Report 2008-129

Optimal network tariffs and allocation of costs

MEMBER OF THE PÖYRY GROUP

Econ-Report no. 2008-129, Project no. 5Z080079.10 Public
ISSN: 0803-5113, ISBN 978-82-8232-017-7
AJE/CSE/pil, 17 November 2008

Optimal network tariffs and allocation of costs

Commissioned by
the Norwegian Water
Resources and Energy
Directorate

Econ Pöyry AS
P.O.Box 5, 0051 Oslo, Norway. Phone: + 47 45 40 50 00, Fax: + 47 22 42 00 40, http://www.econ.no

Table of Contents:

Executive Summary

Abstract

Pricing of transmission and distribution of electricity should be done according to economic criteria, that is reflect the marginal short-term costs of losses and congestion in the grid. Long-term price signals beyond these short-term signals should reflect the cost of customer-specific investments. As the grid constitutes a natural monopoly, these tariffs will however not be sufficient to cover the total need for income in the grid. Hence, there will be a need for tariffs that provide recovery of residual costs. The allocation of residual costs between network customers should be done in a manner which distorts use of the grid and investments as little as possible. In practice, this means relatively low tariffs for generation and large industrial users, while households, the public sector and small businesses should cover the bulk of the residual costs, although there are many variations within these groups and over time. The Norwegian tariff system incorporates many of the economically correct principles, but it is practically impossible to implement a theoretically perfect system.

Background

Tariffs for transmission and distribution of electricity play an important role as price signal for producers and consumers of electricity both in the short and the long run. For instance, tariffs may influence generation investment decisions, localisation of industry and choice of heating systems. Different tariff systems may have vastly different impacts on investment decisions and the short-run utilisation of the existing grid. Also, the distributional effects between different groups of network customers such as households, industry and power generators will be large, depending on the tariff system.

Problem statement

We address the following questions in this report:

> *What are the economic characteristics of optimal tariffs for transmission and distribution of electricity?*
>
> *How should network costs be allocated between different customer groups?*

We focus on the fundamental economic principles rather than detailed description of possible tariff systems, but illustrate the analysis with specific examples taken from observed practice, particularly from the Norwegian power system. The report has been prepared for the Norwegian Water Resources and Energy Directorate.

Conclusions and recommendations

The electricity grid is a natural monopoly

Practically all end-users of energy choose to be connected to the electricity grid in order to cover their demand for electricity-specific goods and services such as lighting, television and computer equipment. In addition, many end-users rely on electricity for heating purposes in certain countries such as Norway. Generators obviously require a grid for selling their electricity in most cases.

Electricity grids are characterised by high fixed costs and low variable costs. In the short run, only transmission losses vary with the consumption level. Typically, losses make up 10-15 per cent of total grid costs. Even in the longer run, there is no clear relationship between consumption/load and the level of grid costs, although the grid must obviously be designed to cover the peak load and the overall electricity consumption in the system, as well as provide sufficient transmission capacity between regions and countries. To simplify, there are significant economies of scale in the grid, and it is not economically efficient to build parallel electricity networks. The grid thus constitutes a *natural monopoly* in economic terms.

Optimal tariffs should reflect short-term marginal costs…

Economic theory on efficient pricing of natural monopoly services can be used as a basis for designing optimal transmission and distribution tariffs. The main criterion for economic efficiency is that the tariffs for use of the network should ensure that the existing grid is utilised to the maximum, subject to demand and the short-run marginal costs of transmission and distribution. The marginal costs consist of marginal losses at each point in the grid, as well as capacity constraints or congestion. If capacity is constrained, a congestion fee should be used to ration the available capacity (peak-load pricing).

Tariffs based on short-term marginal costs also give long-run investment signals. I.e., high congestion fees and marginal losses in a given point in the grid indicate the value of new network capacity – or local generation. Additional long-run price signals can be given through connection charges or project-specific investment contributions from the network customers (both positive and negative) that reflect the impact on system costs from a new connection at a given point in the grid.

General tariffs per kWh or MW capacity which give price signals beyond marginal costs and congestion, such as distance-related tariffs or capacity subscriptions, are not advisable. These price signals will by their very nature be imprecise and too general, thus increasing the risk of distorting operational decisions and investments.

…but supplementary residual tariffs are necessary to recover the full grid costs

Due to the natural monopoly characteristics of the grid, the *marginal* cost of expanding capacity will usually be lower than the *average* cost in the existing grid or indeed the average cost of capacity expansions. Thus, tariffs which give the correct short-term signals will not be sufficient to cover the full costs of the grid. Neither are connection charges or investment contributions likely to give full cost recovery. Hence, there is a need for tariffs which recover the *residual need for revenue* to the grid companies, i.e. residual tariffs.

The economic criterion for efficient residual tariffs is that they should have as little distortionary effect as possible on the utilisation of the grid in the short run and investment decisions in the long run. The relevant investments comprise all decisions about infrastructure for energy use, from generation to localisation of industrial activity and energy solutions for households. In practice, this is an issue of optimal taxation of grid users, both consumers and generators.

Distribution of residual costs according to the price sensitivity of demand is the best solution

A well-known solution to the cost recovery problem from economic theory, is to use so-called Ramsey pricing. In its original form, this would imply a mark-up on the energy tariff which is differentiated according to the price-sensitivity of the demand for electricity. I.e., the more price-sensitive demand, the lower the tariff per kWh. This will yield the lowest loss of economic welfare:

With an optimal two-part tariff, where a variable tariff based on energy consumed or injected into the grid reflects only marginal losses and capacity constraints (if any), a similar principle to Ramsey pricing may still be used for the design of the fixed part of the tariff. For instance, differentiated fixed charges per point of connection or different load-based or capacity charges per MW may be used to recover the residual need for revenue to the grid. In practice, it means that the greater part of the residual costs should be borne by the network customers with the least price-sensitive demand. Both the short-run and the long-run price sensitivity or elasticity are important in this respect.

Power-intensive industry and generation should pay a smaller share of residual costs

In a market-based power system, the price is determined by the marginal capacity which covers demand and clears the market. The marginal capacity will vary over time, but the marginal cost and hence the price will be influenced by the energy tariff. Any mark-up above the cost of marginal losses and capacity constraints will in fact be distortionary. A tariff on installed or maximum capacity will not necessarily influence the merit order curve as such, but will instead influence the price structure through the investment decisions of generators. In fact, if generators are going to contribute towards recovery of residual grid costs, it is better to use a small mark-up per kWh instead of a capacity charge as it will have a smaller distortionary effect. This is particularly relevant in a hydro-based system such as Norway, where the choice of capacity installed is a very important decision variable.

In the long run, the location of generation investments and choice of technology will depend on the residual tariffs in different regions. In an integrated power market, a high residual tariff burden on generators may therefore cause significant welfare losses as high-cost generation may be chosen instead of low-cost generators in another region, which is particularly important if investments are distorted. Hence, generators should pay only a small share of the residual costs. It is also an argument for harmonisation of residual tariffs for generation in a regional electricity market. However, established power plants with low marginal costs of production and a high profitability (such as hydropower in many instances) may carry a relatively higher tariff cost without significant costs to society. It may be difficult to implement such a distinction in practice, though.

A similar line of argument may be put forward for the established power-intensive industries as well. These industries are subject to international competition, and if the opportunity cost of their production is lower elsewhere, they may choose to relocate their business elsewhere. Again, welfare losses may be the result (due to carbon leakage, costs of finding new employment etc.).

In a Norwegian setting, the petroleum sector represents an interesting source for residual income. The return of the petroleum activities is often well above a normal rate of return, and the sector's share of electricity consumption from the onshore grid is

increasing. However, the level is restricted by the opportunity cost of choosing other energy supply options and the oil price.

From an economic point of view, efficient recovery of network costs entail that households and small businesses should pay the greater share of residual costs. This is very much in line with a tariff system based on a "Ramsey principle". Of course, power-intensive industries and generators should still pay the costs they actually impose on the system which are possible to identify and attribute to a specific customer. However, due to the cost characteristics of the grid, this will not be sufficient to achieve cost recovery. The effects on the income distribution from a "Ramsey principle", although the principle is economically efficient, may not be desirable though.

The Norwegian tariff system is built on economic principles…

The current Norwegian tariff system, which has been developed over time since the electricity market reform introduced in the early nineties, is to a large extent based on the economic principles described above. The variable tariff for generation is calculated as the product of the market price of electricity and the estimated marginal losses from injecting power in each point in the grid, and may be positive or negative (i.e. a payment to the generator) depending on the exact location and power flow. The residual tariff is based on the historic ten-year average generation with a rate of around 0.64 €/MWh. In all, generation covers only 5-6 per cent of the overall network costs on average. All generators are required to cover the costs of generation-specific network assets, but there is a discount on the residual tariff available in certain parts of the transmission grid where it is especially beneficial to build new generation capacity. The same basic setup applies to all types of generation capacity.

The variable tariffs for consumption in the transmission grid likewise reflect marginal losses (again, these may be positive or negative), while the residual tariffs are based on capacity (measured peak load in random reference hours or installed capacity) or, in the distribution grid, a fixed rate per point of connection. In the distribution grid, energy tariffs are typically higher than the marginal losses imply, thus covering residual costs as well as the cost of losses.

The tariffs in the central grid (the highest level of the transmission grid) are adjusted through the so-called k-factor, which rewards locating large consumption units with a high load factor in the same point as generation. In practice, this means that the power-intensive industries pay a fairly small part of the residual tariffs and the overall grid costs. Again, this is very much in line with the Ramsey principle. Consequently, the bulk of the tariff bill – estimated at 90 per cent of the residual costs – is paid by households and small businesses, although they only account for around 60 per cent of total electricity consumption in Norway.

The cost of customer-specific investments, including reinforcements if a customer requires a higher quality or capacity, are to a large extent covered by so-called investment contributions, which apply to both generation and consumption. However, investment contributions in the meshed grid are normally not used. General (shallow) connection charges are also used, but these are not directly related to the investment costs as such, unlike the investment contributions.

Finally, large and long-lasting "structural" bottlenecks in the central transmission grid are handled through zonal pricing. Technically, this results in different area prices in the

wholesale market in the presence of congestion between market areas, but it is still an element of the transmission pricing system.

…although there is always room for improvement

Despite the economic foundations of the Norwegian transmission and distribution tariff system, there are still some shortcomings.

Firstly, the widespread use of energy tariffs for recovering residual income in the distribution grid distorts the price signals both in the short and the long run. The long-run effects are particularly important, as they influence the choice of heating systems and overall capacity for consuming electricity.

