired optimum, the sum of these incentives needs to be optimal. It is a difficult task to adjust these three incentives to each other.

### 5.3 Distribution network unbundling

The organization of the energy supply industry is based upon the conviction that not the entire supply chain needs to be a monopoly, but that certain activities can be performed in competition. Other activities, such as network management, need to remain a monopoly. While breaking up the former vertically integrated monopoly causes a loss of economy of coordination, the idea is that this should be more than compensated by the gains which would result from improved efficiency in the newly competitive activities. An important assumption is that all elements of this coordination, both technical and economic, can be recreated in contractual relations between the
independent actors. In the years since liberalization, it was discovered that the transaction costs were higher than anticipated. Serious coordination issues arose between generation, consumers and network operators. Unbundling also increased the customer administration transactions, as network services and energy sales were unbundled and needed to be billed separately. Eventually, the latter two were re-integrated in the Netherlands, with the retail company also billing consumers for their network charges.

The monopoly activities need to be regulated. In addition, the relation between the competitive activities and the monopoly activities needs to be guarded. On the one hand the performance of a monopoly task may not lead to improper competitive advantages for market activities. Network management, on the other hand, should not be influenced by commercial interests. This is the essence of unbundling.

How can unbundling be secured? Traditionally, different levels of unbundling are distinguished:

- separate accounting by the network manager;
- administrative separation: the network manager forms a separate business unit;
- juridical separation: the network manager forms a separate legal person (but may be part of a larger firm);
- ownership unbundling: the network manager is juridically separate and is the owner of the network; he is not allowed to be active in commercial functions.

Until 2003, the Netherlands the network managers were only required to keep separate books. This was to facilitate tariff control by DTe and prevent cross subsidies. Network managers were also not allowed to pass confidential information along to related businesses. In practice, these ‘Chinese walls’ turn out not always to be so firm, as became evident from the aftermath of the bankruptcy of EnergyXS.

The independence of the network manager did not always appear guaranteed. Especially with respect to strategic decisions, such as network investment, network managers are dependent upon the mother company – on whose balance the network assets often remain. The European directives of 2003 made the requirements for the independence of the network managers substantially more rigorous. For instance, a network manager was now required to have sufficient actual control so he can make independent decisions regarding the assets that are necessary for the management of the network.

As was discussed above (Section 3.1 ), the Electricity Directive and the Gas Directive of 2003 require the Netherlands to implement separate accounting and administrative and juridical unbundling of
the network managers. A more rigorous degree of unbundling was optional and eventually implemented.

The European Directives of 2003 did not draw the ultimate consequence of the arguments for unbundling. The Directive does not contain a requirement to unbundle the assets of the network from a vertically integrated company, for the Directive does not concern itself with issues of ownership. That is, traditionally, an issue for the member states. Nevertheless, separating the networks from commercial activities is a pure and consistent solution, from a system point of view.

Where the energy companies point to ‘synergy advantages’, they actually touch upon remaining unfair competitive advantages over energy producers or suppliers who are not related to a network company. Where they complain about the high interest rates that they will need to pay on their ‘commercial’ loans, they forget to mention that the network company will be able to obtain cheaper credit when it is unbundled. The example of KPN’s commercial adventures, which reduced the firm to junk bond status, illustrates that the networks are at risk when integrated with a commercial firm. Another argument proffered by the energy companies was that the unbundled firms are worth less, together, than the integrated company. If this were true, it could mean two things. First, it might have been another indication of the existence of unfair competitive advantages. Second, it might have been an indication that the separation of the energy supply industry into competitive and monopoly activities did not yield a net (economic) advantage... In which case, the introduction of retail competition would have been a mistake.

In 2006 it was decided that the Dutch electricity and gas networks would be required to be fully (ownership) unbundled. The companies were given until 2011 to implement the separation, but had to submit a plan subject to approval by the Minister of Economic Affairs in 2009. All plans were submitted and approved – often with some modifications – in 2009. Of the large companies, Essent and Nuon swiftly unbundled their networks, so the commercial parts of these companies could be privatized. They were purchased by RWE and Vattenfall, respectively. The shareholders of Eneco and Delta were not interested in privatization and therefore opposed to unbundling. They filed a law suit against the forced unbundling of their companies which they won on the grounds that forced was at odds with European rules concerning the free flow of capital. The verdict is currently being appealed by the Dutch State.

However, the juridical separation of the distribution network managers for electricity and gas may be delayed until July 1st, 2007.
Level playing field – for whom?

Juridical unbundling of the network managers and providing them with the complete ownership of the networks isolates monopoly activities from competitive activities. This way, the Netherlands can create a level playing field for the incumbent energy companies and newcomers without networks.

However, the Dutch energy companies compete also with foreign energy companies. If these are not subject to ownership unbundling, their networks will provide them a significant competitive advantage over the incumbent Dutch energy companies.
6  Congestion Management

6.1  Introduction

This chapter provides a theoretical description of the main congestion management methods that are being used in Europe. Europe requires congestion management methods to be market based. As locational marginal pricing – also known as nodal pricing – is currently not politically and institutionally feasible in Europe, this limits the choice of congestion management methods to variations of auctioning (including market splitting) and variations of redispatching. The goal of this chapter is to provide a theoretical framework for understanding the various versions that have been developed in practice.

Cross-border congestion has become an important issue in the process of liberalizing the internal European electricity market. In the past, electricity supply was organized nationally or regionally: the electricity grid was of national (or regional) size and was operated by a monopolistic company which usually was vertically integrated, i.e. which also had the generating stock. Production of electricity was largely domestic. Physical links between different national grids, the so-called interconnectors were built initially only for the purpose of enhancing the overall stability of the system through exchange in cases of emergency. Later, some structural exchanges started to develop, based upon long-term contracts. As these exchanges took place between vertically integrated utilities, they took care that these transports did not exceed the available interconnector capacity.

At least the largest customers must have (third party) access to the grid, so that they are able to use their freedom to buy electricity from their supplier of choice. The exact organization of the national electricity markets is left to the member states (subsidiarity), but in general it is possible for a free customer to buy his electricity from a foreign supplier73. Especially in cases where electricity is significantly cheaper in one member state than in its neighbor, a large demand for cross-border transports can occur. In some of these cases the demand for imports exceeds the available transport capacity on the (cross-border) interconnector(s) and congestion occurs. Congestion may hamper the full integration of the different national electricity markets into a single market.

73 Only if the customer (eligible in his home country) would not have been eligible in the country of his supplier, a cross-border transaction can be prohibited (Article 19(5) Directive).
Examples

In this chapter, the description of each congestion management method will be accompanied by an example. These examples all have the same basic set-up, which is introduced now. Assume there are two countries $A$ and $B$ with a transmission line connecting the electricity networks of the two countries (the ‘interconnector’). Country $A$ has a peak demand of 1500 MW, country $B$ of 3000 MW. Each country has five generators with a capacity of 600 MW. Figure 6 diagrammatically shows the situation. The tables describe the countries’ generation parks. Each G indicates a generator and next to it is its marginal cost (measured in €/MWh), which, for simplicity’s sake, is assumed to be constant. It is further assumed that the generators will bid their marginal costs. In line with the assumptions of the perfect competition model, the fixed costs of generators are ignored.

The following symbols and indices will be used in this chapter:

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>D</td>
<td>B</td>
</tr>
<tr>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td>G</td>
<td>K</td>
</tr>
<tr>
<td>P</td>
<td>T</td>
</tr>
<tr>
<td>Q</td>
<td>R</td>
</tr>
<tr>
<td>R</td>
<td>S</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DA = 1500</th>
<th>DB = 3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marg. Cost (€/MWh)</td>
<td>Marg. Cost (€/MWh)</td>
</tr>
<tr>
<td>$G_{A1}$</td>
<td>$G_{B1}$</td>
</tr>
<tr>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td>$G_{A2}$</td>
<td>$G_{B2}$</td>
</tr>
<tr>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>$G_{A3}$</td>
<td>$G_{B3}$</td>
</tr>
<tr>
<td>24</td>
<td>27</td>
</tr>
<tr>
<td>$G_{A4}$</td>
<td>$G_{B4}$</td>
</tr>
<tr>
<td>26</td>
<td>30</td>
</tr>
<tr>
<td>$G_{A5}$</td>
<td>$G_{B5}$</td>
</tr>
<tr>
<td>28</td>
<td>32</td>
</tr>
</tbody>
</table>

Figure 6: Example set-up with generator marginal costs

For determining the economically efficient dispatch of generation, a key variable is the marginal cost of the most expensive generator
which is producing electricity. The market price is assumed to be equal to the marginal cost of generation. Consumers pay a different price in each country, unless there is enough interconnector capacity to combine the countries into a single market. Consumers pay generators their cost per unit times their output.

Generators with a lower cost of generation than the market price of electricity make a profit, also labeled producer or generator surplus. Total generator surplus can be calculated as the sum of all individual generators’ profits, but also from the difference between total generator revenue and the sum of all generators’ marginal costs.

The analysis is based upon the ideal case of perfect competition, and therefore does not enter into practical problems such as strategic behavior by market parties and imperfect information. Assuming perfect competition means, among others, that:
1. All players have perfect information,
2. There is a sufficient number of market players, so that no individual player can influence the market price,
3. Market parties can freely enter and exit the market,
4. Electricity is assumed to be a homogenous product.
5. The cost of transmission and distribution are assumed to be constant and are not taken into account.
6. Consumer demand is assumed inelastic.

These assumptions generally do not hold in actual electricity markets, except that electricity may be considered a homogenous product, apart from artificial distinctions such as for electricity generated from renewable energy sources. Some electricity markets have enough players to meet the second criterion, but this is not necessarily the case in each market. The other assumptions do not hold in electricity markets: there is always an information asymmetry between the different market players (and government). Free entry and exit does not exist either in the power market. On the production side, the high capital requirements obstruct easy entry and exit of the market; on the demand side, nearly all consumers are so dependent upon electricity that they cannot leave the market. Therefore the framework here provides a theoretical benchmark, a means of assessing the best possible outcomes.

### 6.2 Selected congestion management methods

In this section, the framework that was presented in the previous paragraph is applied to three methods for congestion management. Note that there are more congestion methods available but for the interest of the course: explicit auctioning, market splitting and redispatching are used to explain the purpose of congestion management.
6.2.1 **Explicit auctions**

In this section we will make an economic analysis of explicit auctions for interconnector capacity. There are several forms of explicit auctions of which pay-as-bid auctions and marginal bid auctions are the most relevant for electricity markets. These two methods differ in the price which market parties pay for the capacity. Pay-as-bid auctions will be described first, followed by marginal bid auctions. In a pay-as-bid auction, each participant who wins capacity in an auction has to pay the amount he has bid. In a marginal bid auction, the lowest bid that wins capacity (the marginal bid) is the price that all other participants who win capacity also pay. Pay-as-bid auctions are primarily interesting from a theoretical point of view, because they generate the largest possible congestion rents. Marginal bid auctions are more often applied in practice because they appear to be more fair and provide better bidding incentives.

First we will turn to pay-as-bid auctions. Figure 7 explains the two types of auctions. The figure shows supply and demand equilibria, with generator output and demand on the horizontal axis and price and cost on the vertical axis. Demand is assumed to be perfectly inelastic, which means that consumer demand in \( A \) is equal to \( Q_A \) and consumer demand in \( B \) is equal to \( Q_B \) under all circumstances.

The presence of imports with a volume equal to the interconnector capacity \( K \) cause generator output in \( B \) to be reduced from \( Q_B \) to \( Q_B' \).\(^74\) As \( Q_B'' \) becomes the marginal generator, the price in \( B \) drops from \( P_B \) to \( P_B'' \). To supply the exports, generation in \( A \) is increased with a volume equal to \( K \) from \( Q_A \) to \( Q_A'' \) so that the price increases from \( P_A \) to \( P_A'' \). As generators in \( A \) who export to \( B \) want to receive a net income that is at least equal to their marginal cost of production, their willingness to pay at an auction equals the difference between the price they can receive in country \( B \) (\( P_B'' \)), and their marginal cost, which is represented by the curve \( S_A \). The shaded area therefore represents the combined willingness to pay of all generators that participate in the auction. This amount is in theory equal to the auction revenues in a pay-as-bid auction and represents the theoretical maximum revenue that can be expected from any congestion pricing system. However, the goal should not be to maximize revenues but to allocate the scarce interconnector capacity efficiently.

If a pay-as-bid auction works well, the prices paid by bidders ensure that only electricity from the most efficient generators is transported across the interconnector. Therefore it has the potential of achieving economic efficiency. In practice, bidders who pay more than the marginal bid will see that they could have made a better profit if they

\(^{74}\) Prices and quantities with a double prime refer to the situation with congestion; the single prime denotes the situation without congestion and no prime refers to the starting position, without an interconnector.
had bid less, as long as their bid was above the marginal bid. Bidders will therefore try to estimate the value of the marginal bid, and bid only slightly more themselves. Given the fact that electricity auctions are repeated infinitely, they may become quite good at this game. As a result, the auction revenues will not equal the full willingness to pay of the bidders. In fact, the auction results will resemble less the above case and more the marginal bid auctions, which will be discussed presently.

Figure 7: Explicit auctioning

In a marginal bid auction, the price of the interconnector capacity is set equal to the marginal bid, which is the lowest bid that is awarded transmission capacity. All bidders who have bid more than the marginal bid receive capacity at the price of the marginal bid. This system reduces the theoretical expected revenues, but should result in the same generators obtaining capacity as when a pay-as-bid auction is used. Bidders make a profit equal to the difference between their willingness to pay and the marginal bid. As a result, a bidder’s optimal bidding strategy is to bid according to his willingness to pay: if the auction price turns out to be higher, the price is too high for this bidder; if the price is lower, then he is not punished for the fact that he has “overbid”. This makes bidding easier and the auction process more transparent than in the case of a pay-as-bid auction. This improves the incentive to participate in the auction compared to a pay-as-bid auction, so that a liquid market is more likely to develop.
This in turn may also improve the function of the auction to select the most efficient generators. Therefore a marginal bid auction is also more likely to indicate the market value of the congested interconnector at the time of the auction.

The marginal cost of the last generator to win capacity is equal to $P_A''$, so he is willing to pay $P_B'' - P_A''$: the price difference between the two countries in the presence of the interconnector. This is the marginal bid which sets the price for all auction participants in a marginal bid auction. Therefore the auction revenues equal the interconnector capacity times the price difference between countries $A$ and $B$ per unit of energy. In Figure 7 the area indicated by the double-lined box represents the revenues from a marginal bid auction. This revenue is always smaller than the revenue from a pay-as-bid auction, unless all bidders have the same marginal cost (in which case $S_A$ is flat between $Q_A$ and $Q_A''$). Therefore a marginal bid auction should in theory always yield less revenue than a pay-as-bid auction. In practice, the difference may be small or absent due to the incentives to underbid in a pay-as-bid auction and the incentives to bid according to willingness to pay in a marginal bid auction, and due to the latter’s greater attractiveness to participate. The theoretical congestion rents from a marginal bid auction are the same as from an implicit auction and market splitting, as will be seen in the next sections.