Secondly, the long-term price signals are incomplete for two main reasons. One important factor is that investment contributions or project-specific connection charges in the meshed grid are normally not used. While this is obviously challenging in practice, it still remains a shortcoming compared to the ideal system. Also, a theoretically optimal system should include negative investment contributions to reward investments in generation or consumption that reduce the need for grid investments. This is necessary in order to compensate for the fact that the short-term price signals will not give perfect long-term signals. A typical example is large-scale generation investments in an area with a substantial power deficit, which may reverse the sign of the marginal losses in a given point and remove area price differences, thus removing the very price signals designed to reward new generation capacity in areas where it is particularly beneficial. It is however difficult to envisage practical solutions to this issue.

Using the residual tariffs to give price signals where they are absent may seem a possible solution, but it such a strategy will have other distortionary effects which may be significant. This is due to the fact that the perfect price signals are time- and location-specific, while the residual tariffs necessarily entail generalisation and thus imperfection in some sense. We cannot rule out that such tariffs may increase economic welfare; however, the reverse may also happen. Caution should therefore be applied when considering whether the residual tariffs should give price signals.

On the whole, the regulator must strike a balance between the theoretically correct system and a feasible system which can be implemented in practice. As generation and demand changes, the tariff system may also need to undergo alterations. However, economic efficiency in the short and the long run should always be the target for the regulatory authorities and the network companies.

1 Introduction

1.1 Background

Tariffs for transmission and distribution of electricity play an important role as price signal for producers and consumers of electricity both in the short and the long run. For instance, tariffs may influence generation investment decisions, localisation of industry and choice of heating systems. Different tariff systems may have vastly different impacts on investment decisions and the short-run utilisation of the existing grid. Also, the distributional effects between different groups of network customers such as households, industry and power generators will be large, depending on the tariff system.

1.2 Problem statement

We address the following questions in this report:

> *What are the economic characteristics of optimal tariffs for transmission and distribution of electricity?*
>
> *How should network costs be allocated between different customer groups?*

We focus on the fundamental economic principles rather than detailed description of possible tariff systems, but illustrate the analysis with specific examples taken from observed practice, particularly from the Norwegian power system.

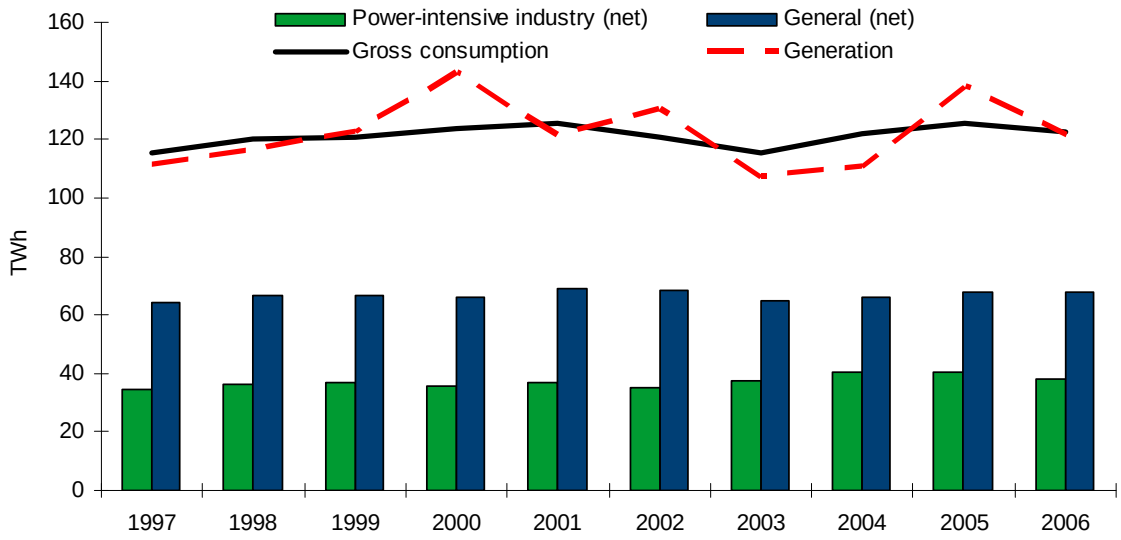
The report has been prepared for the Norwegian Water Resources and Energy Directorate (NVE).[1]

1.3 The Norwegian power system and network structure

The Norwegian power system serves a gross annual consumption of around 125 TWh on average. Over the last decade, more than 99 per cent of the power generation is hydro-based, although some gas (CCGT) and wind power has been built in recent years. Power-intensive industries account for 35-40 TWh of the total consumption, while the petroleum sector is the fastest-growing sector, although from a comparatively low level (around 4 TWh per year). The remaining consumption is characterised by a high share of electricity for heating purposes, which has its historic roots in the availability of inexpensive hydropower and the demographic pattern of Norway, with a large part of the population spread out across the country.

[1] The material in this report is based on previous work by ECON/Econ Pöyry on pricing of transmission and distribution services in Norwegian, in particular ECON (2006) and Econ Pöyry (2008), but has been edited and adapted for a non-Norwegian audience. For a complete set of references to the Norwegian literature, we refer to these reports.

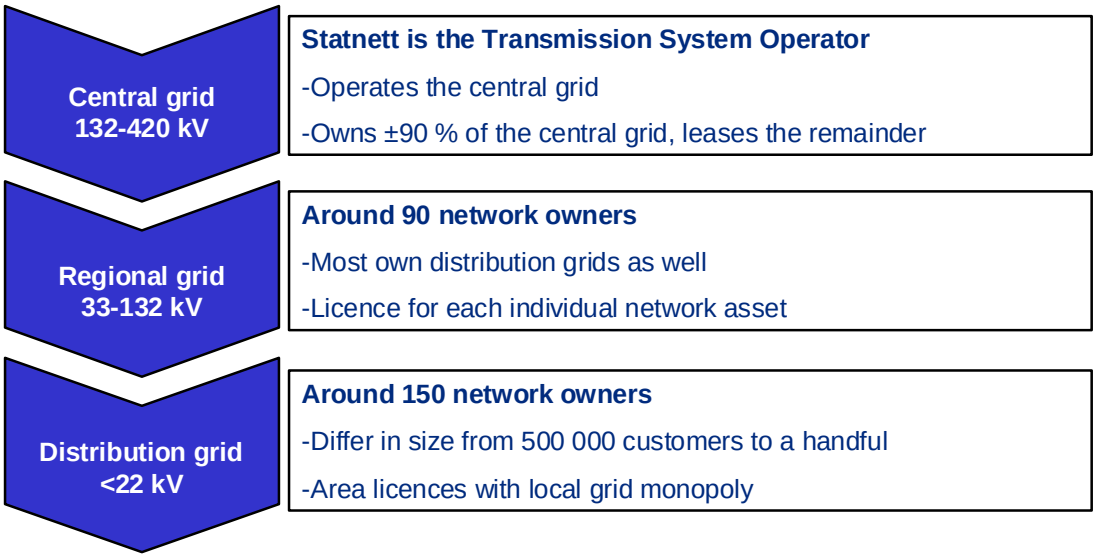
Figure 1.1 Electricity consumption and generation in Norway 1997-2006. TWh/year



Source: Statistics Norway

The figure below shows the Norwegian network structure per 2008. As can be seen, the transmission grid (33 kV-420 kV) is divided into two levels: the regional and the central (or main) grid. The division between these grid levels is set administratively and not according to strict technical criteria. Instead, the division is based on an assessment of the *functional role* of assets in addition to their physical properties such as voltage levels. For instance, there are certain 132 kV assets that are actually covered by an area licence (i.e., the distribution grid), and there are 132 kV assets that are owned by Statnett and included in the central grid. There are also some 300 kV lines included in the regional grid.

Figure 1.2 Overview of the Norwegian electricity network structure



Source: NVE

Customers connected to the central grid pay tariffs to cover the costs of that grid level only. The tariffs in the regional grid reflect the costs at that grid level plus the tariff

costs for the connection to the central grid. Similarly, the distribution grid tariffs cover costs at the lowest grid level plus tariffs for the connection to the higher grid levels. Thus, consumers in the central grid pay a lower tariff per kWh than consumers in the distribution grid. For generators, the tariff bill is at the outset independent of the location in the grid and choice of grid level for connection, although the tariff may vary according to the marginal losses from injecting power into the grid at the chosen location.

Roughly, 2/3 of the total network costs can be attributed to the distribution grid, 1/6 to the regional grid and 1/6 to the central grid. The total costs typically sum up to around €2 billion per year, depending on the interest rate and price of network losses (which is linked to the spot market price).

Historically, almost all generators have been connected to the regional or central transmission grid, but an increasing amount of small-scale hydro is being connected to the distribution grid. Around 60 per cent of the overall consumption is in the distribution grid, while power-intensive industries and certain other large customers are connected to the transmission grid levels (regional and central).

The network companies' maximum revenues are regulated by NVE through an incentive regulation system which combines the cost base of each network company and a cost norm reflecting efficiency scores in NVE's benchmarking model. The network companies' maximum revenues and the rules for setting the tariffs for different customer groups are thus separated in the regulation.

1.4 About this report

The report is structured as follows:

- In chapter 2, we describe the fundamental economic characteristics of transmission and distribution of electricity, both with regard to the demand side and the cost structure of electricity networks.
- In chapter 3, we discuss optimal pricing of network services on the basis of the demand and cost structure as described in chapter 2 and economic theory.
- In chapter 4, we describe the demand for transmission and distribution of electricity for different groups of network customers, both generators and end-users, with an emphasis on the Norwegian power system.
- In chapter 5, we describe the current Norwegian tariff scheme and discuss the strengths and weaknesses of the system in light of the findings from the previous chapters.

2 The economics of transmission and distribution

The subject of this chapter is the demand for transmission and connection to the electricity grid, as well as the cost structure of the electricity grid, i.e. the "supply" of transmission.[2] Understanding the economics of demand and supply of transmission is necessary in order to discuss optimal pricing of network services, which is the topic of the next chapter.

We will first look into the demand for transmission. Thereafter we discuss the cost structure based on national figures from NVE for the Norwegian transmission and distribution sector and some specific examples of grid costs.

2.1 The demand for network connection and use of the grid

2.1.1 End-users

The end-users' demand for connection to the electricity grid is derived from the demand for services that use electricity as an input:

- Some services – such as computer equipment and infrastructure for different media and communication purposes – are electricity-specific in the sense that only electricity can cover the demand of the consumers.
- Other services – for instance heating – can also be covered by other energy carriers, such as district heating or natural gas.
- Several industries use electricity to operate production processes, whether petroleum refineries, aluminium plants, or smaller businesses such as dairies or print offices. This will also be electricity-specific consumption.

We consider the typical demand in the household sector first. If the energy demand is completely dependent on connection to the grid, the willingness to pay will be higher than if the dependency of grid connection is lower. There will of course also be other factors that determine the willingness to pay, such as the possibility of consumption at specific hours, at specific places and with a certain quality and voltage level. The fact that there is a substantial electricity-specific consumption with very high utility per kWh means that the consumer surplus of connection is high. Therefore almost all consumers will want to be connected to the electricity grid. In principle, the electricity-specific consumption can be covered without grid connection, but in normal cases the costs will then be far higher.[3] Consumers and power producers will also demand reinforcements and new lines as the consumption and production capacity increase. Over time it will also be necessary to replace installations, and it can also be desirable to upgrade the grid solutions.

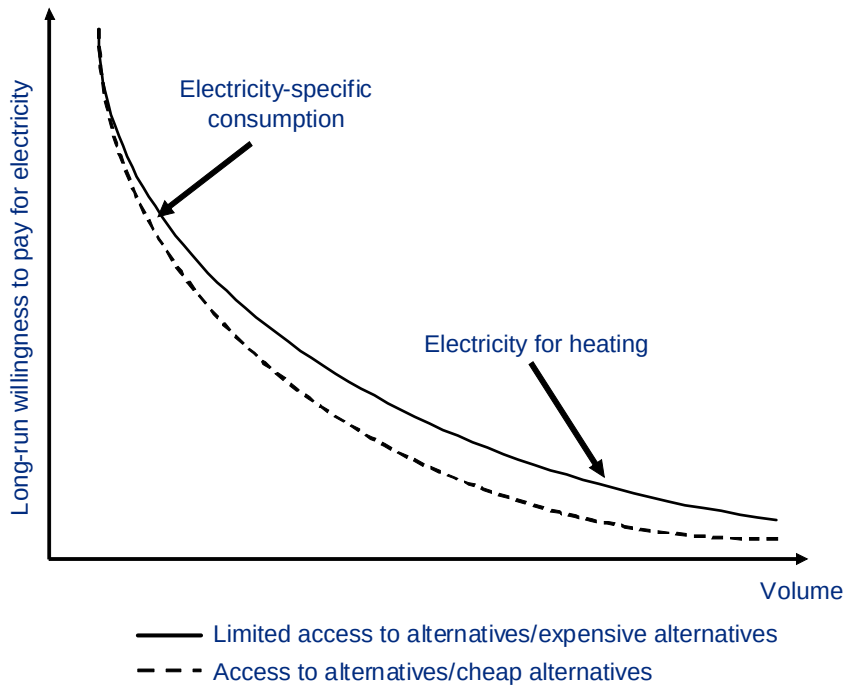
---

[2] In this report, we use the term "transmission" synonymously with transmission and distribution unless otherwise is specified

[3] A possible exception can be for example the use of solar cells to supply cabins located far from the existing grid. Local power generation at the end-user can generally serve as a replacement for grid connection, but often such end-users will want to be connected to the grid to have the possibility to sell excess generation and have access to a backup solution.

The figure below illustrates the possible shape of the long-term demand curve for electricity, which is based on the customers' willingness to pay for being connected to the grid. The willingness to pay for electricity is measured at the second axis, while the demand is measured at the first axis. The willingness to pay for the electricity specific co-sumption is very high and hardly sensitive to changes in the price (the steep part of the curve). The demand for electricity for heating is on the other hand sensitive to changes in the price (the flat part of the curve). In the figure we show two possible demand scenarios. The continuous line reflects that there is limited access to alternatives to electricity for heating (such as remote heating), or that the costs of the alternatives are high. The dashed line shows a situation where there is access to alternatives at relatively low prices. The willingness to pay for the electricity-specific demand is however the same irrespective of access to alternatives.

Figure 2.1 The long-term demand for electricity from end-users



Thus far, we have discussed the willingness to pay for *connection* to the grid. The willingness to pay for expansion of capacity or increased quality of supply will follow a similar path. Here also, grid customers with access to (cheap) alternatives will have a lower willingness to pay than customers without such access. Generally the willingness to pay for electricity grid investments will depend on the following:

- The costs of supplying the demanded energy through local energy production without grid connection (electricity-specific consumption) and/or other energy carriers (thermal consumption)
- Operation and maintenance costs
- The costs of transmission losses in the grid
- The costs of outages and variations in voltage quality
- Bottleneck or congestion costs

The demand for use of the existing grid will follow a roughly similar pattern to the long-run demand. In the short run, customers with alternative heating sources installed

will have a more price-sensitive demand than customers without such alternatives. The most flexible group will be those who have both electric heating and access to one or more alternatives. These will be able to switch between energy carriers when (the relative) energy prices vary. Customers with only alternative heating will very likely have the least price-sensitive demand. A place between these two extreme points we will find customers with only electric heating. For these it is possible to respond to the power price by reducing indoor temperature, which gives some degree of flexibility. For customers with only alternative heating a reduction in the electricity-specific consumption will be the only possibility, but this part of the demand is generally not very price-sensitive.

For end-users in the industrial sector, the willingness to pay for *connection* to the grid is dependent on the value of economic activities which use electricity as an input, i.e. the market price of the products minus the non-grid costs. For industries subject to international competition, the opportunity cost of producing their goods and services in other locations will determine the willingness to pay for connection – or indeed *remaining* connected – to the grid in a given country.[4] As for the demand for *use* of the existing grid, this depends on the short-term marginal costs of using the grid compared to the marginal short-term value of the industrial production. This varies among industries, and may for instance depend on the costs of closing down production at short notice, contractual obligations to deliver goods and other factors.

2.1.2 Generators

The willingness to pay for grid connection from a generator's viewpoint depends on the value of the electricity injected into the grid. This will depend on the expected power price minus the costs – capital costs, operational costs, fuel costs if any, taxes directly related to the generation (i.e., taxes that are not based on profits, but instead per MWh generated or similar). The willingness to pay for grid connection will thus be limited by the expected power price minus the expected non-grid costs. This margin will have to cover all expected network costs, both connection charges and tariffs for using the grid.

Regarding use of the grid for existing power plants, the demand for transmission depends on the market price of electricity compared to the short-run marginal costs of generation, including any network tariffs that are paid per MWh fed into the grid. A tariff on maximum generation in a given hour may also influence the marginal cost. If for instance marginal costs plus network tariffs per MWh exceed the market price of electricity, the generator will choose not to produce.

2.2 Cost structure

2.2.1 The macro picture

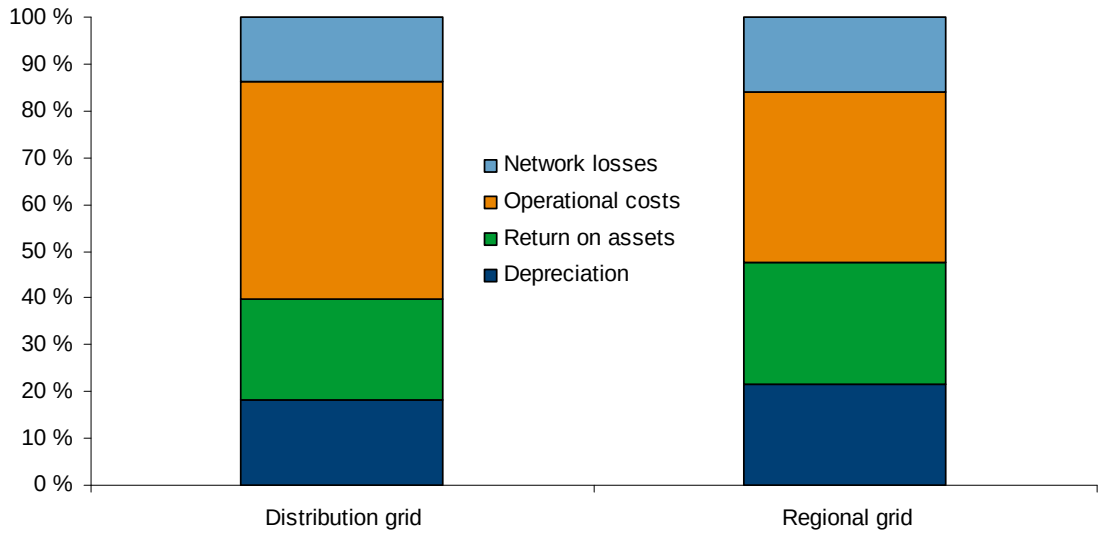
In the distribution grid, the capital costs given by depreciation and return on capital make up roughly a little less than half of the total grid costs. On higher grid levels the capital cost share is significantly higher. The variable costs consist mainly of transmission and distribution losses, which normally constitute 10-15 per cent of the total costs (depending on the power price, that is, the price of the energy that is needed

[4] The same argument can be made regarding choice of location within a country.

to cover the losses). Operation and maintenance costs are generally independent of the consumption level in the short run. In the long run, these costs will also vary with the level of consumption, but the relation between energy consumption and costs to operation and maintenance is not linear.

In the figure below, we show the cost structure in the Norwegian distribution grid and regional transmission grid in 2006 based on figures from the network companies financial and technical reporting to NVE. The return on capital is based on the book value of network assets and a nominal interest rate of 8 per cent. The cost of losses varies with the market price of electricity. The losses in the regional grid are lower measured as the share of energy delivered because of the higher voltage level, but the *monetary* value of the losses constitute a higher share of the total costs in the regional grid.

Figure 2.2 Cost structure of Norwegian electricity networks



Source: NVE, Econ Pöyry

In addition, limited capacity in the grid constitutes an economic cost if the demand for transmission is greater than the available capacity. In an economic sense, this is a short-run marginal cost along with the marginal losses.