While market parties may pay less in a marginal bid auction, the prices are just high enough to exclude the less efficient generators from access to the interconnector. Therefore this system also has the theoretical potential of being economically efficient in the short term, that is, of achieving efficient generator dispatch. Because this type of auctioning encourages bidders to bid equal to their full willingness to pay, it may actually prove better at selecting the most efficient generators, and may therefore be more efficient in practice.

Apart from the choice of bidding system, other important variables in the design of auctions are the time periods which are auctioned (days, weeks, months, years) and the firmness of capacity rights. While the above description concerns the theoretical auction results, these other variables will have a significant impact upon the actual performance in practice.

Both types of explicit auctioning separate energy flows from transmission capacity, which corresponds to the principle of unbundling. An important advantage is that firm transmission access is provided ahead of time. A disadvantage is that this system requires separate transactions for trading electricity and obtaining transmission capacity. The additional transaction increases the complexity of cross-border power trade, and may therefore pose a barrier to trade. When a transaction leads across multiple congested borders, for all of which capacity needs to be obtained in auctions, the complexity increases quickly.
Auctions generate a revenue stream which is indicative of the market value of the congested line. It is important to note that this is not an indication of the value of capacity expansion, but of the value of the existing capacity. The marginal bid equals, in theory, the marginal value of interconnector capacity. The value of capacity expansion depends upon the additional benefits from trade which it would enable and these depend, in turn, upon the cost curves of the generators that were not able to produce due to the congestion but that would have been in merit otherwise.

Example – Pay As Bid

Assuming that all parties have full information, $G_{A3}$ can only sell half its generation capacity in its domestic market so it will want to bid for interconnector capacity. As the sales across the interconnector will drive $G_{B5}$ out of the market, the price in $B$ will become 30 €/MWh. $G_{A3}$ has a generation cost of 24 €/MWh, so its maximum bid is 6 €/MWh of line capacity. He will bid up to 6 €/MW of line capacity, for each hour that it is available to him. Since he wants to export 300 MW, that means a total bid of 1800 €/h. Similarly, $G_{A4}$ will bid 4 €/MW per hour of use. He bids for 500 MW of line capacity, which means a bid of 2000 €/h.

When each bidder is made to pay his full price, revenues from the auction will be 3800 €/h of use.

Generation cost is 109 600 €/h, the same as in the case of optimal allocation. Generators’ surplus is 15 600 €/h, less than if there were no congestion. This is due to the auction fees they pay. In this example, the generators pay the auction fees. In practice, consumers, traders or combinations of consumers and traders with generators may also bid for capacity.

Consumers pay a price of 26 €/MWh in $A$ and 30 €/MWh in $B$, which means in total they pay 129 000 €/h. This figure lies between the amount they would pay in the case of no congestion and the case of no interconnection, as can be expected.

Example – Marginal Bid Auction

If the price is set according to the marginal accepted bid, the price will be 4 €/MW per hour for both bidding generators ($G_{A3}$ and $G_{A4}$). This means revenues are 800 x 4 = 3 200 €/h. The market prices are the same as in the previous example, so consumers pay the same. In this theoretical example, the generators which bid for the auction are the only ones to benefit from the lower auction costs. In particular, $G_{A3}$ saves 600 €/h on his bid for 300 MW. As a result, generator surplus increases to 16 200 €/h.

For simplicity, we assume the congestion exists continuously. In reality, parties will only bid for access during periods of congestion; outside these periods, bid prices would fall to zero.
6.2.2 Market splitting

When market splitting is used to manage congestion, the market is divided by the congested interconnector. There either needs to be an organized market with a separate price on each side of the interconnector, or there need to be two closely co-operating power exchanges. Market parties bid into the organized market on their side of the congestion. In a first step, the two markets are cleared independently. Then the market operator buys electricity from the organized electricity market with the lower price and sells it in the market with the higher price, precisely so much that the interconnector capacity is fully used. The result is that the prices in \( A \) and \( B \) move closer together.

Because the market operator provides an additional demand in \( A \), the market price increases from \( P_A \) to \( P_A' \). The reverse happens in the more expensive market: by increasing supply, the market operator lowers the price in \( B \) from \( P_B \) to \( P_B' \). He provides electricity more cheaply than some of the more expensive domestic generators can offer. Thus the market operator buys at \( P_A'' \) and sells at \( P_B'' \). Because the volume of electricity bought in \( A \) equals the capacity of the interconnector \( K \), the TSO’s revenues are equal to the price difference times the interconnector capacity. Again, the revenues are equal to those from marginal-bid explicit auctions. The revenues can be earmarked for capacity expansion, but they can also be given to the TSO in return for a corresponding reduction of transmission tariffs, as is done in Norway.

Graphically, market splitting can be represented as follows. The market operator buys an amount of \( K \) electricity in country \( A \). Thus he increases demand by \( K \), so it shifts from \( D_A \) to \( D_A'' \). See Figure 8. By adding an amount of \( K \) to the demand curve, the new equilibrium price in \( A \) becomes \( P_A'' \), at the intersection of \( D_A'' \) and \( S_A \). As the market operator sells the electricity he bought in \( A \) to the market in \( B \), he moves the supply curve in \( B \) to the right. In Figure 9, the new supply curve is labeled \( S_B'' \). This causes the price in \( B \) to drop from \( P_B \) to \( P_B'' \), the intersection of the demand \( D_B \) and the new supply curve \( S_B'' \).

---

76 The market operator buys electricity at the new market price of the cheaper market, even though he knows all the bids. The rule at organized electricity markets is that all parties receive the same price for the electricity they offer. It would not only seem to be an unjust use of its information for the market operator to only pay the bid price, when the bid price is less than the market price, but possibly also distort the bidding process in a way similar to a pay-as-bid auction.
Figure 8: Market splitting; situation in A, the lower-priced country

Figure 9: Market splitting; the situation in B, the higher-priced country
Market splitting leads to an efficient dispatch of generation, simply because the market operators buy electricity from the cheapest generators upstream of the congestion. As a result, generation cost is the same as in the previous cases. Market prices are also the same, so welfare effects are the same too.

This congestion management method has as an important advantage that market parties only need to bid into or buy from their own markets. In fact, market parties do not even know whether their bids are used for their local market or for export across the interconnector. As a result, transaction costs and risks are considerably lower than in the case of an explicit auction. As a consequence, there is less of a barrier for small market parties to participate, so market liquidity can be expected to be higher. A disadvantage is that physical bilateral contracts between the two countries are problematic. However, they can be replaced by financial contracts that provide the same benefits to the contract parties. In the Nordic market, the only place where market splitting currently is practiced, there is a strong trend towards financial instead of physical contracts.

Market splitting is currently being implemented on a number of borders in the EU. The Netherlands, Belgium and France have a joint system for managing congestion on the two borders through an improved version of market coupling. Germany and Denmark are working on market coupling, and the goal is that the capacity of the NorNed cable will be allocated this way and that it will be implemented on the Dutch-German border. The latter creates the problem of a meshed network, as flows between Germany and the Netherlands, Belgium and France can follow parallel routes. This leads to complex algorithms for the determination of prices.

Example

Initially, the market operator separates the markets in A and B. The markets are cleared as if there is no interconnector, which results in a price of 24 €/MWh in A and 32 €/MWh in B. As the market price in A is lower than in B, the market operator then buys 800 MW in A and sells it in B. Purchasing from the cheapest generators in A, this results in $G_{A3}$ increasing its output from half to full capacity and $G_{A4}$ providing an additional 500 MW. Thus, $G_{A4}$ becomes the marginal generator in A, and the price in A increases to 26 €/MWh. The market operator sells the 800 MW in B, where it forces $G_{B5}$ out of the market. $G_{B4}$ becomes the marginal generator, so the price in B drops to 30 €/MWh.

The market operator pays generators in A the new market price, which is 26 €/MWh as $G_{A4}$ now is the marginal generator. The market operator’s purchase price for 800 MW is 20 800 €/h. He sells the 800 MW at a market price of 30 €/MWh, so he receives 24 000 €/h, making a profit of 3200 €/h. Revenue streams and market prices are exactly the same as in the case of marginal bid explicit auctions and implicit auctioning.
6.2.3 Redispatching

Redispatching is the default congestion management system, as it involves direct intervention of the TSO in the dispatch of generation. There are different ways to implement redispatching. We describe a system in which the market experiences as little as is possible of the existence of congestion. It works as follows: The market trades as if interconnector capacity is unlimited. As a result, a single price develops in A and B: $P_T$. This results in a demand $D_e$ for electricity flow across the congested line that exceeds its capacity $K$.

Contrary to the congestion management methods that were described above, market parties are not required to change their transactions as a result of the existence of congestion. The generators in A receive market contracts to produce at level $Q_A'$ and those in B at level $Q_B'$ (see Figure 10). Clearly, this would lead to a net flow from A to B in excess of the interconnector capacity. To avoid physical overloading of the interconnector, the TSO intervenes directly in the dispatch of power plants in both countries. He reduces output in the exporting country and increases generation in the importing country up to the point that the net flow across the interconnector matches the available capacity. This process of adjusting generation output by the TSO is called redispatching. In the case of congestion on interconnectors between different systems, the two (or more) TSOs need to cooperate closely.

![Figure 10: Redispatching](image-url)
Redispatching costs the TSO money. This can be seen as follows. In B, the importing country, the TSO needs to dispatch additional generation. He can choose from the generators that did not receive generation contracts in the unconstrained situation. Under the conditions of perfect competition, the only reason they would not run would be that their marginal costs of production are higher than the market price $P_T$. Of course, the TSO will still try to find the cheapest remaining generators. In Figure 10, they are the ones between $Q_B'$ and $Q_B''$.

In A, the exporting country, production must be reduced from $Q_A'$ to $Q_A''$ in order to bring the exported volume of electricity within the capacity limit of the interconnector. The decrease in A is equal to the increase in B (because demand is assumed inelastic). The generators in A whose output is reduced still receive the unconstrained market price for their generation contracts (which are still valid) and by not generating they save their variable costs of production. The TSO demands a reimbursement from these generators in A equal to their avoided costs, so that the redispatching process leaves them financially indifferent. The reimbursement which the TSO receives from the cancelled generators in A will always be less than the cost of the extra generation in B. This is logical: had the generators in B been cheaper, they would already have run.

Had there been sufficient interconnector capacity, the TSO would not need to redispatch. Therefore, the cost of redispatching is precisely the cost savings that can be obtained by enlarging the interconnector to the point that there is no congestion. In other words, the cost of redispatching indicates the value of interconnector capacity expansion, when the value of interconnector expansion is taken as the potential savings in operation cost (which is equal to the increase in aggregated consumer and generator benefits).

Whether it is economically efficient to actually expand the interconnector depends on the cost of expansion: this should be less than the cost savings during the life span of the interconnector. It will therefore not be efficient to expand the interconnector to the point that all congestion is alleviated; at that point surely the marginal cost of expansion is larger than the marginal cost savings from a more efficient redispatch. An advantage of redispatching is that the system can be arranged in such a way that the TSO pays both the cost of congestion management and of capacity expansion, so that he can balance the latter against the cost of prolonged congestion. He therefore has an incentive to make efficient investments in the network.

Here a choice has been made for a system in which the TSOs do not change the market prices, even though their redispatching actions have in fact changed the marginal generators in both countries. Rather, the financial transactions associated with redispatching are made outside the regular electricity market. A different approach to
redispatching is possible, by allowing the redispatching actions to influence the market prices in both countries. This alternative approach starts to resemble market splitting, as different prices develop in the two countries. However, this variation of redispatching will still cost the TSO, as opposed to market splitting which generates revenues (typically for the market operator). Here the choice was made for a system in which redispatching is performed completely outside the market for two reasons. This method corresponds to the way regular redispatching traditionally is applied within electricity systems. The second reason is that it results in a unique case in which congestion management takes place completely outside the market.

A disadvantage of redispatching is that the market does not receive any signals regarding congestion, and will therefore not adjust its trade patterns accordingly. In theory, this could lead to inefficient siting decisions of new generators; in practice, there are so many other constraints on the location of new power plants that the effect appears to be negligible. A second disadvantage is that there is a high potential for strategic behavior by the generators. Firstly, the TSO’s choice of generators is limited. The solution to the congestion may depend on only a small number or even upon unique generators. Secondly, it may be difficult for the generator to obtain true cost information from the generators. Traditionally, redispatching took place within vertically integrated companies. In an unbundled system, the generators will try to make a profit by overcharging the TSO or offering less capacity, depending on the case. A particular case is when a generating company has power plants on both sides of the congested network link. In this case, he may be able to contribute to the congestion by scheduling more production in the exporting market and less in the importing market. The generating company may benefit from this if it is subsequently called upon to relieve the congestion. Thus the company can create congestion and then be paid to relieve it, leading to a continuous income transfer from the TSO to the company.

Whether consumers benefit more in the short term from this system than from the previous options depends entirely on the manner in which the congestion rents in the previous options are spent and on how the money to fund congestion management is raised in the case of redispatching. At first glance, the congestion pricing methods appear less beneficial to consumers. They need to pay more because the congestion rents (the price of the scarce capacity) need to be paid. However, this does not need to cause a net loss for consumers. The apparent higher cost of the congestion pricing methods may be compensated, from a consumers’ point of view, by spending the congestion rent on something that consumers would otherwise have to pay, for instance by reducing transmission prices, or using the congestion rents for capacity expansion that otherwise would be paid by the network users. Conversely the TSO may pass the costs of congestion management on to the network users by raising.
transmission prices. In conclusion, it is clear that the type of congestion management method alone does not determine the effects on net consumer surplus. Consumer surplus also depends on the general design of the electricity market, including the transmission tariff structure and, in the end, on the details of which costs are passed through and how.

Example

The market parties trade as if the interconnector has abundant capacity. The market price therefore is 28 €/MWh. In an unconstrained market, 1200 MW would be imported into B across the interconnector. As it has a capacity of only 800, the TSO must redispatch 400 MW. He will cancel GA5 and reduce GA4 to an output of 500 MW. In B, the cheapest way to make up the lacking power is to purchase an additional 400 MW from GB4.

Again this leads to the economically efficient generation pattern. Financially, however, the situation is quite different from the preceding solutions. The starting position is that consumers pay and generators receive as if there is no congestion at all. This means that the consumers pay only 126,000. In first instance, generators would also have the same generation cost as in the case of no congestion.

The next step is that the TSO redispersches generation in an economically optimal manner until there is no congestion anymore. He does this by cancelling some generation in A and dispatching more generation in B. He reimburses generators that he dispatches by paying them their marginal cost of operation. Similarly, the TSO receives the avoided costs from the generators that are cancelled. We assume that the TSO knows the marginal costs of operation of all generators, so that he can dispatch as efficiently as possible. Thus he receives money from the reduced output of GA4 and GA5, but he has to pay more to make up for that in B. The net cost to the TSO is 1000 €/h.

In this system more generators earn just their cost of generation than in other systems, as a result of which the generators’ profit is lower (15 400 €/h). This is another reason why this system appears more attractive from a consumers’ point of view. However, keeping these costs low is contingent upon the TSO knowing the real marginal costs of each generator, a situation which may not exist in practice.