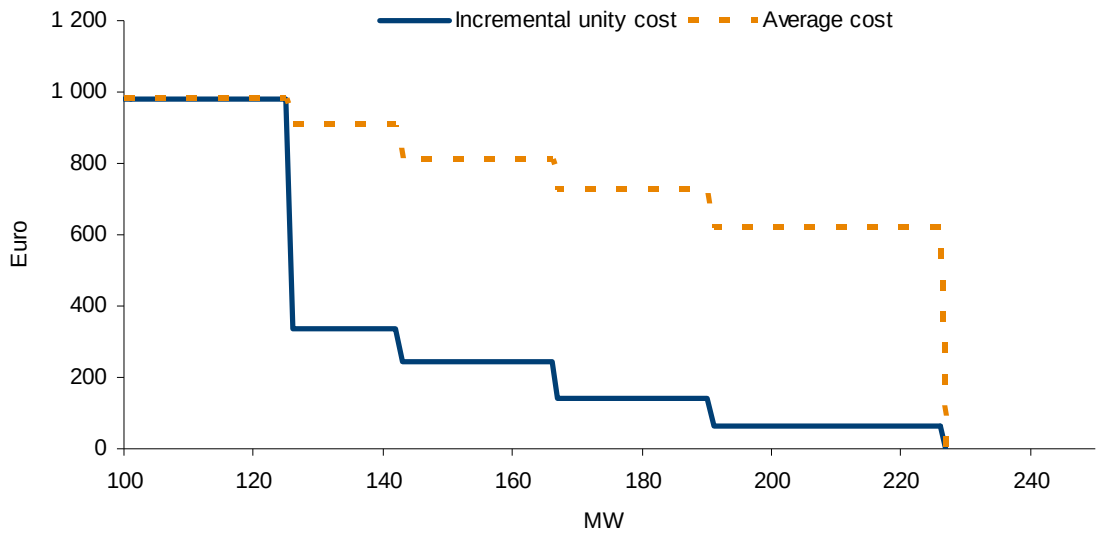
2.2.2 The micro picture

In this section, we look at a specific example of how the cost structure in the grid may be at the level of a single installation, using an investment in a 132 kV line as a starting point (based on data from SINTEF Energy Research). There are five different alternatives for the cross-section of the line with different thermal transmission capacity, where the total investment cost is an increasing function of the cross-section. The most expensive solution costs about 14 per cent more than the cheapest. The largest cross-section gives about 1.8 times as large maximum transmission capacity as the smallest at the voltage level of 132 kV.

In the figure below, we have illustrated the incremental unit cost or marginal cost of increasing the cross-section and thus the capacity. The first part of the curve shows the average cost in €/MW (for one km of line length) of building the line with the lowest

transmission capacity. The next part shows the average cost per MW capacity of building a cross-section with extra capacity. It costs about € 1000/MW if we choose the smallest capacity. If we choose the second smallest cross-section, the extra capacity (above the capacity given by the smallest cross section) costs about 340 NOK per MW. It is also interesting to look at the relationship between the incremental unit cost and the *average* cost at the different capacities. The average cost is shown by the dashed line in the figure. We see that the average cost is higher for all cross-sections except the smallest – then the two costs are by definition the same.

Figure 2.3 Incremental unity cost vs. average cost of building a transmission line. Euro/MW



Source: SINTEF Energy Research, Econ Pöyry

Similar economies of scale will also apply to other kinds of grid installations apart from lines such as cables and transformers. For investments in meters and IT equipment, the cost structure will be different. Characteristics such as lumpy investments and falling marginal costs are likely to be significantly less important than for other grid installations (with exception of the establishment itself of new IT infrastructure or similar investments of a one-off nature). Such investments however make up a rather modest part of the total grid investments compared to lines, cables and transformers.

Another element is that the example above shows the cost structure for a separate grid installation. If we instead it consider a complete network structure, the picture may be different. For example, choosing high capacity on a line may lead to different requirements for transformer capacity etc, than a line with a smaller capacity would do. However, significant economies of scale must be expected in such cases also.[5]

The cost characteristics described above mean that the grid is usually considered to be a *natural monopoly* in the economic sense. A natural monopoly is characterised by a sub-additive cost function. To simplify, a sub-additive cost function is characterised by falling average costs per produced unit.

[5] ECON (2006) contains an analysis of an example of network investments in connection with a new housing development where a more complete set of network installations and reinforcements is considered, which supports the argument above.

2.3 The economics of transmission and distribution – some observations

Above, we have studied the demand and supply of transmission and distribution of electricity in detail, with the aim of uncovering the fundamental economic characteristics. These characteristics have to be taken into account in order to design economically efficient transmission tariffs, which we will turn our attention to in the next chapter. The most important factors to consider are the following:

- Most end-users have a very high willingness to pay for connection to the grid in order to use electricity for specific purposes. However, both the long-term willingness to pay for connection and use of the grid will vary between network users. In particular, industry and generators will only be willing to pay for connection and use to a certain degree, depending on the
- The marginal cost of using the grid, i.e. the cost of consuming or injecting an additional MWh, is limited to the cost of marginal losses in transmission and distribution. In addition, capacity may be constrained in the short run.
- The fixed costs of the grid are high, and may constitute as much as 85-90 per cent of total costs in the short run. Electricity networks are natural monopolies, exhibiting large economies of scale. The marginal cost of expanding capacity will often be small, and it will always be efficient to expand the existing grid rather than building an entirely new one (which however does not preclude that building local generation capacity or reducing demand may be a better option than expanding grid capacity).

3 Optimal network tariffs and allocation of costs

In this chapter, we describe the criteria for economically efficient pricing of transmission and distribution of electricity, under the assumption that the electricity grid constitutes a natural monopoly. Firstly, we discuss how the network tariffs can be designed to reflect the short term marginal costs of the grid and contribute towards optimal *use* of the grid. We then consider how long-term price signals can be given through the tariffs and influence the decision to *connect* to the grid, before discussing how the total grid costs should be recovered in an optimal manner. The method for recovery of grid costs should be evaluated both from the impact on the use of the grid and the choice of connection. In this context, we will discuss both principles for allocation between different groups of network users, including generation, and whether it matters if we look at a closed economy or an open economy where production and consumption of power can be moved between countries in the short and long term. Lastly, we review some results from earlier research regarding the design of residual tariffs.

A general analysis of economically efficient pricing of power, including transmission,[6] can be found in Schweppe et al. (1988). Brunekreeft et al. (2005) gives an overview of the newer discussion on the area. See also Laffont and Tirole (1993) and Joskow (2005) for historic overviews of the development of theory on pricing of monopoly services.

3.1 Optimal prices in the short run

In chapter 2, we gave several examples of the cost function based on macro and micro data from actual network companies and examples of investments, concluding that the electricity grid constitutes a natural monopoly. A very important insight from economic theory is that transmission pricing based on marginal grid costs will yield the optimal grid utilisation, even when the grid is a natural monopoly. The optimal short-run prices give the correct signals for *use* of the grid. When capacity is given (as it is in the short term), it is desirable that this capacity is utilised to the fullest. This happens when the price is set to the short-term marginal cost, which we will explain in the following. Optimal short-run prices will also give information on the long-run cost of *connecting* to the grid. Pricing according to marginal short-run costs will however not be sufficient to cover the total income need of the grid, however, and we return to the subject of optimal cost recovery below.

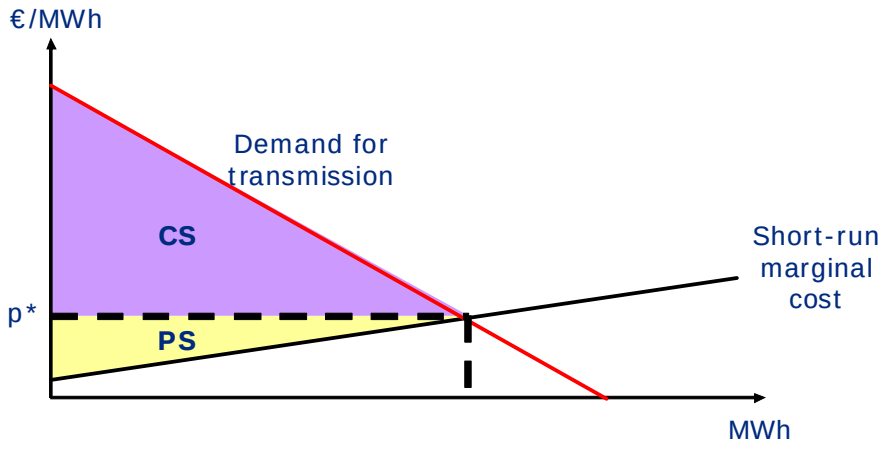
3.1.1 Optimal prices when the grid capacity is not constrained

In the figure below, we show the optimal short-term utilisation of the grid and thereby the optimal tariff. From economic theory, it is an insight that fixed costs do not affect the optimal short-term utilisation, but the variable costs do. This is exactly because they are fixed and need to be taken as given, while the variable costs will be affected by the consumption level. Fixed costs, both capital costs and costs of operation and maintenance, do not vary with the consumption level in the immediate term. The short-term variable costs consists as we have seen in chapter 2 mainly only of transmission

[6] We use the term “transmission pricing” as a common term for pricing of both transmission and distribution of electricity.

and distribution losses. The losses are approximately quadratic as a function of energy delivered, that is, we have a linearly increasing short term marginal cost. For now we will also disregard shortage of transmission capacity. The optimal price is then the price that equates demand for transmission and the short term marginal cost (p^* in figure). The economic surplus of the transmission will then be the largest possible (the total of the producer's and consumer's surplus, respectively *PS* and *CS* in the figure).

Figure 3.1 Optimal network tariff in the absence of congestion



According to the physics of the electricity grid, the marginal cost of injecting or consuming power in a given point may actually be negative. For instance, increased load in an area with a surplus of power may reduce total losses in the grid, i.e. the marginal losses are negative. Thus, the transmission tariff per MWh of use should also be negative. A similar argument may be made for generation in a deficit area.