6.3 Analysis

After comparing the theoretical economic performance of the different methods for congestion management and considering their

---

77 Whether he can access this information in practice is doubtful. Generators would, of course, like to make a profit. Generators whose output is increased will try to receive more than just their marginal production costs. Thus they may engage in a sort of implicit bidding process, where they balance the advantages of being receiving a higher reimbursement from the TSO against the odds of the TSO choosing a different generator. If few generators are available for redispach this may give them market power. Generators which are curtailed will behave the opposite way: they will underestimate their marginal costs, so that their payments to the TSO are limited.
long-term effects, we will now evaluate the effects of the assumptions that underlie the model applied. This section concludes with an evaluation of some practical issues that impact their performance and feasibility.

6.3.1 Economic efficiency and income effects

In theory, each of the congestion management methods which were analyzed in this chapter is capable of achieving the least-cost dispatch of generation. The generators which are allowed to export via the interconnector are in each case selected upon their marginal production cost. As the interconnector capacity is given to the cheapest available generators, this means the overall cost of generation is minimized. Therefore all reviewed methods are economically efficient in the short term.

How much consumers pay depends, first of all, on the actual costs of generation and transmission and, secondly, on the profits that are made by the generators and the TSOs, that is, on the distributive effects of the congestion management method in place. In the short term the system costs are the same for the reviewed methods. This means that the net social benefit is the same for all these methods in the short term. As the revenues of generators and the TSOs equal the payments by consumers, one party’s gain means another party’s loss.

One cannot draw general conclusions regarding welfare effects from this analysis. In particular, one cannot assume that congestion pricing methods will cost consumers more than the corrective methods, even if the examples suggest this. The congestion pricing methods generate revenues which may be used for purposes which otherwise would need to be financed by consumers in a different way, such as network capacity expansion. They may also be returned directly to consumers, for instance by lowering the general network tariffs. The cost which the TSO incurs in the corrective measures, on the other hand, will be passed on to the consumers in one way or another.

It follows that undesired welfare effects of congestion management methods can be compensated elsewhere in the system, for instance through adjustment of the general transmission tariffs. The methods may differ with respect to how easily this can be done, however. In the form that was chosen here, redispatching leaves the generators with higher profits than the other options, and it may be difficult to return these profits to consumers. The rents from the congestion pricing methods can be used much more easily to the benefit of consumers.\(^7^8\)

\(^7^8\) The EC proposes to limit the possible applications of congestion rents, see the proposed Regulation Of The European Parliament And Of The Council on conditions for access to the network for cross-border exchanges in electricity, Art. 6.6.
An important difference lies in the effects of market power. Market power may develop if a constraint separates a relatively small part of the market from the larger European market. If prices in the smaller market are lower than in the general market, the smaller market is considered export-constrained. In this case, market power will probably have a limited effect. But if this market in import-constrained, this means that a certain volume of local generation is physically essential for meeting electricity demand. If the involved generating companies are aware of this, they have – in theory – unlimited market power, for without their plants there will be a power deficit. In practice, their room for increasing prices is limited by legal constraints, if nothing else. (For instance they may also need to balance short-term profits against the need to maintain relations with the regulator.) Because this market power has a physical cause, it will exist regardless of the congestion management method.

However, in case of congestion pricing, market power leads to higher consumer prices, whereas in case of redispatching market power will lead to higher payments by the TSO. This is an important difference, as the TSO may have fewer options for countering market power and its effects will be less visible, as the additional costs to the TSOs are socialized in the network tariffs. In case of congestion pricing, market power affects the electricity prices, which is more transparent. Because consumers are affected directly, they will have an incentive to counter the exercise of market power, for instance through legal procedures.

### Long-term effects

Ideally, a congestion management method is economically efficient in the short term and also provides efficient long-term incentives to both the network managers and to generation companies. None of the compared methods combines these goals. The different income effects produce different incentives for long-term behavior. Congestion pricing systems signal the cost of congestion to market parties, who may adjust their investment behavior accordingly if the congestion persists. Thus they may, in the long run, reduce congestion, although the effect upon siting decisions for new generators is probably limited. Congestion pricing methods provide no incentive to the network manager, however, to relieve the congestion; if he is allowed to keep the congestion rents, the incentive is even to increase congestion. Nevertheless, congestion pricing methods yield a revenue flow which can be dedicated towards projects to relieve the congestion. The corrective methods have precisely the opposite effect: while they do not provide market parties with an incentive to change their behavior to reduce congestion, they do provide the TSOs with an incentive to minimize congestion. This is an important advantage,

---

as the TSO can balance the costs of congestion against the costs of network capacity expansion.

It may be tempting to base interconnector capacity investment decisions upon the incentives and revenues which the congestion management methods provide. There are at least two significant problems with this information. First, while the marginal bid equals the marginal value of capacity, it does not say anything about the need for capacity expansion. Therefore the revenues from the congestion pricing methods bear no relation to an efficient level of investment in capacity. That depends on the potential benefit from further trade, which in turn depends on the generation bids that were not allocated capacity on the congested interconnector but that otherwise would have been in merit. This problem may be overcome in market splitting, where the market operator knows the entire supply and demand curves, which provides him with the information to calculate how much capacity is needed to reduce congestion, as well as the value to the market of capacity expansion.

The second problem is more significant. The marginal bid only indicates the momentary value of a marginal expansion, while investment in interconnectors takes a long time. Capacity expansion projects can take a decade or more to complete, at which time market conditions, and hence the demand for interconnector capacity, may have changed significantly. Moreover, network components last decades, and it is impossible to estimate the demand for network capacity so far ahead. Therefore the value of the information that is provided by congestion management revenues is limited, with respect to network investment decisions.

6.3.3 Barriers to trade

Explicit auctions are likely to create higher transaction costs than the other methods, as they require two separate transactions for cross-border trade of electricity, whereas the other methods require market parties only to make a single transaction. The complexity of auctions may pose a barrier to market parties, in particular to those who do not have the expertise to handle the associated risk. As a result, explicit auctioning may have a distorting effect upon the market. Small generators may be least favored, as for them the transaction cost of the auction is relatively highest. On the other hand, a well-designed auction may provide transparency to the market, which would reduce the barrier for newcomers.

The other methods appear to have lower transaction costs than auctions. However, this may be offset by the lower transparency. In the case of market splitting, the transaction costs of trades across a congested link are not different from those of trades which do not involve a congested link, so the congestion management system does not impose an extra barrier to trade.
6.3.4 Congestion in a network

This chapter has provided a mostly theoretical analysis of a simple case of congestion. The reality of the existing electricity grid is much more complex. One must take network effects into account when considering a solution. Two complicating issues may arise. First there may be parallel routes to the congested route, which may or may not also be congested. The second issue arises when multiple congested links need to be crossed for a single transaction.

If there are multiple parallel paths that connect two countries, the actual transmission capacity between the two countries is not simply determined by the sum of the capacity of the individual interconnectors, but influenced by the electricity flow pattern.\(^{80}\) The location of generators and loads influences the flow of electricity through networks, and therefore the occurrence of congestion. In general, little influence can be exerted upon the location of loads. Generation is more flexible, as most regions in Europe currently have an over-capacity of electricity generation. An efficient combination of a transmission pricing system and congestion management might be able to influence the generation pattern in such a manner that the network will be used optimally, meaning that it results in the economically most efficient dispatch of generation, given network constraints.

Of the investigated methods for congestion management, redispatching certainly has this potential. By definition the TSO can influence the generation pattern when he redispatches. In counter trading, this depends upon the type of contract between the TSO and the generators. The congestion pricing methods leave the decision which generators should run to the generating companies. Because the decision process in these methods is more decentralized, their success depends upon the incentives which these methods provide. If generators would receive perfect economic incentives which accurately conveyed to them the cost of transmission and congestion, their response could be optimal. However, the discussed congestion management methods do not provide such refined incentives and are therefore not likely to use the available network capacity optimally. However, it may be possible to refine them and adjust them to the conditions of a meshed network.

A second issue is how well a congestion management system performs when sales of electricity are made across two consecutive congested borders. Such a series of congested links may pose a significant obstacle to trade, unless the congestion management methods are easy to handle for market players. The same conclusions that were reached with respect to transaction costs can be applied to this case, but they will weigh even stronger. If capacity auctions are used, combinatory bidding must be implemented to allow

\(^{80}\) See among others Harvey et al. and Haubrich et al.
transactions involving the multiple congested links. Redispaching, of course, still pose no obstacle to trade. Explicit auctioning and market splitting could work, but would require a significant level of organization between all the involved TSOs and market operators.
7 Security of supply of electricity

7.1 Introduction

This chapter provides an analysis of the question of generation adequacy in competitive electricity markets. Secondly, a policy framework is developed for selecting among the policy options to maintain generation adequacy. The existing generation stock is considered adequate if it can be expected to meet demand under all reasonable conditions, considering normal outage rates. Concerns whether competitive electricity markets provide a sufficiently strong and early enough investment signal rose after the crisis in California’s electricity market in 2000 and 2001. Shortages in other places, such as in New Zealand, Scandinavia and, most recently, Italy, have fueled these concerns. A number of adjustments to the market design have been proposed with the purpose of stabilizing the volume of generating capacity. A systematic framework for the selection of such a capacity mechanism has not been developed yet, however.

The focus of this chapter is upon Europe, because European electricity markets have several specific features. First, most European markets do not have a mandatory power pool. Market parties may sell their electricity bilaterally and only need to notify the system operator of their physical programs. Second, many European markets have significant trade volumes with neighboring markets, while the connected market models often vary greatly. Third, hydropower plays a limited role in many European markets, the exceptions being Scandinavia and the Alp countries. The latter factor means that most European power markets are capacity-constrained, rather than energy-constrained.

Capacity mechanisms vary widely in the way they are intended to work and with respect to their implementation requirements. Some provide financial incentives to generating companies, while others control the volume of generating company. Some are designed for mandatory pools, which means they might need to be adjusted for implementation in Europe. In most cases, little attention has been given to the issue of trade: how to prevent the investment incentive from ‘leaking’ abroad, and how to make a capacity mechanism immune from regional shortages? This chapter develops a set of criteria to evaluate the different proposed capacity mechanisms, describes the advantages and disadvantages of the different capacity mechanisms and, most importantly, develops a framework for deciding which capacity mechanism to implement under which circumstances.

---

81 This chapter is based on the dissertation of De Vries (2004).
The next section starts with a brief summary of the reasons to implement a capacity mechanism. Section 7.3 develops a set of performance criteria which capacity mechanisms should meet. This set is used to evaluate proposed and actual capacity mechanisms in Section 7.4. However, none of the currently proposed capacity mechanisms appears satisfactory for implementation in a market without a mandatory pool but with strong interconnections, such as many European countries have. Therefore Section 7.4.6 proposes an innovative version of the capacity mechanism called reliability contracts. Section 7.5 is devoted to the development of a policy framework for the selection of a capacity mechanism. The conclusions are summarized in Section 7.6.

7.2 Why a capacity mechanism?

7.2.1 The narrow investment optimum

Electricity markets have a different dynamic from other markets due to three characteristics:

- Electricity is a strongly time-limited product. It cannot be stored, other than in pumped-hydro facilities, in a commercially viable way. However, the electricity supply system can only function in a stable manner if supply and demand are continuously balanced.

- The supply of electricity is only partly characterized by a gradually increasing marginal cost function. When all available generation units are producing electricity, no marginal increase is possible in the short term. As a result, the marginal cost curve ends with a perfectly price-inelastic section.

- The demand for electricity also is highly inelastic. This may be caused by the fact that there is no readily available alternative for most applications of electricity. At least as important is, however, that few consumers receive the required price information in time to adjust their behavior. Moreover, electricity consumption is usually measured over long periods, so consumers have no incentive to shift consumption from peak hours to off-peak hours. As a result, few consumers adjust their electricity consumption to the current price of electricity, so that the observed price-elasticity of consumers is extremely low. There are multiple experiments aimed at increasing consumer price-elasticity, but in most electricity systems their impact still is small (Nilssen and Walther, 2001; Roberts and Formby, 2001; Sæle and Grønli, 2001).

The combination of these three characteristics is the reason that most mechanisms which aid the clearing of other markets, such as a delay in the delivery of the good, consumers switching to other goods or higher prices leading to a reduction in demand, are not available in current electricity markets. This has significant consequences: wholesale electricity prices are highly volatile, and secondly, there is a chance of service interruptions.
With some modifications, the theory of spot pricing still holds, even if demand is assumed to be fully inelastic. The main consequence of insufficient demand price-elasticity is that there is a risk that the market does not clear, obviously in the case that physically there is not enough generation capacity available to meet demand. At this point, the market does not reach a price equilibrium and some load will need to be shed. The cost of installing so much capacity that the chance of load shedding would be reduced to zero would exceed the social cost of the load shedding which would be avoided.

However, if the market does not clear, it may be necessary to institute a price cap to protect consumers against overcharging (e.g. Ford, 1999; Hobbs et al., 2001c; Stoft, 2002). If consumers are not involved in real-time price setting, they otherwise might find themselves paying more for electricity than their value of lost load. This price cap needs to be determined carefully, as it impacts the attractiveness of investment in generation capacity. The price cap needs to equal the average value of lost load (VOLL), because at this price consumers should, on average, be indifferent whether they receive electricity or not. Stoft (2002) shows that in a perfectly competitive market, this results in an optimal level of investment in generation capacity, with an optimal duration of power interruptions. Therefore the theory of spot pricing still is valid, even if demand is fully inelastic. Price caps can be problematic, however, because it is difficult to determine the optimal level, as the value of lost load is difficult to measure (Willis and Garrod, 1997; Ajodhia et al., 2002).

Although theoretically sound, the reliance upon periodical price spikes to signal the need for peaking capacity has some significant weaknesses. To begin with, there is the risk that the price cap is set at the wrong level, resulting in over or under-investment. However, there are more fundamental issues. The first is that investment in peak generation units is quite risky, so that small distortions of the investment signal may have large consequences. The second is the argument that there is a positive externality associated with investment in peaking units, because security of supply is a public good (due to the network character of electricity supply). The third factor is the inevitable development of market power during periods of supply scarcity. These issues will be addressed in the next sections.

7.2.2 Market Failure

Now we will discuss a number of factors which may disturb the narrow investment optimum. The following types of market failure can be discerned (based, in part, upon Hobbs et al., 2001b):

- price restrictions,
- imperfect information e.g., regarding consumer willingness to pay or future supply and demand,
- regulatory uncertainty,
- regulatory restrictions to investment, and
- risk-averse behavior by investors.
The fact that a price cap may be needed to protect consumers against overcharging in times of scarcity represents a significant risk, because the optimal level of the price cap is difficult to determine. While the theory is clear that the price cap needs to be equal to the value of lost load, there are many methods of measuring the value of lost load with widely varying outcomes (see for instance Willis and Garrod, 1997; Ajodhia et al., 2002). The cost of erring is high. A price cap that is not equal to the value of lost load likely results in a sub-optimal level of investment in generation capacity.

Producers lack the information needed for socially optimal investment decisions (Hobbs et al., 2001b; Stoft, 2002). This increases the investment risk and therefore reduces the willingness to invest. In order to calculate the probability that peak units will operate and to calculate the expected return on investment, generating companies need to know both the stochastic distribution of the demand function (so they know the distribution of the frequency, duration and height of price spikes) and the expected development of total available capacity (Hobbs et al., 2001a). The exact characteristics of the demand function are difficult to estimate, especially in newly liberalized markets for which no long time sequences of empirical data are available. Moreover, the basic characteristics of demand change over time (for instance due to the introduction of new technologies) which reduces the validity of demand functions based upon historical data.

Regulatory uncertainty increases investment risk and therefore adversely impacts the willingness to invest. Regulatory uncertainty can be considered as a negative externality associated with changes in public policy. Especially in newly liberalized markets such as most electricity markets, regulatory uncertainty can be a significant factor. Consider, for example, a few of the policy changes which currently are underway in Europe:

- On November 25, 2002, the European Council on Transport, Telecommunications and Energy reached political agreement on amendments to the electricity and gas Directives and a regulation for electricity (EC, 2002).
- The European gas market is in the middle of a liberalization process. Most notably the development of the gas transport tariff system, including charges for flexibility and imbalance penalties, is highly uncertain. This has a considerable impact on a business plan involving today’s state-of-the-art gas-fuelled generators.
- Additionally, there is uncertainty about future European environmental rules, such as cooling water regulations or the specifics of the proposed CO₂ emissions trading scheme (EC, 2001).

A second source of regulatory uncertainty, with an equally significant impact upon the willingness to invest, exists with respect to the question whether a period with high prices will give cause to the government or the regulator to implement a maximum price or, if a maximum already exists, to lower it. Volatile prices are not only a
risk for investors, but also for regulators due to the public protests
they give rise to. Most electricity systems start liberalization with
ample capacity. In fact, the desire to reduce excessive reserve margins
was a motivation for liberalization. If, after the initial excess capacity
has disappeared, a period develops in which prices are many times
higher than their historical levels, consumers may consider this a
failure of liberalization and demand intervention. This occurred in
San Diego at the beginning of the crisis in California, when even a
brief period of high consumer prices proved politically unacceptable
(Liedtke, 2000). The political risk of being held responsible for high
electricity prices, whether these are economically efficient or not,
translates into a risk for investors of political intervention. Hence
price volatility itself brings about regulatory risk, at least until
sufficient experience has been gained with liberalized markets that
investors know whether they should expect political or regulatory
intervention or not (Oren, 2000; Newbery, 2001).

Obstacles to obtaining the necessary permits may be another cause of
underinvestment. While the social benefits of a proper licensing
process are not disputed here, it should be taken into account that
they may incur negative side-effects. Firstly, the permitting process
can be lengthy, thereby increasing the response time of generation
investment to an increase in demand. Especially in a situation of
incomplete information about the future development of supply and
demand, this may contribute to investment risk. A second effect of
increasing the lead time for the construction of new plant is that it
may contribute to investment cycles. This subject will be further
discussed below. A third effect of permits is that they may impose
additional requirements on generators, leading to operational
constraints to the response to market signals. An example is that
cooling water regulations may restrict operation during periods of
hot weather.

The theoretical approach by Caramanis et al. assumes that generating
companies behave in a risk-neutral manner with respect to
investment. This is not necessarily the case, especially when many
risks themselves are not well understood. Given the many
unquantifiable risks in a liberalized electricity market, it is not
unlikely that investors in generation capacity choose a risk-averse
strategy with respect to generation investment (Vázquez et al., 2002).
If all investors do so, none of them lose market share, so the penalty is
limited to a loss of sales during periods of supply shortage. However,
this loss of volume is small, compared to overall production of
electricity, and is likely to be more than compensated by the high
prices that develop during a period of supply shortage. Therefore a
collective strategy of risk-averse investment behavior is beneficial to
the generation companies, as long as this does not attract newcomers
to the market. Such a risk-averse investment strategy would lead to
less installed capacity than would be socially optimal.
7.2.3 Reserve capacity as a public good

Another argument is that there is a positive externality associated with the provision of generation capacity, as a result of which a competitive market will under-provide it. The existence of a chance of service interruptions is key to the existence of market failure in generation capacity, according to Jaffe and Felder (1996), because service interruptions are random. The more generation capacity is available, the higher the reliability of the supply of electricity is. Therefore, they argue, the presence of generation capacity in excess of the capacity which is contracted by market parties ('reserve capacity') provides an additional benefit to all consumers of electricity in the form of higher reliability of service.

This benefit to all users of the system is a positive external effect of the provision of capacity, as the owner of the generation capacity cannot charge consumers for increasing the reliability of service. As the added reliability is non-excludable and non-rival (the reliability of service to all consumers increases), the reserve capacity can be characterized as a public good. The same is true for the withdrawal of firm load: when a consumer reduces his load, system demand goes down and the chance of a shortage decreases. Withdrawal of load and the provision of additional capacity have the same positive external effect: they both increase system reliability. Because part of the socially optimal amount of generation capacity is a public good, liberalized electricity markets will tend towards an equilibrium volume of installed generation capacity which is lower than the social optimum. This analysis is corroborated by Pérez-Arriaga and Meseguer (1997), who consider generators to deliver three distinct products: energy, operating reserves and capacity reserves. This implies that when generators are not paid for their capacity reserves, they provide an external benefit.

In a market equilibrium, this positive externality would be reflected by consumers not revealing their true willingness to pay. If service interruptions are the consequence of, for instance, a 2% shortage of generation capacity, this means that service interruptions affect only about 2% of the customers at a time during a period of scarcity. (This is the maximum percentage of load that was disconnected during the crisis in California in 2001.) This means that each individual consumer can expect to be without electricity for 2% of the time that the shortage lasts. The consumers who caused the shortage by under-contracting therefore do not suffer the full consequences; instead, they still can consume as much electricity as they want for 98% of the time. In a market equilibrium, this means that those consumers who show a lower willingness to pay, benefit from those who show a higher willingness to pay and thereby attract more peak capacity. The public good character of reserve capacity therefore provides consumers with an incentive to understate their willingness to pay.
This effect could in principle be offset by the phenomenon, described in Section 7.2.1, that during periods of scarcity prices may rise above consumers’ willingness to pay, dependent upon the level of the price cap. Therefore, if the price cap is set high enough, the average price over time should still be high enough to recover investment in peaking units. However, Jaffe and Felder are right that during periods of ample generation capacity, the electricity price in an energy-only market does not include a premium for the availability of unused capacity; it merely reflects the marginal cost of production of the marginal unit. The combination of the fact that ordinary prices do not reflect the full social value of electricity, while scarcity prices may be too high, may lead towards a tendency of cyclical investment behavior.

7.2.4 Investment cycles

A year before the California crisis started, Ford (1999) published a paper in which he used a computer model to show that investment in electricity generation facilities is inherently unstable in a system with rules such as in California. His explanation is that investment is not aimed at dampening business cycles, which it would do if the right amount of new capacity became available at the right time, but at making a profit. Because investors tend to wait until they are reasonably certain that they can make a profit, and because they tend to overreact (in part because they do not know their competitors’ plans), Ford considers the interaction between the price signal which a power exchange provides and investment inherently unstable.

Ford’s argument is essentially that a combination of risk-aversion and an insufficiently long time horizon leads to a delay of investment. Due to the low elasticity of supply and demand, the price signal will not indicate scarcity until the capacity margin is so slim that the chance of service interruptions has become unacceptably large. The long lead time for new investment means that, once a shortage has developed, this shortage becomes worse before it is alleviated with new generation capacity. Ford’s argument is reinforced by the argument from the previous section, that generation capacity is undervalued during periods of abundant supply and overvalued during periods of scarcity.

Visudhiphan et al. (2001) contend that investment cycles are not inevitable, as long as investors are able to anticipate market developments. However, as we saw above, sufficient information about future supply and demand is lacking. In their simulation, Visudhiphan et al. also find that backward looking investment, that is, investment based upon recent experience in the market, will lead to investment cycles. Stoft (2002) arrives at the same conclusion. He notes that the distribution of price spikes may be such that investors would need to have a time horizon of several decades to determine the real average revenues from price spikes. If they use a shorter time
horizon, they are bound to overestimate or underestimate their expected revenues.

### 7.2.5 Long-term contracts

The investment risk in peaking capacity could be greatly reduced by the use of long-term contracts. Moreover, long-term contracts would reveal the expected future demand for peaking capacity to generating companies, as the retail companies (who purchase power on behalf of their customers) would reveal their peak supply demand when negotiating the contracts. This would improve the availability of information to generating companies and therefore reduce their investment risk. Long-term contracts would remove much of the price volatility, which is a risk for generators and consumers alike. An important additional benefit of long-term contracts for consumers would be the removal of the incentive to withhold capacity during periods of scarcity. So if there are so many benefits to long-term contracts to all involved parties, why are peaking units not covered by long-term contracts in practice? The answer is two-fold:

- There is an opportunity for consumers to free-ride, as a result of which retail companies are discouraged from engaging in long-term contracts for peaking capacity;
- Long-term contracts tend to have too short a duration to dampen the business cycle.

The opportunity for free-riding is caused by the fact that, for a large part, generators do not sell their electricity directly to the final consumers but to retail companies which act as intermediaries. This creates an opportunity for free-riding by consumers (Neuhoff and De Vries, 2004). This can be seen as follows. If long-term contracts are to provide an adequate signal to generators to install sufficient generation capacity, they must pay the generators the average cost of peaking capacity. During periods in which the peaking capacity is not used (which is most of the time), the spot market price will be below the cost of these contracts. Competitive spot market prices reflect the marginal cost of production, which does not include capital cost. Retail companies which hold long-term peak load contracts will therefore have higher costs, as they contribute to financing the capital cost of the peaking unit. Since rational consumers who are free to choose their retail company generally choose the cheapest one, they will therefore normally prefer companies without long-term contracts for peaking capacity. Hence, retail companies will be reluctant to purchase long-term peak load contracts during periods with ample generation capacity.

Long-term contracts for peak load become attractive again to retail companies during periods of scarcity, but then the reverse effect takes place. The retail companies will try to obtain average-cost based peak load contracts from generators, as they now are cheaper than spot prices. However, now the generators are unwilling to engage in these contracts, as they can sell their peak load for much higher prices in
the spot market. So the fact that supply companies intermediate between generators and end-users creates an opportunity for free-riding by consumers, which discourages timely investment in peaking units.

If retail companies were vertically integrated with production companies the dynamics change, but the companies would still not receive a sufficient incentive to invest in peaking capacity. Now the reason is the public good character of security of supply, which was mentioned above in Section 7.2.3. Vertically integrated companies would still have an incentive to keep insufficient peaking capacity, because this would allow them to sell electricity at a lower price than competitors with a complete portfolio. Consumers who purchase electricity from a company with insufficient peaking capacity would be free-riding on consumers who purchased (more expensive) electricity from a company with a complete generation portfolio.

The second failure is that long-term contracts generally are not long enough (Ford, 1999). Long-term contracts would need to extend beyond the current phase of the business cycle to cover at least the next phase in order to dampen the business cycle. The physical inertia of the electricity sector and the close relationship between demand growth and the general economy cause the business cycle of the electricity sector to be long, probably in the order of one or two decades. Because neither investors nor consumers have this long a time-horizon, the contract length generally is shorter. The result is, however, that long-term contracts do not represent a true average price of electricity, so that they tend to extend the current phase of the business cycle.

Even if generating companies would receive the appropriate long-term demand signals, there is a final problem with long-term contracts: the slow learning curve of consumers. The long time it takes to develop new capacity and the long life cycle of generating plant provides a serious obstacle to reaching an efficient equilibrium (Vázquez et al., 2002). If investment signals depend upon consumers entering into long-term contracts to hedge their risk of supply interruptions, consumers need to have the opportunity to learn how to find contracts which are attractive to them. As they would mainly learn through trial and error, this would require repeated periods of shortage and high prices. Due to the length of the business cycle, consumers have few opportunities to learn how to find attractive contracts which hedge their risk of supply interruptions. Moreover, it is likely that each period of shortages will result in changes of the market rules by the regulator, so that the learning curve would need to be started over.

The result is that consumers will probably never learn to cover all of their future demand with long-term contracts, so that electricity shortages will reoccur time and again and the market will never reach an equilibrium. Even if end consumers or their suppliers have a
proper incentive to enter into long-term contracts for peaking capacity, it will take unacceptably long before they would know what their actual (long-term) needs are and how to negotiate these contracts.

Even if consumers would consider a long enough period to average the swings of the business cycle, the risk to generation companies of such contracts would probably be too large. In the course of several decades, the fuel markets are likely to change, generation technology and environmental regulations may change and the uncertainty about the development of demand is large. So here is the Catch-22 of long-term contracts: a short time horizon does not isolate the contracts sufficiently from the business cycle, while a long time horizon carries too much risk.

We may conclude that investors lack the incentive, the time horizon and the education to engage sufficiently in long-term contracts. Moreover, consumers have an option to free-ride. They also lack the time horizon and the required sophistication. Therefore we may conclude that while long-term contracts may cover a significant portion of generation in a mature market, we cannot expect them to cover peak load capacity. Especially during a period of excess capacity and low prices, a shortage of long-term contracts for peaking capacity appears to be likely. Consequently, it is to be expected that this period of excess capacity is followed by a period of power shortages.

7.2.6 Market power

When high price peaks occur, the incentive may be quite large to withhold generation capacity from the market, as was demonstrated during the electricity crisis in California (Joskow and Kahn, 2002). When the capacity margin is slim, or when acute shortages already exist, the low price-elasticity of demand means that a small reduction in the supply of electricity may lead to steep price increases. In that case, even a small market share may provide enough market power to raise prices by keeping some generation capacity off the market. The temptation will be large to withhold generation capacity, for instance by listing generating units as requiring unscheduled maintenance.

Stoft (2002) points out that if the price cap is absent or very high, for instance equal to the value of lost load, the increase in profits from withholding can be so high that it becomes attractive even for small generators who would have to withhold the majority of their generation capacity. As a result, many generating companies, not just the large ones, have market power during a period of scarcity. The increases in profit which result from withholding are large, while it is difficult to take judicial steps against this behavior (because one would have to prove for each hour during which it occurs which generators were withholding illegally, and not legitimately out of
service). The strong incentives to withhold capacity when it is needed most is a fundamental weakness of electricity markets which rely on price spikes to signal the need for investment, even in the absence of other forms of market failure.

Besides the chance of being caught for abuse of market power, the only disadvantage of this strategy, from the point of view of generating companies, is that withholding electricity during a period of scarcity may cause such a crisis that it prompts a complete overhaul of the market design, as it did in California. Therefore an established oligopoly of large generators may choose a more stable, long term strategy.\textsuperscript{82} If generation companies are able to keep prices above the competitive level during normal market conditions, they may opt to overinvest in order to discourage new entry. The presence of excess generation capacity would serve as a threat to new market participants that the incumbents would be able to meet all demand, if necessary at a price equal to the marginal cost of production. A second reason for an oligopoly to invest more than would appear to be economic, could be that it would place an extra value upon reliability, because service interruptions would attract undesired (political) attention. Allowing an oligopoly to develop is hardly an attractive option, however, because it would undo many of the gains from liberalization (Newbery, 2002).

7.2.7 Trade between electricity systems

An entirely different aspect of the issue is how to ensure generation capacity in the presence of significant volumes of trade between systems. In theory, trade between liberalized electricity systems should not change the basic market dynamics. If the involved systems are liberalized in similar ways, trade between them only represents a scale increase. The scale of the system does not impact the question of generation adequacy, as it is addressed in this chapter. A benefit of a larger interconnected system is, however, more stability, as the relative impact of individual generators and capacity additions becomes smaller.