3.1.2 Optimal prices when the grid capacity is scarce

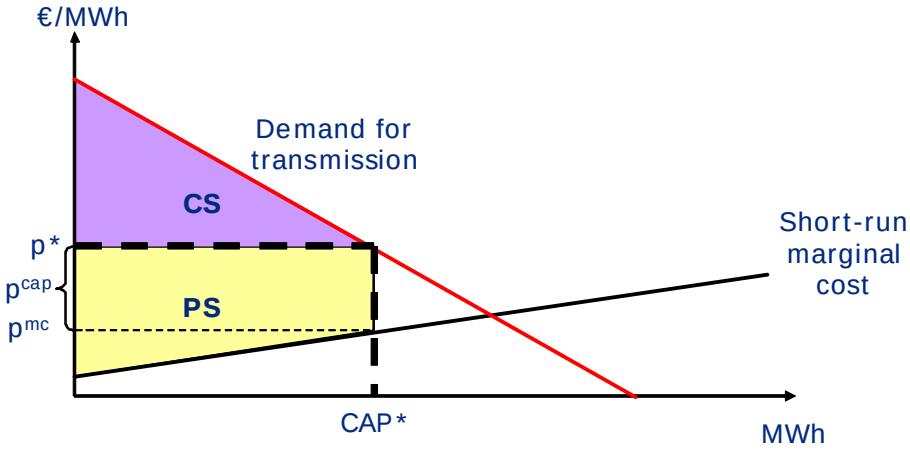
If the available grid capacity is fully utilised, i.e. the intersection between price and marginal cost is to the right of the available capacity in figure 3.1 above, it is optimal to introduce a tariff to ration the capacity. This is in effect what the price areas at the Nordic power exchange Nord Pool do, where the price areas reflect bottlenecks in the Nordic transmission system.[7]

In the figure below we have introduced a capacity restriction *CAP** that shows the maximum possible transmission within a period (for example an hour). We then see that the optimal price p^* can be split into one part that reflects the short term marginal cost (p^{mc}) and a capacity part (p^{cap}) which ensures that the demand for transmission does not exceed the available capacity. This again ensures that the economic surplus of the power transmission is the greatest possible, given the available capacity. This is because the price mechanism takes advantage of differences in the consumers' willingness to pay to allocate the capacity (as opposed to rationing in accordance with administrative

[7] Congestion between the Nordic countries and between Norwegian regions is managed through a system of area prices (zonal pricing), using spot price differentials to reduce flows across congested links to the maximum available capacity. The price will be higher in the deficit area (with full imports into the area) and lower in the surplus area (with full exports out of the area).

procedures, which do not necessarily give conformity between willingness to pay and access to the grid).[8]

Figure 3.2 Optimal network tariffs when capacity is congested



3.2 Optimal long-run prices

Above we looked at the stylised situation where only the short-term costs and capacity constraints are considered. However, in the long run, the grid company also has fixed costs that they need to recover in order to avoid bankruptcy. In the long run, the average costs of the electricity grid are falling due to the natural monopoly characteristics. Therefore, the optimal short-term prices will not cover all the costs.

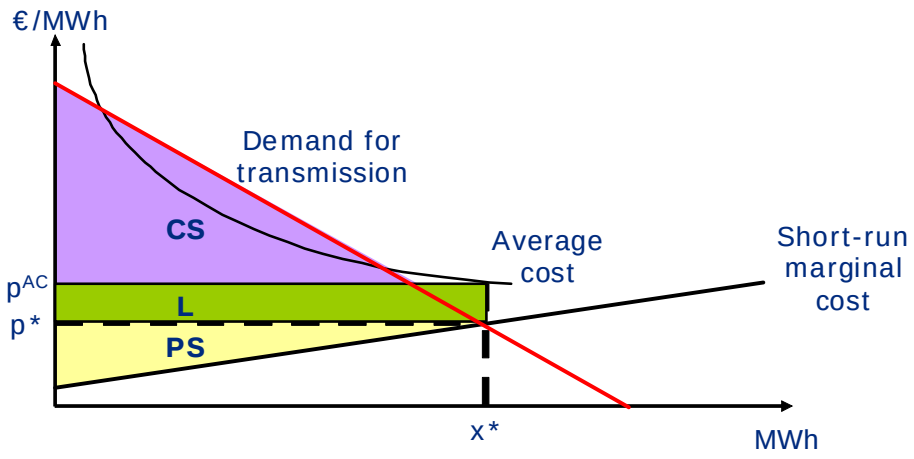
3.2.1 Optimal tariffs vs. tariffs which give full cost recovery

In the figure below, we show the optimal price according to economic criteria compared to the price which gives full cost recovery, if we include fixed costs. The short term marginal cost is still rising, but the long term average cost, where also the fixed costs are included, is falling as shown in the figure. We have disregarded capacity constraints and subsequent capacity fees in this specific case. We see directly that a network tariff equal to the marginal cost (p^*) gives a financial loss equal to the area marked L (limited by p^{AC}, p^*, x^* and the second axis). Total income to the grid company will be the area p^*x^*, while the costs will be equal to the area $p^{AC}x^*$. This tariff will however give the optimal utilisation of the grid provided that the total surplus of both producers and consumers exceeds the financial loss to the grid company (so that it is actually beneficial to have an electricity grid, i.e. the benefits outweigh the costs). This is the case in our example. Price equal to average cost (p^{AC}) leads to reduced consumption compared to the optimal level.[9]

[8] Pricing of transmission according to short-term marginal costs is closely linked to the principle of nodal pricing introduced in Schweppe et al. (1988).

[9] As can be seen from the figure, the price will actually have to be set above p^{AC} if all costs are to be recovered. If the price is set to p^{AC}, which gives a consumption level lower than x^*, the fixed cost per kWh increases. Hence, the network company will still run a financial loss at the price p^{AC}. The intersection between the average cost curve and the demand curve determines the long-run price from a purely profit-maximising point of view.

Figure 3.3 Optimal network tariffs vs. tariffs that give full recovery of costs



We also see that a tariff which reflects marginal losses will contribute towards cost recovery. That is because the marginal losses are an increasing linear function of the electricity consumption. Hence, the revenues from the tariff (the area p^*x^*) will be higher than the total cost of losses (the area under the marginal cost curve). This contribution will however in practice not be sufficient to cover the total costs of the grid, as illustrated by the area *L* in the figure.

The fixed costs create a residual income need *L* in excess of the income provided by tariffs that reflect short term marginal costs. The question is then how this residual income need can be covered. Capacity fees can undoubtedly contribute to this, but will hardly be sufficient to cover the total grid costs. Consider the following illustration:

- Assume that the total costs in the distribution grid are about 1.2 billion euro including return on capital (given an interest rate of 8 per cent).[10]
- The total consumption in the distribution grid is about 70 TWh. Assume for simplicity that the marginal loss rate is 10 per cent for all consumption points and that the power price is 50 €/MWh. The total income from the marginal losses is then given by 70 TWh x 10 per cent x 50 €/MWh, which sums up to €350 million. As the marginal losses are higher that the average losses, the revenue from this tariff will be higher than the cost of losses.
- The residual income need is then €850 million. The capacity tariff must then be around 12 €/MWh on average to cover the total grid costs. This must be regarded as a high average level – and it requires that the capacity is scarce in the whole distribution grid.

In practice, capacity tariffs are not used in the distribution or regional grids, but in the Nordic central grids, there are de facto capacity tariffs through differences in area prices on Nord Pool in the case of congestion. These tariffs gives rise to bottleneck revenues that are divided between the Nordic system operators. These revenues are set off against

[10] This is approximately equal to the cost base for the period 2004-2006 according to NVE's latest analyses of the efficiency in the distribution grid.

the revenue cap of Statnett and make it possible to lower other residual tariffs. These bottleneck revenues have however never been sufficient to cover Statnett's total costs.[11]

3.2.2 Supplementary price signals

Above, we considered tariffs that reflect the short-term marginal costs of the grid including capacity constraints. Theoretically one can also imagine tariffs that reflect the likelihood of outages, which is an increasing function of the load in the grid. Again it is not likely that full cost coverage will be achieved if the tariff levels are to reflect the expected economic costs of outages. The risk of outages is higher in grids with limited capacity (in addition, the consequences of an error will also be larger if the consumption is already high), so the tariffs will in practice have similarities to capacity- or load-based tariffs reflecting capacity constraints.

NVE have proposed to introduce mandatory automatic metering for all end-user customers in the grid by 2013. This gives increased practical possibilities for continuous pricing of not just electricity consumption (i.e. continuous signals from the spot market), but also network tariffs based on the capacity situation in the grid together with marginal losses. Also, the possibility to receive income from different tariffs based on the use of the grid will increase. As seen from the example above, such tariffs will nevertheless still leave a residual income need.

3.2.3 The impact of lumpy investments

The optimal short-run prices for using the grid are *necessary*, but they will not be sufficient as investment signals. A particularly important factor is that the expansion of capacity typically has to happen in greater steps, not in incremental small steps, as we showed in chapter 2 (ref. Joskow and Tirole, 2005, for a discussion of the implications of so-called lumpy investments). When the capacity is to be expanded, it can be profitable to build more capacity than the marginal demand for transmission capacity and the price signals would imply. This is because some network components only exist in given sizes, and partly that planning cost, construction work etc. exhibit large economies of scale. Thus, when an investment is made, the cost of building a large capacity at once will tend to be lower than building a small capacity first and then add extra capacity later, as long as there is a certain probability that it will be necessary to expand at a later date anyway. The example of a transmission line investment from chapter 2 may again serve to illustrate the point.

Also, lumpy investments make it difficult to finance grid investments through tariffs that give optimal short-term price signals. The reason is that capacity increases often happens in such a way that the price signals that are present ex-ante (for example higher area prices in a deficit area), disappear after a grid investment is undertaken (for example a new transmission line to the deficit area). The investment is all the same economically beneficial to society if the willingness to pay is higher than the cost. If it had been possible to expand capacity in infinitely small steps, it would have been

[11] Statnett handles bottlenecks inside price areas with the help of so-called special regulation or counter-trade. This arrangement means that Statnett pays generators behind bottlenecks to increase or decrease generation in order to balance demand and supply given the available transmission capacity. This is a financial cost for Statnett in contrary to the capacity tariffs from the spot market. Statnett could potentially have larger bottleneck revenues if all bottlenecks had been handled through price areas or node prices. However, this would not be sufficient to cover the residual revenue need of the central grid.

possible to expand until the marginal willingness to pay was exactly equal to the marginal cost of expansion.

Under these circumstances, it may be tempting to use tariffs which exaggerate the price signals from the short-run tariffs, or even include tariffs which give general price signals which are more or less in accordance with the underlying long-run marginal costs.[12] However, such tariffs will by the nature of their generality be imprecise and give the wrong signals in many instances. The distortionary effects may be severe both in the short and the long run, and one should therefore be careful about introducing extra price signals without a detailed analysis of the consequences.

The implication of this is that it is impossible to price transmission and distribution, i.e. use of the grid, economically correct, and at the same time recover all the costs of the grid. Still, tariffs that give correct short-term price signals are necessary in order to give an optimal utilisation of the existing grid and an optimal dispatch or power flow.[13] These tariffs are also important as long-term investment signals, although not sufficient.

3.3 Utility of investments as the ideal model for cost recovery

Given that the optimal tariffs for use of the grid are insufficient to cover the total grid costs, the residual income need must be covered otherwise. A theoretically ideal model for coverage of the residual income need can be based on widespread use of *connection charges* or *investment contributions*[14] (positive and negative) that allocate the grid costs on the basis of the benefits which producers and consumers gain from the grid investments. This could cover a very high share of the total grid costs. In addition, there are some genuinely joint grid functions, for instance related to system operation. These are costs, which will not be possible to allocate between the parties even in a perfect theoretical model, but they will add up to a relatively limited amount of the total costs.