In practice, interconnected electricity systems often have quite different market rules, and the rules for using interconnectors are different from the regular transmission access rules within the systems. Therefore the markets which function within the interconnected systems are not fully integrated, but incompletely linked. This has repercussions upon the generation adequacy in the different markets.

\textsuperscript{82} Many markets are dominated by a few large suppliers. For example, the French, Belgian, Portuguese, Italian, Greek, Danish and Irish markets are dominated by one or two generators, while only three or four producers serve two-thirds or more of the markets in Germany, Austria, Sweden, the Netherlands and Spain (EU energy markets, 2002).
In the case of California, for instance, part of the problem was that investment in generation was not only lagging in California itself, but also in neighboring states. There, however, it did not lead to a shortage, but to a reduction of capacity reserves. When the weather suddenly caused a shortage, these states used their own generation resources to meet their own demand first, selling to California only what excess electricity was left. As a result, California, the importing state, bore the full brunt of a crisis, the roots of which were actually spread among a number of states.

In the European Union, a similar scenario is possible. Article 23 of the Directive allows member states ‘in the event of a sudden crisis’ to take unspecified ‘safeguard measures’ (Directive 96/92/EC). This can be interpreted as giving member states the right to close down interconnectors temporarily in an emergency. While there may be sound technical reasons for doing so, this means that in the case of a crisis, the European internal market may fall apart into a number of unconnected national markets. With respect to generation adequacy, this is an important issue as each country therefore should have sufficient reserves of its own to guarantee adequate supply under adverse conditions.

7.2.8 Risk asymmetry

The experience of California shows that the social costs of deviating from the optimum can be high. The outages in California totaled 30 hours, spread over six days. The largest amount of electricity not served at any given time was 1000 MW, although much of the time the outages affected a much smaller volume of load (Hawkins, 2001). To place these figures in perspective: during 0.3% of the year, a maximum of 2% of electricity demand was not served. Despite the seemingly small proportion of time that there actually were outages, the estimated social costs of the crisis are 45 billion USD (Weare, 2003).

Clearly, a relatively small deficiency of the system can have large negative social consequences. This observation is corroborated by efforts to estimate the value of lost load, which is the cost to customers of not being served with electricity. The value of lost load usually is estimated to be some two orders of magnitude higher than regular electricity prices, i.e. well in excess of 1000 €/MWh, although it is difficult to estimate accurately (Willis and Garrod, 1997). For example, in Australia the value of lost load was determined at 20,000 AUS$/MWh, which is about 11,000 €/MWh (Australian Competition and Consumer Commission, 2000).

From these considerations, the conclusion may be drawn that the provision of electricity is characterized by a strongly asymmetric loss-of-welfare curve. The loss of welfare due to under-investment by a certain amount is at least an order of magnitude higher than the loss
of welfare due to over-investment by the same amount. In this view, the likelihood of underinvestment due to the factors which were described in the previous section is a serious risk, which it is worth considerable cost to avoid.

Two strategies to reduce the risk of underinvestment to society can be proposed. One is to consciously overinvest in the electricity system. While the over-investment would constitute a loss of welfare with respect to the social optimum, it can be considered as a social insurance against the much greater negative effects of under-investment. The second strategy is to ‘flatten’ the investment optimum by changing the dynamics of the electricity system. If demand can be made more responsive to price, a shortage would result less quickly in random rationing and extreme prices. Instead, the least valuable loads would reduce their demand first. This would reduce the social cost of a shortage from the average value of lost load to the value of lost load of the least valuable customers.

7.3 Performance criteria

The previous section presented reasons to include a capacity mechanism, a mechanism to stimulate a socially desired volume of investment in generating capacity, as part of the design of competitive electricity markets. We may conclude from the previous analysis that the main goals for a capacity mechanism are to provide incentives to generating companies to provide an adequate volume of generating capacity, even in the presence of imperfect information, regulatory risk and/or risk aversion. Secondly, the incentives to withhold generating capacity during a shortage should be removed; rather, generating companies should receive incentives to maximize their output during shortages. Of course, the capacity mechanism should not introduce new opportunities for the exercise of market power.

In open systems (systems with a significant volume of trade with neighboring electricity systems), compatibility issues arise when neighboring systems do not implement a capacity mechanism, or a different one. This situation may develop in the European Union, where the issue of generation adequacy is left to member states (Directive 2003/54/EC). To be robust, a capacity mechanism should provide incentives to develop generating capacity within the system in which the capacity mechanism is implemented. If the capacity mechanism improves generator revenues by raising off-peak prices, for instance, the effect may ‘leak’ to neighboring systems to the degree that demand is met through imports.

Similarly, a capacity mechanism should provide a means to ensure that the generating capacity which it funded is available to those who paid for it. This means that if one country implements a capacity

---

83 For a similar view with respect to transmission capacity, see Hirst (2000).
mechanism, its benefits such as improved reliability and lower price spikes should accrue to the consumers in that country. Thus in an open market, a capacity mechanism must provide a means to ensure that during a regional shortage, international trade does not cause prices in the country with the capacity mechanism to be equal to the prices in connected countries without a capacity mechanism. Without such a provision, the capacity mechanism would be practically ineffective, with consumers paying for extra generating capacity the benefits of which would be exported.

Finally, there are some more general goals for capacity mechanisms. Improved price-elasticity of demand would reduce the volatility of electricity prices and hence investment risk. It would also reduce the probability of the need for service interruptions. Therefore policy intervention for the sake of securing generation adequacy should also stimulate the development of demand price-elasticity. Supply-side efficiency also is a general goal for market design: the capacity mechanism should not distort generating companies’ investment decisions (with respect to the size or type of generator) or the merit order of dispatch. Finally, any adjustment to the market design should be feasible, physically and institutionally, and, in the case of Europe, it should be compatible with bilateral, open markets.

Summarizing, capacity mechanisms should ideally meet the following criteria:

1. incentives to generating companies to provide an adequate volume of generating capacity
2. incentives to generating companies to maximize output during a shortage
3. no new opportunities to exercise market power
4. effectiveness in an open market
5. robustness against a regional shortage
6. stimulation of demand price-elasticity
7. supply-side efficiency
8. physical and institutional feasibility
9. compatibility with bilateral, open markets.

The next section evaluates the different proposed capacity mechanisms, but to keep the discussion of the capacity mechanisms concise, not every criterion will be discussed in each case. Instead, for capacity mechanisms with significant shortcomings the discussion will be limited to these shortcomings.

---

84 See De Vries and Hakvoort (2003) for an overview of capacity mechanisms.
7.4 Capacity mechanisms

7.4.1 Capacity payments

Several capacity mechanisms have been implemented or proposed to ensure an adequate level of investment in generating capacity. One of the first solutions was to provide generators with capacity payments. Their effectiveness is doubtful, however. First, the payments may not necessarily suppress the tendency towards investment cycles because they do not provide a clearer indication of the total demand for generating capacity than an energy-only market does. Second, it is nearly impossible to determine the optimal level of the payments (Vázquez et al., 2002). Because the demand curve for generating capacity has a steep slope and the supply curve has a gentle slope, a small error in the electricity price leads to a large shift in the equilibrium volume of generating capacity. This is what Oren (2000) refers to as the ‘classic prices versus quantities argument’. This argument holds for all price-based capacity mechanisms: given the many uncertainties regarding the relation between market prices and investment behavior, it is more difficult to obtain a desired level of reliability through a financial incentive than by regulating the volume of generating capacity.

7.4.2 Strategic reserve

Another common proposal is for the system operator to acquire a strategic reserve (sometimes also called a mothball reserve) of power plants, which are used in emergency cases. Purchasing old units itself does not change the overall volume of generating capacity; the effectiveness in improving reliability depends upon the investment signal that is sent to the market. This depends upon the price at which the electricity from the reserve is sold. As the reserve dispatch price determines the investment incentive, a strategic reserve should also be considered as a price-based capacity mechanism.

When the reserve capacity is dispatched, the market price will effectively be capped by the reserve price (until no more reserve capacity is available). The resulting reduction in price spike income reduces the average revenues of generating companies, as a result of which the equilibrium volume of generating capacity, which investors in the generation market provide, can be expected to be lower. The strategic reserve must make up for this deficit. Thus the lower the price at which the reserve is dispatched, the higher the volume of capacity in the reserve must be.

The difficulty with a strategic reserve (and with operating reserves pricing, the next capacity mechanism that we will discuss), is how to calculate the correct combination of reserve volume and dispatch price. As with all capacity mechanisms except capacity subscriptions (see Section 7.4.7), the regulator must first determine the optimal volume of generating capacity. Then it must be decided how large the strategic reserve will be. From this follows the volume of generating capacity.
capacity that is to be provided by the market. With accurate load-duration data, it can then be determined how much time the marginal generator that is to be provided by competitive generating companies (not by the reserve) will operate on average. With an estimate of the costs of this generator, the correct dispatch price of the strategic reserve can be determined. The price must be such that the marginal commercial generator can just expect to make a profit. Errors in this process will either lead to an under incentive to invest or to electricity prices that are, on average, higher than necessary to fund the necessary generating capacity.

In choosing the size of the reserve, there is a trade-off to be made: a small reserve with a high price has a limited effect upon market power, while a large reserve would cause the system operator to become a major actor in the generation market. This may not be desirable with respect to his independence.

A strategic reserve is a modification of an energy-only market: investment is still driven by price spikes, only reliability is enhanced through an extra volume of generating capacity in the reserve. This means that the same tendency exists towards investment cycles, although a large reserve may reduce it. Similarly, the incentive to withhold generating capacity remains unmitigated until the reserve is dispatched (at which point withheld generating capacity is replaced by capacity from the reserve).

Most relevant for individual European countries is that a strategic reserve is not robust against regional electricity shortages in a decentralized market. If consumers pay for a strategic reserve, for instance through an excise tax on electricity, they would expect the benefit of improved reliability of service. However, in a decentralized market, scarcity in a neighboring system will also lead to high prices in the system at hand. If there is sufficient interconnector capacity, trade between the two systems will cause the prices and the reserve margin will be about the same in the two systems. Contrary to systems with a mandatory pool, decentralized markets offer the system operator no possibility to direct electricity to its own consumers. Thus a strategic reserve is not effective in this case.

7.4.3 Operating reserves pricing

Similar arguments trouble the proposal for operating reserves pricing (Stoft, 2002). In this capacity mechanism, the system operator contracts a volume of reserve capacity through daily auctions. The system operator’s maximum willingness to pay for reserve capacity effectively caps the market price: when the spot electricity price exceeds that level, generating capacity that normally is sold to the system operator becomes available to the market. The idea is that the system operator’s demand for generating capacity will cause prices to rise sooner than otherwise, which provides an earlier investment signal, while the height of the price spikes is limited, so that average
prices remain the same. Thus an earlier and more stable investment
signal is provided. Advantages of operating reserves pricing are that
it does not disturb the merit order of the dispatch of generation
facilities, like a strategic reserve is apt to do; that the system operator
does not become a major vendor of electricity; and that it is easy to
implement.

However, the same obstacles will need to be surmounted as for a
strategic reserve to calculate the correct combination of the volume of
the reserve and the optimal willingness to pay for reserves. The latter
corresponds to the dispatch price of a strategic reserve and
determines the investment incentive for generating companies.
Secondly, while operating reserves pricing should dampen
investment cycles, it may not eliminate them. It is true that prices
start to rise sooner than in an energy-only market, giving an earlier
investment signal, but it is still the question how soon generating
companies will respond, and whether enough time is left then before
a real shortage develops. The main objection, however, is again that
operating reserves pricing does not appear effective in an open,
decentralized market. The higher prices during shoulder periods may
not even stimulate investment within the system, but lead to more
imports, while the limiting effect of the reserve upon peak prices may
prompt exports when neighboring systems are short of capacity.

7.4.4 Capacity requirements

The PJM electricity market on the East Coast of the USA, one of the
largest competitive electricity markets in the world, uses a system of
capacity requirements to maintain generation adequacy (PJM
Interconnection LLC, 2003). The principle of this system is that
government requires the load-serving entities to purchase enough
capacity credits from the generating companies to cover their own
peak demand, so that the system as a whole is ensured of enough
generating capacity to meet system peak demand plus a reserve
margin.\(^{85}\) The desired margin between generating capacity and peak
demand is administratively determined. Based upon the expected
total coincident peak demand of the loads served by each load-
serving entity (retail company or large consumer), the system
operator calculates how much generating capacity each load-serving
entity must purchase (PJM Interconnection LLC, 2003).

Reserve capacity may take the form of available generating capacity
or interruptible contracts. Generating companies may sell capacity
credits up to the volume of generating capacity that they have
reliably available. To this end the regulator rates the availability of
their generators. Capacity credits can be traded, so there is a
secondary capacity market. Load-serving entities include the cost of

\(^{85}\) ‘Load-serving entity’ is PJM’s term for parties that are licensed to provide
electricity to PJM consumers. An load-serving entity may be thought of as a
retail company or a large power consumer.
purchasing capacity credits in the price they charge final consumers for electricity. The requirement for load-serving entities to contract generating capacity in excess of the projected peak causes the capacity market to become constrained before the energy market does. Consequently, the incentive to invest in new generating capacity develops before the electricity market becomes constrained. If the capacity margin is large enough, this leaves enough time to bring new generating capacity on line before an electricity shortage develops.

The main advantage of capacity requirements is that they provide a robust way to maintain a certain capacity margin. The investment incentive does not depend upon the generating companies’ forecasts of future electricity prices, but upon the regulator’s projections of peak demand and the resulting capacity requirements. As a result, this system is less affected by information deficiencies and other sources of investment risk than an energy-only market, as long as the regulator’s capacity requirement is reasonably well chosen.

Capacity markets are complicated and need to be designed carefully. PJM has experienced several design problems. One is that a firm reserve requirement creates a perfectly inelastic demand for reserve capacity. Not only does this increase investment risk, it also provides a venue for the exercise of market power. Stoft (2002) suggests to make the penalty to load-serving entities who are short of their capacity obligations elastic: it should increase with the magnitude by which a load-serving entity does not meet its capacity requirement. This would reduce both the volatility of the capacity credit prices and the incentive to withhold generating capacity.

Another practical problem in the initial PJM design was that generators could ‘delist’ their capacity on short notice (Hobbs et al., 2001a). Thus they could earn revenues in the capacity market when electricity demand was low, and sell at high prices in the (neighboring) electricity market when that was more profitable. The solution was to increase the minimum duration for which a capacity credit may be sold, so generators need to decide for a whole season at once whether to offer capacity credits, and to require a longer notice for de-listing reserve capacity. However, the strength of these rules depends upon the penalty for non-compliance.

A related issue is that the system can be gamed by providing reserve capacity that is not actually operational: it rewards ‘iron in the ground’. In PJM the penalty to generators that have sold capacity credits but that are not available apparently is too low, given the probability to be caught, so the expected revenues from selling capacity credits exceed the expected amount of penalties to be paid (Hobbs et al., 2001b). Reliability contracts, a capacity mechanism that will be discussed in the next section, are specifically designed to provide generators with a better incentive to be available.
We may conclude that capacity requirements perform reasonably well on most of the criteria that were presented in Section 7.2.1, as solutions have been developed for most of the problems that were encountered in PJM. The main issue that has not yet been discussed is implementation in decentralized markets (without a mandatory pool) that have significant exchanges with neighboring markets that do not have a similar capacity mechanism in place. As mentioned above, PJM experienced problems with generators who sold capacity credits but exported their power when prices in neighboring systems were higher, so they did not actually contribute to the reliability of the PJM system. In a mandatory pool like PJM, the pool operator has the ability to ‘recall’ exports; this possibility does not exist in decentralized markets such as in Western Europe. This issue will be addressed in the Section 7.4.6.

7.4.5 Reliability contracts

Reliability contracts are designed as an improvement upon capacity requirements, with the purpose of providing generators with better incentives to make their resources available during periods of scarce supply. An independent agent – let us assume the system operator – purchases call options from generators on behalf of consumers. The call options provide the agent the right to the difference between the electricity spot price \( P_m \) and the option strike price \( P_s \). This price difference then is returned to consumers, so the net amount they spend, and generating companies receive, during price spikes is limited by the option strike price.

The volume of the contracts and the strike price are determined by the system operator and/or the regulator. The volume of reliability contracts is equal to the forecast coincident peak load plus a reserve margin, similar to a system with capacity requirements. The strike price should be above the highest marginal cost of operation of all the generators, to make sure it will not discourage any generator from producing. The system operator organizes an auction in which he purchases the contracts from the generating companies. Thus the option premium will reflect the generating companies’ expected loss of revenues at times when the market price exceeds the option strike price \( P_m > P_s \).

The system operator calls the options any time that the market price exceeds the option strike price. When the options are called, generating companies who have sold options pay the system operator \( P_m - P_s \) times the volume (in MW) for which they have sold options. An operational generator will receive \( P_m \) from selling electricity in the market, so his income will be equal to \( P_m \) minus his payment \( (P_m - P_s) \), which is equal to \( P_s \). The generating company is fully hedged against high market prices as long as the generator that backs the option contract is operational.

---

86 The description is based upon Vázquez et al. (2001).
A generator who has sold option contracts but happens to be unavailable when the options are called, still is required to pay \((P_m - P_s)\), but does not have any revenues to compensate these payments. Consequently the payments cause a net loss that increases with the market price, so their incentive to produce increases with scarcity. Therefore generating companies have a strong incentive to make their capacity available when the options are called, which is when electricity is scarce. This is one of the main advantages of this system. A second advantage is that generating companies have an incentive to sell a volume of call options equal to their expected output: selling less would lower their revenues, while selling more would expose them to price risk.

For consumers, the effect is that the system operator has ‘purchased’ a price cap equal to \(P_s\). As this limits the average revenues of generating companies, the latter will demand a price for selling the option contracts that corresponds to the expected loss of price spike revenues, which is the sum of \((P_m - P_s)\) over all hours that \(P_m > P_s\). As the option price is determined by the generation market in a competitive auction, it should reflect generators’ expected price spike revenues. If the system functions well, there should be no net cost to consumers, as the price paid for the reliability contracts should be offset by the reduction of price spikes. If the reliability contracts induce a higher volume of generating capacity, the costs of the extra generating capacity should be offset by the benefit of increased reliability (if the volume of reliability contracts was chosen correctly).

Reliability contracts combine the advantage of capacity requirements of a clear, capacity-based investment signal with better operational incentives and therefore appear an attractive solution. There are two issues, however. The first is that the auctions in which the system operator purchases the reliability contracts must be carefully designed to minimize generator market power. Vázquez et al. (2003) present some suggestions. Key is to minimize the barriers to new market entrants. To this end, auctions should be held a number of years in advance, so parties without generating capacity may first sell reliability contracts in the auction and then construct generating capacity to underlie their options. The risk for newcomers is also lowered by increasing the duration of the reliability contracts. However, this reduces the auction liquidity, so there is a trade-off to be made.

The second issue is compatibility with decentralized markets. The original proposal by Vázquez et al. (2002) was designed for a mandatory pool system, in which the system operator has greater control over the market transactions than in a decentralized market. The first question is how to accommodate bilateral power contracts. Generating companies who have sold their output through long-term bilateral contracts lose on the option payments, even if their output equals the volume of option contracts that they have sold. Vázquez et al. (2004) propose to let these generating companies make the option
payments, but to return the payments through parallel contracts to the extent that the generating company can show that he has sold his output through a long-term bilateral contract. The net effect is the same as in the original version: the generating company receives the strike price for all capacity that was not sold through bilateral contracts and pays the market price minus the strike price for unavailable capacity, when the options are called.

This version of reliability contracts should also be robust against a regional shortage. As the options hedge consumers for their full demand, the load-serving entities can bid any price, if necessary, to obtain enough electricity for their consumers, as they are fully hedged against price spikes. This should allow them to out-bid competitors from neighboring energy-only markets, who should not be willing to bid more than their value of lost load.

While this proposal works on paper, it may be vulnerable to gaming. First, the measures to mitigate market power in the auctions must be balanced with the need for liquidity. Second, the large financial flows that circulate between the generating companies, the system operator and consumers would provide a natural focus point for manipulation.

7.4.6 Bilateral reliability contracts

An alternative, described by De Vries and Hakvoort (2004), would be a bilateral variant of reliability contracts, in which the load-serving entities purchase the option contracts.87 This is more similar to PJM’s system of capacity requirements, except that the load-serving entities purchase options instead of capacity credits. By requiring the load-serving entities to purchase options for power in excess of their demand, it is ensured that they will always be able to purchase enough electricity to meet their demand at a predetermined maximum price. The financial flows would be straightforward: the load-serving entities pay the generating companies the option premium, and the generating companies pay the load-serving entities the market price minus the strike price ($P_m - P_s$) when the options are called. There is no need for financial involvement of the system operator or for reimbursements to generating companies with bilateral contracts. Rather than reimbursing the end consumers for high prices, bilateral reliability contracts limit the purchasing costs of load-serving entities and hence prevent the need for high end consumer prices.

Bilateral reliability contracts could be adjusted to reflect the preferences of the contract parties. For instance, if parties agree to option contracts with lower strike prices, the options would converge with ‘regular’ bilateral contracts, in which electricity is sold for a

87 The seminal ideas for this variant are provided by Oren (2000) and Vázquez et al. (2004).
specific price. Therefore there is no need to make this system compatible with bilateral contracts. The regulator would need to apply some limitations, however. The strike price would need to be limited to a maximum; if it were allowed to be as high as the average value of lost load, the option premium would be zero and the option contract would be meaningless. In this case, load-serving entities would be tempted to free-ride by covering their option requirement with these contracts. This would eliminate the investment incentive that the option contracts are intended to provide, as the free-riding opportunities that lead to an under-development of long-term contracts in an energy-only market would remain (De Vries and Hakvoort, 2003). Similarly, the minimum duration of the contracts would need to be regulated to provide a sufficiently stable investment signal.

A system of bilateral reliability contracts eliminates the need for auctions. A decentralized market should be more robust against the exercise of market power than the auctions for two reasons. First, because the load-serving entities compete with each other in the market for reliability contracts they have a stronger incentive to minimize costs than the system operator has in the centralized variant, as the system operator can pass the cost of the reliability contracts on to the consumers. Second, the load-serving entities have more options to mitigate market power, such as developing demand-side management programs to reduce their peak load and installing their own generating capacity.

The main question about bilateral reliability contracts is whether they can be made to work in a market in which generating companies are vertically integrated with retail companies, as is common in Europe. Option contracts between the retail component and the generation component of a company are meaningless. It might be necessary to rate the availability of the company’s generating assets, like PJM does, and to subtract the in-company volume of available generating capacity from the company’s obligation to purchase reliability contracts. This could diminish one of the fundamental advantages of reliability contracts, which was that they stimulate generating companies to make a best estimate of the availability of their generators when they sell the option contracts. This disadvantage needs to be weighed against the advantage of avoiding a central auction of options, as in the reliability contracts proposal.

7.4.7 Capacity subscriptions

A fundamentally different capacity mechanism, which promises to be the most market-oriented of all, is the system of capacity subscriptions (Doorman, 2000). This capacity mechanism directly involves consumers by requiring them to purchase electronic fuses which, when activated, limited their electricity consumption to a predetermined capacity. As a result, a capacity market develops between generating companies and consumers, which has the triple
advantages that it provides consumers with an incentive to limit their peak consumption, that it produces a clear signal that indicates the volume of generating capacity that consumers wish to have available reliably, and that it provides a steady revenue stream with which to cover the costs of generating capacity.

The original proposal does not appear robust against inter-system trade in a decentralized market. Generating companies who have sold capacity subscriptions to consumers within the system could still sell their output outside the system. A solution could be a financial version, which again uses option contracts (De Vries, 2004). To ensure that peak consumption is indeed within the contract limits, real-time meters would be required. Both variants have significant implementation requirements, as a result of which they do not appear feasible in the near term. Therefore we will not discuss them further.

7.4.8 Conclusions

As was mentioned in the introduction to this section, capacity mechanisms that provide a capacity signal are preferable over those that provide a price-related investment signal. Not surprisingly, the most effective capacity mechanism that has been tried in practice is PJM’s system of capacity requirements. Unfortunately, it appears that this system cannot be implemented in its current form in most European markets. There would be a risk of ‘leakage’ in case of unilateral implementation: if one country implements a capacity requirement, its reserve capacity could be sold outside the country in case of a regional shortage, so that the net contribution to the reliability within the country would be diluted. Reliability contracts could be designed to be robust against this effect, both a central and a bilateral variant. This would be an innovative, untried option, however, which entails a higher risk of policy failure.

7.5 Policy framework

The previous sections provided a brief overview of the advantages and disadvantages of a number of capacity mechanisms. Which one should be implemented, in particular in a decentralized market? Should a capacity mechanism be implemented now or should we wait until we have more evidence of the dynamic nature of electricity markets? This section discusses the policy choices, with a focus upon European electricity systems.

7.5.1 Implementation as a precaution?

The first question to be resolved is whether a capacity mechanism is called for now or whether we should wait and see how the market develops, considering the lack of empirical evidence of market failure?

Policy choice 1: Should a capacity mechanism be implemented as a preventive measure, which is easier, but for which the need has not been demonstrated, or should it only be implemented when the...
Waiting: risk that policy intervention will be too late
Waiting entails a significant risk, because it is not possible to monitor the market and forecast generation adequacy with sufficient certainty, far enough into the future, to allow time for policy intervention when it becomes apparent that a shortage of generating capacity looms. Not only does the development and implementation of a capacity mechanism take time; generating companies will also need time to evaluate the capacity mechanism’s implications in order to adjust investment strategies and, last but not least, it will take time to construct additional generating capacity in response to the new capacity mechanism.

Implementation during a shortage difficult
If, in the meantime, the volume of generating capacity drops below the level that the capacity mechanism is designed to obtain, a difficult transition period will follow. During this period the reliability of service will be lower than desired, while the inability of the market to immediately provide the desired volume of generating capacity may cause high capacity prices in a capacity market or in a system with capacity subscriptions.

Preventive implementation easier
Implementation of a capacity mechanism during a period of excess capacity, on the other hand, is much easier, as it would not require an immediate physical reaction from the market. The market could continue to reduce the capacity margin until the limits of the capacity mechanism would be reached, after which it would stabilize. The smoother transition and the lower risk to the reliability of service are arguments in favor of a ‘preventive’ strategy.

The need for a capacity mechanism is disputed
However, despite the previous analysis, the need for a capacity mechanism is not proven. In fact, it is a matter of hot dispute, with opponents arguing that one should not meddle with the market if there are not clearly demonstrated reasons for doing so. The talk about a capacity mechanism itself would create uncertainty and therefore be contra-productive.

Politically, intervention is risky without a demonstrated need
Politically, however, the balance may shift in the other direction. Implementation of a capacity mechanism is a significant intervention in the electricity market. Without a clear sense of urgency, it may be difficult to gain support, both political and from the sector, for such a change. The expected social costs of not taking any action appear much higher than the implementation costs of a capacity mechanism. However, due to the long time scale at which the generation sector develops, the resulting political repercussions will probably not affect the political leaders who currently are in office.

7.5.2 Unilateral or regional implementation?
Strongly interconnected electricity systems face the question whether to implement a capacity mechanism themselves or whether to attempt to find a regional solution. The latter is not only more
efficient economically, it also is easier and there are more suitable capacity mechanisms available. A system of capacity requirements, for instance, can be implemented without much difficulty in a decentralized market if there are no significant imports and exports. If there are, one of the variants of reliability contracts needs to be used. However, regional implementation of a capacity mechanism may take much time, especially in a network with as many constituting systems as the continental European electricity network. There may not be enough time to develop a regional solution before the first investment cycle develops. A dilemma is the consequence:

*Policy choice 2: Should importing countries implement a capacity mechanism unilaterally, despite the distortion of the international market, or rely upon imports and hope that a collective solution will be developed in time?*

In Europe, the Netherlands and Italy import a large proportion of their electricity. Therefore these countries face the difficult choice between unilateral implementation of a capacity mechanism, which would be more expensive or less effective, or waiting for a European solution, which might take too long.

### 7.5.3 Self-reliance?

If the choice for unilateral implementation is made in a system with strong interconnections, another question immediately presents itself. Should physical self-reliance be the goal? Alternatively, to which degree can imports be relied upon in the long term? The issue is not only the physical availability of imports, but also the price at which they are available. Capacity mechanisms tend to reduce the price volatility of electricity markets; some provide an upper limit to the payments for energy. Imports from energy-only markets could undo this effect, because they would cause price spikes in neighboring systems also to be imported, which would leave consumers to pay both for the capacity mechanism and for price spikes. This would undermine one of the main advantages of having a capacity mechanism.

For these reasons the choice may be made to become self-reliant, if neighboring systems do not implement a similar capacity mechanism. The cost of self-reliance may also be high, however, for electricity systems with a large share of imports. They may compromise by requiring a lower reserve margin in their capacity mechanism than would be considered optimal in an isolated system. While this reduces their security of supply to the extent that the imports are not dependable, it also reduces the cost of supplementing these imports with presumably inactive back-up generation.

*Policy choice 3: If an interconnected electricity system chooses unilateral implementation of a capacity mechanism, should it become fully self-reliant? If not, to what degree should it depend upon imports?*
7.5.4 Innovativeness

The analysis in Section 7.4 showed that the more innovative variants of reliability contracts promise to be more effective, in particular in decentralized markets, but the lack of experience casts some uncertainty upon their practical merits. Theoretically, they should provide better incentives to generating companies and be robust with respect to inter-system trade. However, the vulnerability of these untried systems to gaming, for instance, is unknown. This raises another policy dilemma:

*Policy choice 4: Should a capacity mechanism be chosen that has been tried in practice, but has known flaws, or should the choice be made for a more innovative system with better theoretical incentives, but unknown flaws?*

In the case of unilateral implementation in a decentralized market with strong interconnections, the only choice is to implement the one of the innovative variants of reliability contracts. The alternatives are to do nothing (and perhaps trying to achieve a regional solution) or to implement a capacity mechanism of which the effectiveness is uncertain. In a system with a mandatory pool, PJM’s system of capacity requirements may be implemented, also if it has strong interconnections.