In the figure below we show the principles of the theoretically ideal model. The benefits of the investments given by reduced outages, losses and congestion, and the benefits of connecting new customers to the grid, is shown to the left. These benefits can however not be realised without incurring a cost of investment (the figure shows a marginally profitable investment with a benefit that is exactly equal to the costs). We have included environmental costs caused by the grid investment as part of the economic costs. In real life, different kinds of grid investments will have different environmental consequences, and to some degree there may be a need to balance cost for operation, maintenance and investments with the environmental costs. An overhead line may for example have negative effects on the aesthetic environment, while a cable with lower environmental costs will be more expensive to build and operate. Furthest to the right we illustrate how the costs of the investment are allocated to the consumers and generators through changes in tariffs and positive and negative connection charges, which also compensate for the effects of the lumpy investments as we discussed above.[15] If the benefit for each

---

[12] For instance distance-related tariffs, capacity subscription schemes and similar arrangements.

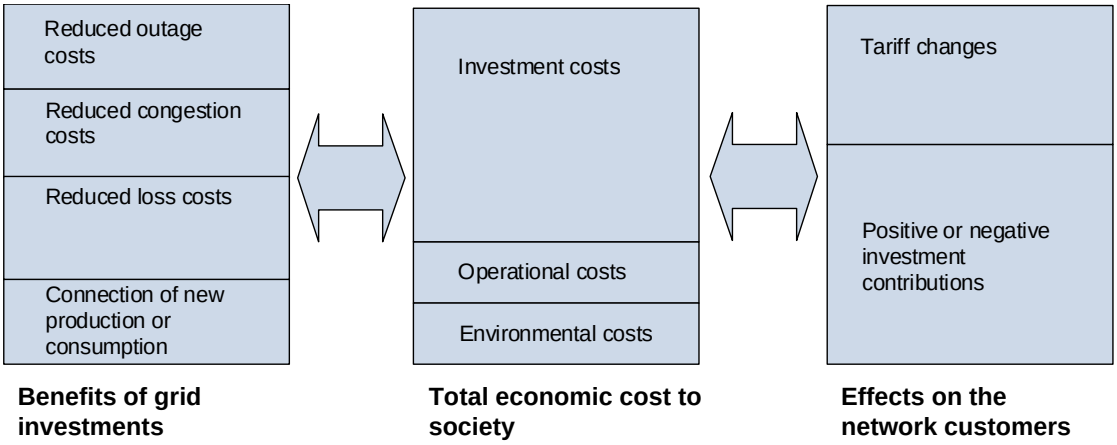
[13] Schweppe et al. (1988) is again a classic reference.

[14] Investment contributions is the term used in the Norwegian tariff system for describing connection charges based on the cost of a specific grid investment rather than a general connection charge. See chapter 5.

[15] Consider for example a new generator in an area with an energy deficit leads to lower network losses. Assume further that an initial negative tariff for injecting in the area due to marginal losses falls to zero, i.e. no extra

customer in the grid exactly matches the changes in the costs, all customers will at all times get perfect short- and long-term price signals, and the income need of the grid will also be covered, except for the costs of the genuinely joint grid functions.

Figure 3.4 Benefits, economic costs and effects on the network customers from investments



In practice, such a benefit-based model for allocation of the grid costs will not be possible to implement, for a number of reasons. For instance, there is no single criterion for the dimensioning of the electricity grid. The choice of grid solution and belonging costs is a function of expected energy consumption, maximum capacity required, customer requirements regarding quality of supply, component standards and different governmental requirements related to security of supply, HES etc. Because of this there is no well-defined relationship between the benefit of each specific grid customer and the cost level unless special circumstances are present (for example in the case of customer-specific radial lines). Furthermore, it will sometimes be desirable to build extra capacity to prepare for later investments in production and consumption, which further complicates the picture (the benefits of the investments fall to customers who do not necessarily exist today). Security of supply in the grid as a whole also has character of being a collective good to a large extent. This means that the supply security of one single grid customer does not affect the security of supply for other customers.

In addition, there will be very big information challenges related to both methodology and collecting the necessary data, and there will be a significant risk of different kinds of strategic behaviour from both producers and consumers. The model also requires a complete evaluation of benefits for all grid investments, which in practice would be too complicated.

3.4 Tariffs that cover the residual income need

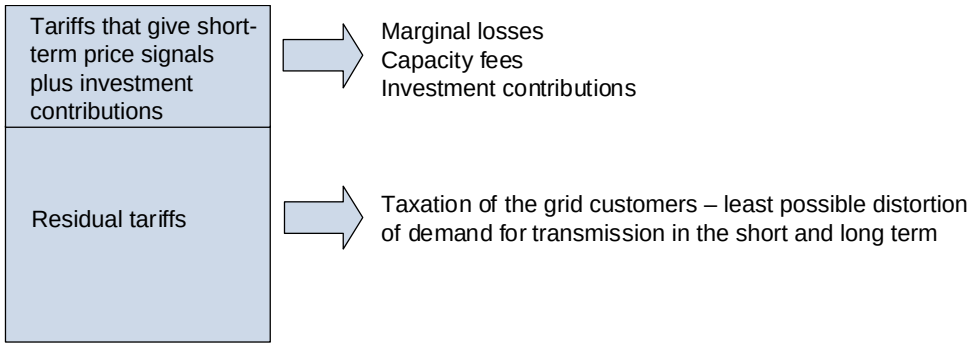
Given that the ideal model is not possible to implement in practice, the residual income need must be covered in other ways. Principally, it is possible to use general taxes to cover the residual income need. In practice, it is common to recover the costs through payments from the grid users, as financing through the tax system generally has

income to the generator from the tariff. This is an economic benefit to society, but the generator is not rewarded for the cost reduction in the grid. Instead, a negative investment contribution based on the impact on losses may be used.

distortionary effects on the labour supply and investments. User payment has also been the norm in the power sector. Historically there has nevertheless been some financing of grid investments through the government budgets by subsidies for the building of grid installations in certain areas. There is also an arrangement for governmental grants to harmonise tariffs between distribution grids, which gives reduced the tariffs for grid customers in areas with particularly high grid costs without reducing the income to the grid company.[16] This scheme has been introduced for reasons of wealth distribution.

The size of the income to be collected through the residual tariffs depends on the total cost level and the income from the tariffs which give optimal short- and long-run price signals according to the criteria we described above. The income from these tariffs will be a function of the optimal rates, which can vary significantly from year to year depending on the capacity situation and the losses in the grid.[17] The figure below illustrates the relationships between the different tariff elements and how they are set.

Figure 3.5 Price signals vs. residual tariffs



3.4.1 General criteria for coverage of the residual income need

The question is then how different kinds of user payments can be used to cover the residual income need in an economically efficient manner. The following economic criteria should be used for the evaluation (obviously, the tariffs must also be possible to implement in practice at an acceptable administrative cost):

- Optimal utilisation of the grid
- Correct investments in grid, production, consumption and alternatives to electricity

The economic literature on coverage of the residual income need in regulated natural monopolies concentrates on two main types of pricing:[18]

---

[16] The scheme has a limited scope though, as it has comprised less than 0.5 per cent of the total network revenues in recent years.

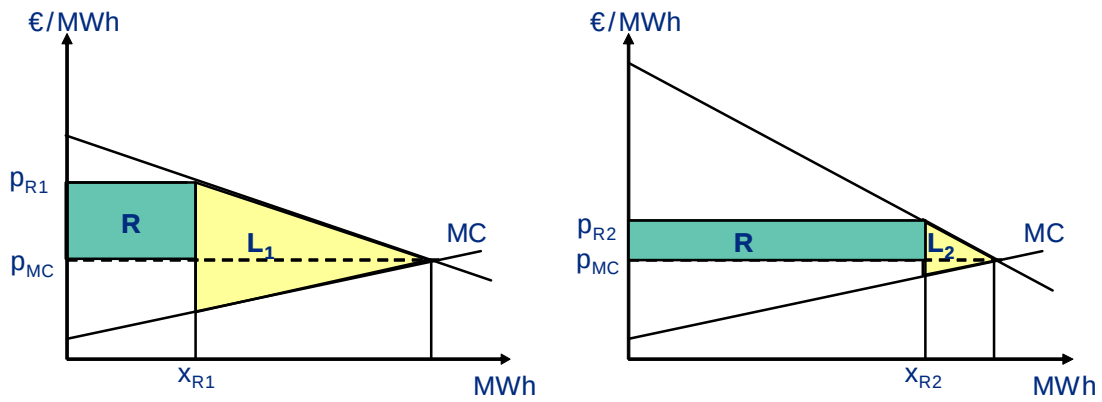
[17] As Norway is a primarily hydro-based system, inflows both nationally and regionally vary significantly, which also affects power flows, losses and congestion in the grid.

[18] Another alternative is so-called Aumann-Shapley prices, which give cost coverage to the monopoly anda fair allocation of the shared costs between different product types. They can be calculated without knowledge of the demand conditions, but require instead detailed knowledge of the cost situation in the monopoly. Aumann-Shapley prices are not discussed much in the literature on natural monopolies, and have little practical relevance for the electricity grid.

- *Ramsey pricing*. Ramsey-Boiteux prices, or just Ramsey prices, are prices that are differentiated in accordance with the demand elasticity of the different customers or customer groups. The idea is that the customers with the least price-sensitive demand should pay the largest relative mark-ups on the marginal short-term cost of supplying these customers. Such a differentiation ensures that the grid company will cover its costs at the same time as the distortion in demand compared to the economically efficient solution (where all customers meet a price equal to marginal costs) is the smallest possible. Using Ramsey prices in the electricity grid will mean a mark-up on the tariff per MWh which varies between customers depending on the price sensitivity of the demand. The drawback of Ramsey prices is that they require detailed information on the price sensitivity of the demand, and they can be perceived as unfair from a wealth distribution perspective.
- *Two-part tariffs*. With two-part tariffs the grid customers pay a tariff per MWh consumed or injected, and a fixed part that can be designed in different ways. The criterion for the variable part is that it reflects short-term marginal cost (such as transmission losses and capacity restrictions). The fixed part should ideally fulfil the criteria of optimal utilisation of the grid and correct investments, that is, give as little as possible distortion on the decisions on the use and development of the grid. The difference from Ramsey prices in the classical sense is that Ramsey prices refer to a situation with only a single MWh-based charge for network use.