7.5.5 Short-term versus long-term options

The choice of capacity mechanism depends upon the specific circumstances of the system within which it is to function. If a capacity shortage already is looming, it may be necessary to implement a capacity mechanism that can be implemented quickly, as a transition measure, even if it does not meet all the criteria. Capacity payments, a strategic reserve and operating reserves pricing are relatively easy to implement, which makes them attractive as short-term solutions. Unfortunately, their effectiveness is limited and they entail a risk of distorting investment incentives. Whether to implement a short-term solution is a judgment call: if it is estimated that enough time remains to develop a more elaborate, but also more effective and efficient capacity mechanism, this will be preferable.

*Policy choice 5: Should a capacity mechanism be chosen that can be implemented quickly, or one that requires more implementation time, but probably also more effective and efficient?*

If the decommissioning of old units threatens the capacity margin, the system operator may choose to purchase them as a strategic reserve (providing he has the authority to do so). Creating a strategic reserve this way was Sweden’s response to concerns about generation adequacy in recent years. An alternative that can be implemented just as easily is operating reserves pricing. Expanding the operating reserves when the reserve margin is below the target level would immediately create an investment signal. The disadvantages of these methods are that their effectiveness in stimulating investment is
uncertain (so an investment cycle may yet develop), that they mitigate, but do not eliminate the problem of capacity withholding in the electricity market, and that they are not robust against regional shortages. Therefore they should only be considered as temporary solutions. If more time is available (in the order of five to ten years before a shortage is projected), a version of reliability contracts appears more effective and efficient.

7.5.6 Overview of the policy choices

The above policy choices are summarized in Figure 11. The diagram shows the consecutive choices that present themselves, as well as to which capacity mechanisms they lead. The first choice to be made is whether a capacity mechanism will be implemented right away, as a precaution, or only when it becomes clear that the market is not providing sufficient generating capacity. Waiting is a risky policy, because failure of the market to provide sufficient generation capacity cannot be predicted far enough in advance to allow time to implement a capacity mechanism.

The next issue is whether regional implementation is feasible, as this is preferable over implementation by individual systems within a larger interconnected network. If regional implementation of a capacity mechanism does not appear possible in time to avoid a shortage, individual systems may decide to take action. If they are weakly interconnected, the absence of regional measures does not matter much, as all options are still open.

Development of a capacity mechanism jointly by the countries that form the continental European electricity network, or even for part of it, may take too long for some member systems. The Netherlands and Italy are the largest importing countries in the EU (UCTE, 2002c), which may give them cause for concern with respect to future reliability. If these countries would decide to implement a capacity mechanism independently, they would need to choose one that is robust with respect to imports and exports. Unilateral implementation of a capacity mechanism in a strongly interconnected systems is a difficult issue. None of the capacity mechanisms presented in Section 7.4 appears robust to a regional shortage, except the two pool-based systems (capacity requirements and reliability contracts). This means that in a decentralized market one of the innovative versions of reliability contracts needs to be implemented.

The ovals on the right hand side of Figure 11 indicate which options are available to which European countries. The interconnected electricity grid as a whole is so large, relative to the exchanges with its neighbors, that it can be considered as an isolated system for the purposes of security of supply. Examples of other relatively isolated electricity markets are those in the UK, Ireland and the Iberian Peninsula. These markets are in the comfortable position of having all
options available, because they also are not under pressure to implement a temporary solution as a transition measure.

In Nordel, the shortages in Sweden and Norway have prompted implementation of a strategic reserve and a form of operating reserves pricing, respectively, as short-term solutions while a longer-term solution is being developed.

- **Norway and Sweden: emergency measures**
  - **Implementation issues**
    - If the choice is made for capacity requirements, the ample experience in the PJM system provides the opportunity for empirical study of this capacity mechanism. Much has already been written about PJM's ICAPs; however, when considering implementation of capacity requirements in another system, the potential impacts of the...
differences between the two systems should be assessed. For instance, large imports have not been an issue in PJM.

Unilateral implementation of an effective and efficient capacity mechanism in one of Europe’s decentralized markets requires innovation. None of the available systems are fully satisfactory, with the possible exception of capacity subscriptions in a closed system. As there are some questions about the feasibility of the latter, the possibilities of developing reliability contracts should be further explored.

New systems must be thoroughly tested with respect to their ability to stabilize the generation volume in the presence of insufficient information regarding future supply and demand conditions and risk-averse behavior by both producers and consumers. In addition, they should be robust against the exercise of market power among generating companies, both in the short and the long term.

The art of developing a system based upon reliability contracts is to guard against new possibilities for strategic behavior. In general, the combination of a regulated volume of generating capacity and the fact that generating capacity cannot be expanded on short notice creates an opportunity for capacity withholding somewhere in the system. In the case of reliability contracts, the generating companies may manipulate the contract auction; in the case of capacity subscriptions, the generating companies may be able to artificially raise the prices of the subscriptions. The vulnerability of these capacity mechanisms must be tested, in a model and/or in practical tests, before they can be implemented with any confidence. Other opportunities for manipulation may occur through exchanges with neighboring systems with different market models. Implementation of a capacity mechanism should lead to a larger reserve margin and lower prices at times when the system otherwise would have been under stress.

Care must be given that the consumers who pay for reliability also benefit from the capacity mechanism. In this respect direct contracts between consumers and generating companies, such as in a system of option requirements or capacity subscriptions, appear more robust than reliability contracts in which a central agency purchases options on behalf of the consumers.

7.5.8 Assignment of responsibilities

It is the task of the system operator to preserve the operational reliability of the electricity system. To this end, he contracts system reserves (also called regulating power) with which he can correct imbalances between supply and demand in real-time. This obligation places system operators in energy-only markets in an awkward position with respect to the long term, as they do not have any means to influence the volume of available generating capacity. A survey of European countries shows that the responsibilities for generation
functions generally are restricted to monitoring by the system operator or by a government agency (UCTE, 2002a). In some cases, there is a planning requirement, however without a means to implement the plans. The actual provision of adequate generation resources is generally left to the market in Europe.

If a choice is made to implement a capacity mechanism, responsibilities need to be assigned for:

- choosing the desired level of reliability of electricity service (except in the case of capacity subscriptions)
- operational decisions regarding the capacity mechanism
- how to monitor and enforce the system.

Except for capacity subscriptions, all capacity mechanisms have in common that the desired generating capacity margin is the same for all consumers. The choice of the level of generation adequacy, which determines system reliability, could in theory be made through a benefit-cost analysis. The marginal cost of providing a large capacity margin should equal the marginal social benefits of the resulting reduction in power interruptions. However, these calculations are difficult to make, in particular because the social cost of service interruptions is difficult to determine. As a result, the level of reliability becomes a political choice, in which the cost of electricity is weighed against the perceived acceptability of occasional service interruptions.

The second issue is who makes the operational decisions. This depends upon the capacity mechanism that has been chosen. In a centralized system, such as operating reserves pricing or a strategic reserve, the system operator decides when to dispatch the reserve units. Capacity requirements and reliability contracts leave this decision to the market: they place the responsibility to provide a certain level of generating resources with the market. The same is true of capacity subscriptions.

Monitoring, finally, is a function that should be performed by an independent agent, so a government body (such as the regulator) and the system operator are likely candidates. Table 3 shows an overview of the distribution of responsibilities under the different capacity mechanisms.
### Table 3: Responsibilities with respect to generation adequacy

<table>
<thead>
<tr>
<th>government or system operator</th>
<th>producers</th>
<th>consumers</th>
</tr>
</thead>
</table>
| energy-only market           | • operation of the reserve capacity  
• monitoring                      | determination of the reserve margin |
| strategic reserve            | • determination of the reserve margin  
• operation of the reserve capacity  
• monitoring and enforcement   |
| operating reserves           | • determination of the reserve margin  
• operation of the reserve capacity  
• monitoring and enforcement   |
| capacity requirements        | • determination of the reserve margin  
• monitoring and enforcement   |
| reliability contract         | • determination of the reserve margin  
• monitoring and enforcement   |
| capacity subscriptions       | • monitoring and enforcement          |
|                              | operation of the reserve capacity     |
|                              | determination of the reserve margin   |

7.6 Conclusions

Competitive energy-only markets appear prone to investment cycles. Resulting episodes of scarcity may provide generating companies with substantial market power: by withholding generating capacity, they may be able to raise the electricity price substantially. These factors, combined with the asymmetry of welfare losses due to deviations from the optimal volume of generating capacity (the social costs of having insufficient generating capacity increase faster than the social costs of excess capacity) are reasons to implement a capacity mechanism. This chapter provided an analysis of investment in competitive electricity markets, evaluated possible capacity mechanisms and developed a policy framework for selecting the most suitable capacity mechanism. The main factors that influence the choice are how much time is available, whether the market in energy-only markets vulnerable to investment cycle and manipulation.
which the capacity mechanism is to be implemented has a mandatory pool and whether it has strong interconnections.

Capacity requirements (PJM’s system of ICAP) and reliability contracts best meet the policy goals of stabilizing the volume of generating capacity in an efficient way. Capacity requirements have as an advantage that practical experience is available, whereas reliability contracts promise to provide more efficient incentives to generating companies. Implementation of these capacity mechanisms in European markets is not straightforward, however. Both capacity mechanisms were designed for markets with a mandatory pool, which provides a means to control to which consumers electricity is delivered. This is an important issue when the capacity mechanism is to be implemented unilaterally in a market with significant interconnections such as exist in many EU member states.

The EU leaves policy with respect to generation adequacy to the individual member states. If a member state wishes to take measures to secure a certain volume of generating capacity, it would want to ensure that the generating capacity is available to its own electricity consumers. In a mandatory pool, this is easier to achieve than in a decentralized market, which happens to be Europe’s dominant market model. This does not mean that European countries should institute mandatory power pools, but rather that the proposed capacity mechanisms need to be adjusted. If not, the capacity mechanisms will probably be effective in stimulating investment in generating capacity, but it will be difficult to prevent exports from the generators that were financed through the capacity mechanism.

In order to ensure that the consumers who paid for the capacity mechanism also reap its benefits, the best solution appears a form of mandatory call options between producers and a party that acts on behalf of the consumers, ‘bilateral reliability contracts’. A bilateral version would resemble capacity requirements, except that load-serving entities would purchase options rather than capacity credits. It is also possible to adjust the proposal for reliability contracts to a decentralized market in such a way that it is robust against trade.

Making a capacity mechanism that is implemented by a single market robust against international trade means that it distorts international trade. Therefore joint implementation by a group of strongly interconnected systems would be much preferred. Implementation of a capacity mechanism in a large region, for instance in the continental Europe, would reduce the need to make it robust against trade, as the impact of trade outside the region would be limited. In this case, capacity requirements could be used without much need for adjustment, so the choice of capacity mechanism would be wider. Therefore stronger EU policy in this field is called for.

In the absence of a regional capacity mechanism, individual countries are faced with a dilemma. To do nothing entails a risk of the
development of an investment cycle, whereas to implement a capacity mechanism unilaterally is complicated and carries a risk of regulatory failure. Unilateral implementation also has a risk of being superseded by European policy later on, while waiting for this policy to develop may mean that the investment incentive comes too late to avoid a first investment cycle.
Literature


Algemene Energieraad (AER) 2004. Behoeftzaam Stroomopwaarts, Beleidsopties voor de Nederlandse Elektriciteitsmarkt in Europees Perspectief, Den Haag, AER.


CEC (Communication from the Commission) 2006b. “Sector Enquiry under Art. 17 of Regulation 1/2003 on the gas and electricity markets (final report)”, SEC 1724


Dehdashti, E.S. 2004. ‘Developing countries; restructuring with benefits from competition (or not)’. IEEE power & energy magazine September/October 16-23.


Doorman, G. 2000, Peaking Capacity in Restructured Power Systems, Thesis (Ph.D.), Norwegian University of Science and Technology, Faculty of Electrical Engineering and Telecommunications, Department of Electrical Power Engineering.


Klink, A. 1990, Publieke gerechtigheid: een christen-democratische visie op de rol van de overheid in de samenleving (rapport van het Wetenschappelijk Instituut voor het CDA), Houten.


Minister van Economische Zaken 2004. Leveringszekerheid in een geliberaliseerde elektriciteitsmarkt, Den Haag, Ministerie van Economische Zaken.


and Time-Tables” on Friday, 9 September 2005 in Brussels, Sessa. Obtained from:


http://www.ppic.org/content/pubs/R_103CWR.pdf.


Annex: The design of the Dutch gas sector

A.1 Goals and functions

The second European Gas Directive (2003/55/EG) is much firmer in its requirements to separate commercial and monopoly tasks. Who will be made responsible for which tasks, however, depends upon choices which still are to be made. Which activities should be regarded as monopolies? How should they be regulated? This chapter identifies the choices by analyzing how the institutional arrangement of the necessary functions in the gas sector influences the opportunities for achieving the stated policy goals.

Like for electricity, the general policy goals for the gas sector are ‘affordable, reliable and clean’, which is to say an economically efficient production, good reliability of service and an acceptable environmental impact. The idea is that the policy goal of affordability is achieved through liberalization, that is, by introducing competition where possible in order to reduce the costs of production and supply. For the monopoly activities, regulation will need to ensure affordability and quality. So far, the general approach is the same as for the electricity sector. What makes gas special in the Netherlands is the presence of large supplies; optimal utilization of them can be considered as an additional policy goal.

Divergent interests play a role in the utilization of natural gas supplies. The Dutch gas supplies contribute significantly to the security of supply of gas in Northwest Europe. This contribution is increased by the small fields policy (see page 11), which is intended to spare the Groningen field and utilize the other fields first. It makes it possible to deploy Groningen’s flexibility to meet the seasonal fluctuations in supply and demand and to handle temporary supply interruptions. If the large domestic supplies and the flexibility of the Groningen field are utilized cleverly, in the future they may preserve the Netherlands’ role as a hub in the gas supply of Northwest Europe. Utilization of empty gas fields for storage offers perspectives for reinforcing this role. An entirely different issue is the Dutch State’s interest in the revenues from selling natural gas resources. Short-term maximization of natural gas revenues is at odds with the goal of security of supply, so a trade-off must be found between these goals. Finally, the price for the sale of domestic gas supplies influences the price that consumers (domestic and foreign) pay for gas. Here, too, a trade-off needs to be made, namely between the goal of keeping consumer prices low, the goal of high gas revenues for the State – also from exports – and the goal of stimulating consumers to use energy efficiently.
Summarizing, the following four policy goals can be identified:

- affordable (liberalization and regulation)
- reliable (security of supply, short-term reliability)
- clean (minimization of negative environmental effects)
- optimal utilization of the Dutch natural gas supplies.

Environmental issues, such as climate change and the question whether natural gas resources underneath the Waddenzee national park should be developed, are not discussed here. The focus is on questions concerning the design of the gas sector; environmental issues are a somewhat different matter. Natural gas can, however, be considered to contribute to the transition to a more sustainable society.

Keeping these policy goals in mind, the functions that need to be performed in the gas sector can be analyzed and an understanding can be developed of the available choices concerning the design of the gas sector. Looking back at Figure 3 on page 11, the following groups of related functions can be identified:

- production (resource extraction; from the Groningen field, the small fields and the other fields) and imports,
- transport (and transport network management) and system management,
- flexibility,
- distribution (network management).