The figure below gives a stylised presentation of the principle behind Ramsey prices. We look at an electricity grid with two customers, where one has a demand which is more price-sensitive than the other. As earlier the short term marginal cost curve is rising. We disregard capacity constraints. The residual income need is given by the area F in the figure. Assume that both customers start by paying the same price p_{MC} which is equal to the marginal cost. In the left part of the figure we show the effect of covering the whole residual income need with a mark-up on the price of the customer with the most price-sensitive demand (the flattest of the two demand curves). The price must then be increased to p_{R1}, which leads to customer 1 reducing his consumption to x_{R1}. This gives rise to a welfare loss due to the reduced consumption. Both the producer and the consumer's surplus are reduced, but some of the loss to the producers is compensated by a reallocation of consumer's surplus to producer's surplus due to the higher price. The loss is given by the area marked with L_1. In the right part we show the effect for the customer with the least price-sensitive demand. The consumption is reduced to x_{R2} when the price is increased to p_{R2}, and the welfare loss will be L_2.

Figur 3.6 Residual cost recovery through Ramsey pricing



We see directly that the welfare loss L_1 is larger than L_2. That is because the price to customer 1 must be increased relatively more so that the equal amount R is to be covered. Given that the residual income need is to be covered through a mark-up on the variable tariff per MWh, it will be desirable to set the price to customer 1 equal to the marginal cost and let customer 2 cover the whole residual income need. That gives the smallest total reduction in consumption.[19]

With optimal two-part tariffs both customers will pay a variable price equal to the short term marginal cost p_{MC}. That will give the maximum utilisation of the grid. In addition both of them will pay a fixed charge, which can be equal for both or differentiated depending on the tariff design. The only condition is that the total of the two fixed charges is equal to the area F. For such charges also, it is possible to use the Ramsey principle as a guideline, although they are not Ramsey prices in the strict definition. The charges are then designed so that they to the least possible degree affect the demand for transmission of electricity in the short and long term.

A note on power generation

The discussion above was based on different kinds of end-users of electricity, but the fundamental economic logic is to a large extent also valid for power generation. In practice, this means that the most profitable generators should pay a greater share of the residual costs. The profitability of generators will obviously vary over time and between technologies.

3.4.2 Other allocation principles

It is possible to use other criteria than willingness to pay and price sensitivity of the demand in a practical tariff model. Some grid customers will for example have larger requirements for quality of supply than others. This can be an argument to say that these should pay a larger share of the residual tariffs than other customers. However, there will be a relationship between the required quality of supply and willingness to pay for grid investments, so this is not necessarily an independent allocation criterion, but rather a derived one.

Another question is whether the system load in a wider perspective should be the basis for allocation of residual grid costs. The system load will partly be reflected through tariffs that reflect the short term marginal costs (together with the power prices), that is, a marginal loss tariff and capacity fees, but it is also possible to imagine the system load reflected through other kinds of tariffs. A related question is whether different technologies for power production inflict different costs on the power system (for example due to different need for ancillary services and balancing power), and if this should have any consequences for the tariffs.

3.4.3 Consumption versus production

We have so far discussed the allocation of the residual income need between different groups of end users (or groups of producers). Another question is how the income from

[19] Actually the price to customer 1 needs to be raised even more to cover the fixed costs. That is because the grid company receives a net profit from the variable tariff, which reflects marginal losses that are in increasing function of the consumption level. Hence, a lower consumption level will reduce the contribution from this tariff, thus increasing the need for residual revenue. We have for simplicity disregarded this effect in the figure.

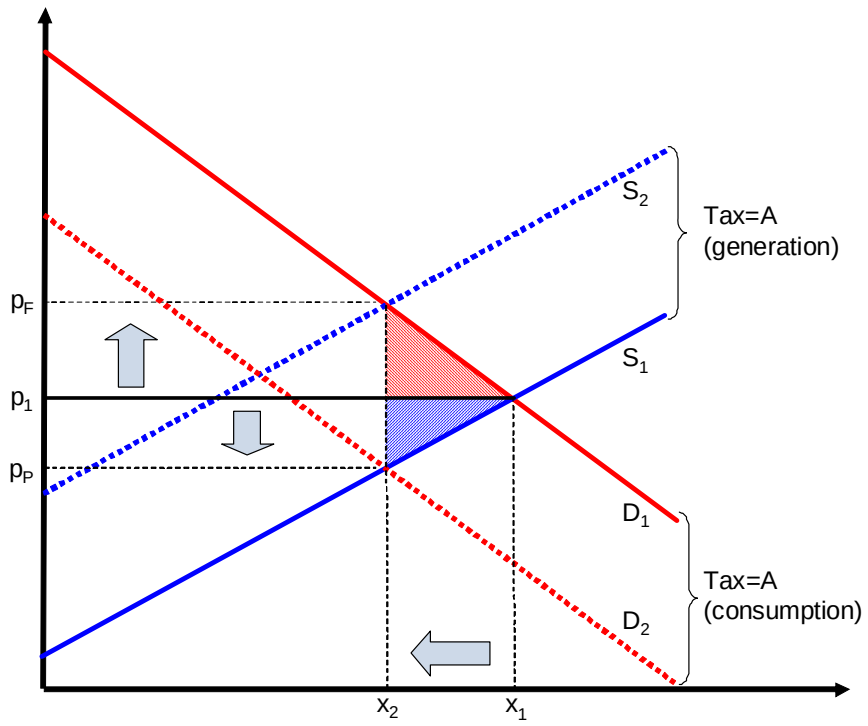
residual tariffs should be allocated between consumption and production. The answer depends on whether we look at a closed or an open economy.

Closed economy

We will first consider the impact of a network tariff on the supply and demand curve in the power market to clarify the effects of a tax or tariff in a closed economy.

It is a well-known result from economic theory that the outcome is the same whether a tax is placed on the demand side or the supply side in a homogeneous market. This principle is illustrated in the figure below. Before the introduction of the tax we have a supply curve S_1, demand curve D_1, traded volume x_1 and marked price p_1. If a tax of A per produced unit is applied for the generators, the supply curve will shift outwards with this amount to S_2. The traded volume will be reduced to x_2, the consumer price increases from p_1 to p_F and the producer price goes down from p_1 to p_P because of lower demand. If the same tax is levied on the consumers instead, the demand curve would shift inwards to D_2. The result on the producer price, consumer price and traded volume will however be just the same as when the tax was levied on the producers.

Figure 3.7 Impact of a generation or demand tax in a closed economy



How much of the *real burden* of the tax which is placed upon producers and consumers will depend on the relative steepness of respectively the demand and supply curve. In our example we can see that the price increase for the consumers is somewhat larger than the price decline for the producers, resulting in a larger relative loss for the consumers. The welfare loss is given by the total of the area given by the shaded triangles, where the red on top shows the loss for the consumers and the blue below shows the loss for the producers. The most important result is however that the economic burden of the tax is independent of who pays it.

At a general level, thermal production technologies have approximately constant returns to scale, which means that an increase in production by 10 per cent also increases the

costs by 10 per cent. In such a case, the long-term supply curve is horizontal. The entire real burden of a tax will in this case in the end be carried by the consumers, regardless of how the tax is collected. It is not necessary that the whole supply curve is horizontal, it is sufficient that the relevant part of the supply curve is flat. With a totally inelastic demand (a vertical demand curve), there will be no loss at all.[20]

Open economy

We have shown that if a tax is placed on all producers or all consumers in a market, it is (in the long-term equilibrium) the same who actually pays the tax. The real burden is not affected by who pays the tax, but by the slope of the supply and demand curves.

In a market where more countries participate and they practice different principles for tariffs, and perhaps also with different historical cost levels to cover, the situation will be different. Differences in the residual tariffs between countries can then distort competition and lead to a suboptimal allocation of production and consumption. The economic consequences will however depend on a range of factors such as power generation technology, the composition of consumption (the amount and kind of industry) etc. We return to this issue in chapter 4.

One qualifying remark should be added, however: If tariffs in other countries are “too low” in an economic sense (i.e. some kind of subsidisation is taking place), reducing the domestic residual tariffs for generators or certain end-users will increase the tariff burden on other domestic customers above what an economically efficient level would imply (i.e. a form of cross-subsidisation through the tariff system).

3.4.4 Tariffs on capacity/load versus tariffs on energy

The residual tariffs can be based on parameters that are more or less fixed or variable, at least in the long run. Two obvious alternatives are energy (MWh) or capacity/load (MW). The question of whether the residual tariffs should be based on energy or capacity has been discussed in former work both in Norway and internationally (see for instance Stoft et al., 1997). We will limit ourselves to give an account of the main conclusions below.

The interaction between residual tariffs and the power market

The residual tariffs will impact the demand and supply curve in the power market depending on the design of the tariffs. As the residual tariffs by definition are tariffs that come in addition to the efficient short-run and long-run price signals, the question is how they move the power market equilibrium away from the social optimum. Both demand and supply may be distorted in this manner.

The optimal residual tariffs will yield the economically efficient price structure in the power market. With price structure we mean *how the prices are in different load periods* from off-peak load to extreme peak load. From general economic theory we know that an optimal price structure reflects the marginal costs of producing a good and ration the available capacity. In the long-term optimal solution, capacity should be built to a level where the prices cover the long-term marginal costs. In a power market with

[20] The demand for electricity-specific consumption will be almost insensitive to the price, at least up to a certain price level.

efficient competition and with no residual tariffs the short-term prices will be correct, and in the long-run equilibrium the market will also realise the optimal generation capacity and technology mix, and thereby the corresponding optimal price structure.

To simplify, we can say that the market will realise a solution where the peak load price reflects the cost of producing peak load and the off-peak load price reflect the costs of producing base load. With such a price structure, there will be an optimal balance between production and demand-side measures. The consumers will have correct incentives with regard to the extent that they should invest in demand-side measures, i.e. move load from peak load to other load periods, or reduce consumption through energy saving or conversion from electricity to other energy carriers if possible.

Generation

The optimal residual tariffs for generation and consumption may be different. We consider generation first. Both tariffs on installed or maximum capacity and tariffs on produced energy move the market away from the ideal solution. Tariffs on capacity, however, have larger distortionary effects than tariffs on produced energy. This happens because tariffs on capacity lift the prices under peak load significantly, while tariffs on produced energy give the same small price increase in all load periods. Tariffs on capacity do not influence the short-run marginal cost of generation; rather, such tariffs influence the supply curve and the consumption and generation of electricity in the long run through the relative profitability of different generation technologies and the impact on the price structure and equilibrium outcome.

An important general result is that in the long-term equilibrium, the burden of a tariff on installed capacity will be borne by the consumption in the relatively few peak load hours where the effect capacity is limited. When a small share of the hours needs to carry the whole cost of the tariff, the price effect of the tariff will be substantial in these hours. A tariff on capacity will tend to reduce the supply during peak load. A large price wedge is then created between the cost of producing at peak load and what the consumers must pay. This price wedge results in too little generation during these hours, and the consumers are forced to take unnecessarily expensive measures to reduce their peak consumption. Instead of taking advantage of “cheap” production possibilities, expensive measures are undertaken to reduce electricity consumption during peak load.