The first three of these groups of functions will be discussed in this chapter. For each function, the relation between the different system design options and the policy goals will be indicated. Section A.2 provides an analysis of the production of natural gas and imports. Section A.3 discusses transport and system management. Section A.4 describes the issue of flexibility in supply and demand. Distribution is not discussed here, as it was the subject of Chapter 3.

A.2 Production and imports

Two questions are addressed in this Section. The first question is how the stated policy goals can be achieved in a liberalized market. At issue is the structure of this market: assuming that there is competition, how the market should be designed in order to achieve the policy goals, such as security of supply. The second question is, essentially, whether restructuring is worth the effort. Is there sufficient potential for competition, that is, with enough pressure on market parties to improve their efficiency and to innovate, considering the strongly oligopolistic character of the international gas market?

It is said that unbundling and privatization of the Dutch gas sector would lead to a speedy emptying of especially the Groningen field, which would not contribute to the security of supply. Continuation of the involvement of the State would provide a guarantee for a more
cautious policy. Reality is more complex, however, because draining the Groningen field in a short period of time does not appear economically interesting for anyone, among others due to the limited market for the low-caloric Groningen gas. On the other hand it is possible to envision different regulations and control mechanisms to implement a depletion policy in the different models for restructuring the gas sector. Government could, for instance, put a cap on production volumes or influence the production of the different fields through fiscal measures, as it currently already does.

A related issue is the small fields policy. In order to continue this in a liberalized environment, it is important that the agent that is responsible for its execution is independent and well-regulated, in order to keep him from abusing his market power. Shell en ExxonMobil, the intended heirs of Gasunie for implementing the small fields policy, would obtain access to all kinds of market information, while of course they have strong interests of their own in the gas market. Other oil companies, who are involved as shareholders in the production of a large number of gas fields in the Netherlands, like Total, BP, Wintershall, Gaz de France/Suez and Centrica, will resist even the appearance of a conflict of interest. It is possible to conceive of solutions for this issue, too, in the form of strong regulation of the party that is executes the small fields policy. An option is, for instance, to allow EBN – the current holder of the State share – to play a role. The question remains, however, whether GasTerra is not the most suitable party for this role.

The last policy goal to be discussed here is the optimal utilization of the Dutch natural resources. Different aspects play a role. First, Dutch domestic production needs to be large enough to be able to execute the small fields policy. For diverse reasons, the output of these fields is not flexible. Because GasTerra purchases most of the gas from these fields, the volume of sales of GasTerra need to correspond to the production by the small fields. The second aspect is, of course, that the State’s revenues from gas sales depend upon domestic production. A lower production rate of Dutch gas does not only mean lower revenues for the State, but also for the private gas producing companies in the Netherlands. This will probably reduce their willingness to invest, which could have negative consequences for the future position of the Netherlands in the international gas market. This would not benefit the development of the Netherlands as an international hub.

Almost by definition, newcomers in the gas market mean a reduction of the market share of GasTerra in the Netherlands. If the small fields policy is maintained, this means a reduction of the output of the Groningen field, because the demand for gas is not expected to increase significantly. When the large consumers gained the right to ‘shop’ for their gas in 1998, Gasunie saw part of this market lost to imports from the U.K. A difficulty was that Gasunie – despite its apparently attractive purchasing position – could not underbid these
companies, because its regulation as a near-monopolist required it to apply the offered price reductions to all similar customers. This loss of market share may conflict with the task of purchasing the gas from the small fields.

Yet, later on, GasTerra managed to expand its export, particularly to the UK, via the newly constructed BBL pipeline, so Dutch gas production has not declined. As was mentioned before, the Groningen field can be produced at low cost, so it generates high revenues. As opposed to the small fields, the Groningen field has a progressive revenue arrangement. As a result, the State’s share of the profits may increase to between 70% and 90% of the net revenues, while for the small fields this is limited to 40% to 50%. Therefore decrease of the output of Groningen weighs heavily on the State’s gas revenues.

The last aspect of the presence of gas in the Dutch soil is that this has provided the Netherlands a key role in the gas supply of Northwest Europe. Despite the limited nature of these resources, it may be possible that the Netherlands retains this position for a long time to come and even reinforces it. The small fields policy and the development of gas fields underneath the Waddenzee may delay the exhaustion of the Groningen field, but by and by increasing quantities of natural gas will need to be transported from far away sources to Europe. The Netherlands could import Russian or other gas and export it again at advantageous moments, after having treated and stored it. The use of old gas fields for seasonal storage of gas and the already present infrastructure would support the Dutch hub function. In addition, there appear to be opportunities for the Dutch gas industry to offer other services, such as storage near consumers, stabilizing the composition of the gas and the providing seasonal flexibility and peak supplies. These activities can be developed now on the basis of the current Dutch gas production. If that decreases in the future, the Netherlands may remain able to play the role of gas ‘banker’ by using the available infrastructure and its geographic advantageous location.

Summarizing, it may be said that achieving the stated policy goals (security of supply, small fields policy, maximization of gas revenues, development of the Dutch gas industry) has become more difficult, but not impossible in a liberalized market. The next question for this Section is to which extent we may expect that these reforms actually lead to an increase in competition. For the time being, it is unclear what the future structure of the European gas market will be and how other importing and exporting states will act on that market. As a flexible gas exporter in the geographic center of the European market, the Netherlands is in a special position, which it shares with no other EU member state or gas producing state. Therefore the Netherlands has much to lose from irreversible interventions in the gas system, if later these turn out not to be compatible with the new market conditions.
As was mentioned before, the development of international competition falls short of the expectations. An important reason is the way in which gas producers, wholesale traders and electricity producers, such as Shell, ExxonMobil, BP, Total, Ruhrgas/E.On and Gazprom are reorganizing themselves. In a competitive market environment, horizontal and vertical integration at the European level – important contributors to the development of an oligopoly – are a substitute for the traditional public-private organization of the market. In order to finance the necessary investments in infrastructure, gas producers want to secure their sales by means of long term contracts. Consequently, the majority of the continental market still is supplied in this way. Pressured by the most important export countries (Russia, Algeria and Norway), the European Commission has accepted that these contracts are inevitable. Even the market-oriented British government accepts that these forms of security are a condition for gas imports from Norway (as an alternative to the decreasing domestic production).

In addition, the role of sovereign states in the production and international trade of gas is far from over. In contrast to some member states of the EU, the governments of most gas exporting countries (Norway, Russia, Algeria and other potential suppliers) are not at all convinced of the need to place themselves at a distance from the market. The reasons are the geopolitical and economic national interests (on the supply side of the market), the need for coordination of the development of supply and demand, and the role of governments in the development of the necessary long-distance infrastructure. The increasing geopolitical complexity of the expanding international natural gas system is reason to believe that the role of governments may increase sooner than decrease. Therefore it is an open question to what extent real competition is possible in the different segments of the gas industry, in the Netherlands, in the European gas market and between the suppliers in the Commonwealth of Independent States and North Africa.

Considering the limited degree of competition in the production of gas and the strong involvement of governments, it appears practical that the Dutch government is sufficiently equipped to defend the position of the Netherlands in international negotiations. The question is which structure of the Dutch sector best meets the interests of the State of the Netherlands, consumers and businesses, without reverting to the traditional pattern of over-regulation. The Netherlands has relatively much room for maneuvering in this matter, because the European Gas Directive only concerns the liberalization of trade and end-user supply of gas. Production is left outside the scope of the Directive because of the national sovereignty over natural resources.

The existing public private partnership between the State and Shell and Exxon has prevented sharp controversies between the government and the gas producers. This is probably due to the relatively symmetrical access to information and the mutual involvement in important decisions. In addition, the public private partnership reduces the direct influence of politics to a degree (and therefore also the impact of hypes), as a result of which public policy had a significant degree of continuity. If government had been further removed from the industry, it would have had more difficulty retaining control over the developments. On the other hand, a completely private industry would have been less easily able to use the government for its interests, both domestic and foreign.

One of the main questions with respect to the restructuring of the gas sector is the way in which the State of the Netherlands should remain involved in the production of gas. The main argument for a reduction of its role is that that fits best with the general policy of liberalization. However, the potential for competition in the upstream part of the gas market (production) is limited. There are good reasons for continuing the current public private partnership of the State with Shell and Exxon in one way or another. The depletion policy and the small fields policy are more easily implemented and the growth of the Dutch role as hub will be more easily stimulated. Direct involvement of the governments in other gas producing countries may also be a reason for the Dutch government to remain involved, in order to be able to negotiate on an equal footing.\textsuperscript{89} The government will therefore need to make a clear choice with respect to the degree to which it wishes to remain involved in the sector. It is important to realize that dismantling the public private partnership in the current gas sector of the Netherlands will be an irreversible act.

A.3 Transport network management and system management

Gas Transport Services, GTS, is the responsible party for the monopoly tasks, other than distribution, within the gas sector. This section will discuss the regulation of GTS. With respect to unbundling, it is the question for which networks and activities the manager of the gas transport network should be responsible. Besides the main Dutch gas network and the connections to the gas fields, GTS is also responsible for the export and import pipelines. Although these networks all are part of the total gas supply system, the rationale for their management, their ownership and their regulation differs. For other activities, such as storage, quality conversion and flexibility services, it is the question whether they should be considered as monopoly activities, which should be regulated strictly, or as competitive activities (and for which monopoly regulation is unnecessary). With respect to the ownership issue, the Dutch

\textsuperscript{89} Deze visie is verder onderbouwd en uitgewerkt in een recente studie voor de Europese Commissie (Van der Linde \textit{et al.} 2004).
government decided in October, 2004, to buy the shares of Shell and Exxon in GTS, so that the state will become the sole owner.

The transmission network is the lynchpin of the gas sector. It forms the connection between the Dutch gas fields and the domestic and foreign consumers, while it also transports imported gas to domestic consumers. Four groups of tasks can be identified with respect to transport and system management:

- The network manager needs to perform his typical tasks, such as transport and supporting services.
- In the short term, the system operator (who usually is the same actor as the network management) needs to balance supply and demand.
- In the medium term, the system needs to be developed and expanded in such a way that it continues to have enough capacity to meet supply and demand. The main issues are the adjustment of the infrastructure to the development of new gas fields and the development of new pipelines to foreign fields.
- In the long term the management of the gas transport network influences the role of the Dutch gas infrastructure in the European gas market when the domestic Dutch reserves become exhausted.

In the short term, especially the regulation of the current use of the transmission network is important. Essentially, the view of DTe implies that competitive and monopoly activities are unbundled. These monopoly activities need to be regulated in such a way that the shippers can obtain custom combinations of transport, storage and other services with transparent contracts and for a regulated, cost-based price. This would allow the shippers to save money for their services, since they would only need to pay the cost of essential (regulated) services. Moreover, (the availability of) flexibility in the transport of gas and storage are considered conditions for achieving an efficient gas market in which actual competition is possible between the different producers. The question arises which activities can be competitive and to which extent these activities are independent of each other and can be separated. While it probably is not realistic to duplicate the transport network itself (despite the exception provided by the existence of the ZEBRA pipeline), it is a matter of discussion to which extent gas storage and quality conversion are monopoly activities. An important consideration in this respect is that contracted flexibility in the production and consumption of gas can be considered as alternatives to storage (see below.)

Maximization of access should not have as a consequence that system stability is threatened. Therefore the second important question with respect to network regulation is how the non-competitive activities should be regulated and, especially, how the necessary coordination can be realized. In the short term, the balance between supply and demand needs to be maintained. For this purpose, a system of tariffs
and penalties is needed that stimulates shippers to take care of balancing themselves as much as possible, while enabling different types of transactions on the network. This system is being developed; the commencement of a spot market will be an important step forwards.

In the medium term, it is more difficult to realize the necessary coordination if the transmission network is unbundled. Unbundling the supply system could come at a cost to the efficiency of system management and, eventually, to the security of supply. Especially in a market environment with many short term transactions, it is quite difficult to forecast the demand for capacity in specific parts of the system early enough. Moreover, unbundling leads to two more or less independent organizations with separate (profit) goals. In the long term, it is the question whether the possible gains in efficiency of the new unbundled structure outweigh the disadvantages of less coordination and less information for the purpose of system expansion and adjustment. Market parties, on the other hand, will demand higher revenues as a compensation for increased risk and additional uncertainties.

In order to achieve a position as a regional gas hub, an attractive investment climate needs to be created for the necessary facilities. At issue are mostly commercial activities with a certain market risk that have little to do with Dutch public tasks. An investment regime is required that provides the possibility to incorporate these risks in the prices and access rules, while the risks and investment costs should not be placed upon the shoulder of the Dutch consumers. A possibility is to allow exceptions to the normal system of regulation, such was done for the Balgzand Bacton pipe line (BBL) from the Netherlands to the UK.

A.4 Flexibility in supply and demand

Due to the large variation in the use of gas on a time span of hours, days and seasons, it is necessary to incorporate flexibility in the supply. In addition, flexibility of gas is desired in order to prevent supply disruptions in case of calamities. This flexibility can be created in different ways. Some traditional measures are:

- flexibility in the production of gas;
- temporary storage of gas, short or long term;
- interruptible contracts (for instance with consumers who have alternatives, such as dual firing);
- line-pack: storing gas in the transport network by increasing the pressure.

In a system that has not been liberalized, flexibility simply is part of the (all-in) supply contracts. The costs of these measures are part of the integrated tariffs which are (or used to be) charged to consumers. Before liberalization, especially physical measures were used to
create flexibility. These were aimed at keeping the necessary volume of gas and the necessary transport and distribution capacity available.

Liberalized markets require different solutions, but also offer new opportunities. Not only gas, but also flexibility can be offered as an (independent) service and traded for market value. Prices reflect the degree of scarcity at the moment of supply. This way, they have a function in guiding the volume of demand at different times. This results in a more efficient utilization of the price elasticity of demand. To obtain flexibility, market parties can of course make use of the traditional ‘physical’ sources of flexibility. But new services may develop, such as spot markets, supply of ‘peak gas’ and forward markets. Market parties may also opt for in-house solutions, such as investing in private storage facilities or interruptability (for instance with the aid of consumers who have dual firing possibilities). Another possibility is that large consumers, such as electricity producers, do not consume all the gas that they have contracted, but re-sell part of it (‘gas-power arbitrage’). The development of financial instruments to cover risks may also contribute. These financial instruments provide signals to investors in the different forms of physical capacity.

Liberalization does not only lead to new forms of flexibility, but also makes flexibility an issue of customization. In non-liberalized systems, a certain appointed actor applies flexibility instruments and facilities for all system users together (and therefore the same for all consumers and with socialization of costs), based upon a certain policy. In a liberalized system, on the other hand, consumers may purchase flexibility services à la carte, rather than only as a packaged deal. This means that consumers need to determine what the value of these service is to them: the value of a certain type of flexibility, in a certain quantity and at a specific moment. As a result, the value they attach to their own security of supply becomes a determining factor for the maximum price which these consumers are willing to pay for flexibility.

The above scenario rises or falls with the possibility to offer flexibility as a competitive service. Besides the TSO, the shippers, suppliers and consumers in principle are able to provide certain types of flexibility. The question is whether these parties are able to offer flexibility as a commercial product, or whether it actually is a natural monopoly, for instance due to economies of scale or because GasTerra holds key facilities.