The most important conclusion can be summarised like this: Residual tariffs on the producers’ installed or maximum capacity will in the long run lead to an unnecessarily large scarcity of power at peak load and therefore unnecessarily high prices at peak load. This gives an welfare loss which is higher than the losses given by flat residual energy tariffs.

Consumption

Residual tariffs based on *maximum consumption* (during peak load) or installed capacity will function as a tax on peak load consumption This will in turn stimulate excessive measures to move consumption from peak load to off-peak or lead to lower (sub-optimal) investments in capacity for consumption.

For consumption there will nonetheless be an important exception: Consumers that only pay an average price will not receive the market price signal during peak load and will

therefore in an economic context understanding have an excess consumption of power during this period.[21] If a tariff on peak load consumption can be administered to this group in an easy way to reduce their consumption during this period it will give an increase in economic welfare. (Strictly speaking the off-peak consumption of this group should then also be *subsidised* because they "under-consume" power in this load block). The need for this kind of tariffs will probably decrease as the consumers are confronted with continuous price signals to a larger extent (for example as a result of implementation of automatic metering).

A residual tariff based on energy will have a distortionary effect both on use of the grid and long-term investment decisions. In practice, a tariff based on installed or maximum capacity may be the least costly solution, at least at the higher grid levels. For small end-users, a fixed residual tariff per point of connection will be neutral in principle.

[21] With excess consumption we mean that these consumers adjust to the lower average price and will not reduce their consumption in the relevant hours as they would have if they had needed to pay the higher peak load price for the consumption in these hours.

4 Demand for transmission in practice

In the previous chapter, we argued that the residual income need should be covered by tariffs which distort the demand for transmission to the least possible degree. This again is a question of the customers' willingness to pay for transmission. In the following, we analyse important characteristics of the demand for transmission for different customer groups in the short and long term, with empirical illustrations from the Norwegian power sector in particular, but also from generation in the other Nordic countries. We concentrate on the long term willingness to pay for grid connection, including the willingness to pay for remaining connected, but we also discuss the price-sensitivity for transmission in the short run (use of the existing grid) along with some other issues such as quality of supply. We focus on tariffs based on energy, capacity/ load and other factors directly related to the consumption and generation of power.

4.1 Power generation

4.1.1 Willingness to pay in the short run

A residual tariff can give rise to a welfare loss if it affects the decisions to generate power in existing plants, so that the demand for electricity must be covered at a higher cost, or by affecting the power price so that the total demand is reduced (or a combination of the two effects). If the resulting price reflects more than the production cost of the marginal unit, including environmental costs related to CO_2 emissions and other relevant factors, this is by definition a suboptimal outcome. The question is in other words how residual tariffs affect the supply curve of the power market. In line with the conclusions from the previous chapter, we focus on energy-based residual tariffs in the following. We discuss the impact of the tariff from a Norwegian perspective, under the assumption that Norwegian generators compete in a common Nordic market.

A residual energy tariff based on actual power fed into the grid per hour, or average historical energy production, will obviously increase the marginal cost per MWh of a power plant. Even if the marginal cost increases above the social costs of generation, however, it will not necessarily affect the production decisions to any large extent. The effect is highly dependent on the technology:

- For wind power the marginal cost of production is relatively small, in practice close to zero in a given hour. An energy-based tariff at the present Norwegian level of 0.64 €/MWh[22] will therefore not affect the production unless the power prices are very low. The tariff can probably be several times higher before the production is affected in any significant way. Even if what is normally classified as variable costs per kWh will typically be around 1.1 € in Norway,[23] these are costs that are independent of the production in a given hour. This can for example be operation and maintenance costs that in the longer term are related to the

[22] Given a NOK/€ exchange rate of 8.75.

[23] Average operation and maintenance costs for a Norwegian wind power plant in 2007 according to the Norwegian Water Resources and Energy Directorate production statistics for wind power. There are large variations, however. Per MWh produced the costs varied between 0.25 and 2.5 €/MWh.

production volume, but which do not affect the production decision in the short run (costs to staff for stand-by during production and maintenance are not affected by a marginal increase in production).

- A hydropower plant without regulating ability (without a reservoir) will to a large extent be operated in the same way as a wind park and will not be affected by a energy-based tariff, as long as the tariff is not higher than the expected power price less the variable operation costs.
- For a hydropower plant with regulating capacity the picture is more complicated, but it is probable that the production decisions in Norway today are mainly unaffected by a moderate energy-based residual tariff. The production decision of such a hydropower plant is based on the estimated water value, which reflects the opportunity cost of using the water for production in another time period.[24] The water value depends among other on expected inflow to the reservoirs, the reservoir capacity, production costs of thermal power plants in Norway and other countries, and finally transmission capacity constraints. The marginal cost itself of a Norwegian hydropower plant is normally very low and in practice consists only of grid tariffs and certain production taxes,[25] which normally have little influence on the water value. A flat energy-based tariff will affect the water value just the same in all periods and will by definition not change the incentives to move the production. It is possible to construct some special cases with extreme inflow conditions (for example during spring floods with full reservoirs) where production is actually affected, due to the fact that prices will tend to be very low under such circumstances, perhaps even lower than a small or moderate residual energy-based tariff. These situations arise too rarely to influence the social costs and benefits noticeably, though.
- In thermal power plants based on coal, gas or biomass, the marginal cost is affected by an energy-based residual tariff. This is capacity that in many situations will be marginal in the Nordic power market, even if the competitive situation will vary quite a lot over time with changes in fuel prices for gas and coal. For bio (and waste) the access to fuel and the corresponding fuel costs may also vary. There is not much thermal power generation in Norway today, but there is one gas power plant with 430 MW capacity located at Kårstø and another large CHP plant under construction to be located at Mongstad. Furthermore there is a strong political support for increased use of bioenergy, which will probably lead to more bio-based co-generation capacity in the future. Because the marginal cost of such technologies is likely to be close to the cost of the marginal capacity in the Nordic market, a residual energy tariff on thermal generation in one country will affect the production decisions. It will also cause a welfare loss if more expensive capacity in countries without such residual tariffs is dispatched instead.

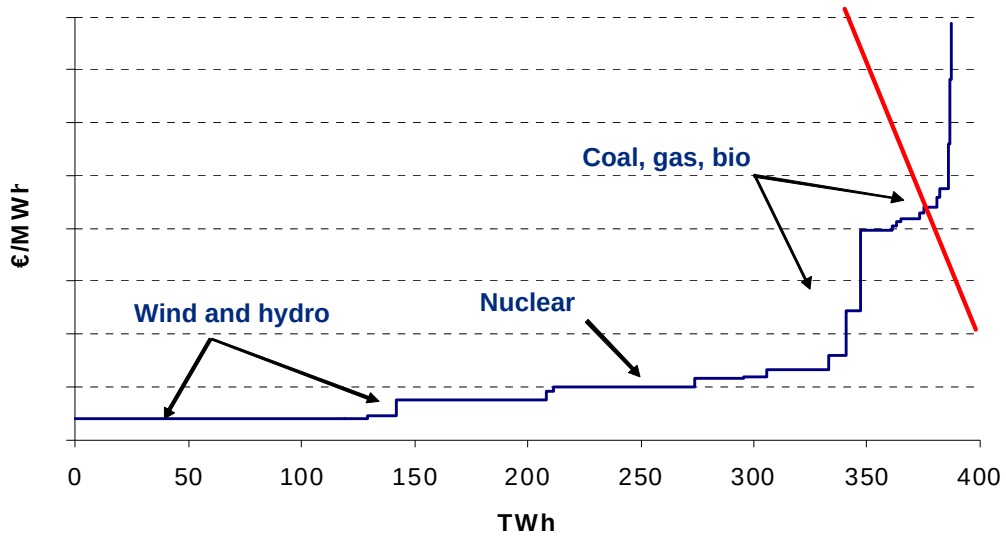
[24] This is true even if the hydropower system is part of a larger power market where the power price is set in interplay with countries with other tariffs. The total production in a hydropower system is given by the inflow to the reservoirs and run-of-river plants, and will not be changed by an energy-based tariff. The allocation of production over time is given by the regulating ability and price differences. A tariff that is the same in all periods will not affect the optimal allocation of the production over time, and when the production pattern is given the market prices will also be the same as without the tariff. The consumer prices are then not changed. In the short run, an energy-based tariff will work as a confiscation of profit from the generators and is not passed through to the consumers. The same will also be the case in a hydropower system with a tariff based on installed capacity.

[25] The Norwegian property tax for hydropower can be seen as a tax on production in this context as it is based on average generation.

Obviously, the exact effect on the supply curve will depend on the *relative* principles and levels of the residual tariffs in the different Nordic countries.

The figure below shows a stylised version of the short term supply curve in the Nordic power market based on the annual production per technology. The supply curve varies during the day and between seasons, and can change substantially over time with changes in coal prices, gas prices and other factors. Of course, demand will also vary and influence what is the marginal equilibrium capacity. The main picture with regard to the competition between the technologies is however as shown in the figure.

Figure 4.1 The supply curve in the Nordic power market – illustration



Source: Econ Pöyry

4.1.2 Willingness to pay in the long run

In the long run, the important question is how investments in power production are affected by the residual tariffs, both investment in new power plants and reinvestments or upgrades/expansions of existing power plants. Seen from Norway, the question of investment must be viewed in a Nordic perspective just as the short term production decisions. We base the discussion on residual tariffs set by the average energy production in line with the current Norwegian rules and the criteria for optimal residual tariffs that we discussed in chapter 2.

Investments in Norwegian power production will generally be less profitable as the result of a residual energy tariff. There are however interesting differences between the technologies:

- New hydro power projects with unit costs (including capital costs) that are lower than the expected future market price can be profitable also after a potential energy tariff. At the same time the costs of hydro power projects cover a wide range, and some projects will always be marginal. But the more profitable the project, the larger the share of the residual grid costs it can carry. The profitability of single projects is however difficult to estimate without detailed information of the costs and production potential, and the profitability will obviously be exposed to significant uncertainty.

- In established hydro power plants, a residual energy tariff will at the outset simply reduce the profits and not lead to plant shut-down, but certain investments will however be affected, especially upgrades and expansions in existing power plants, but also ordinary reinvestments given certain conditions (typically replacement of capital equipment in relatively expensive plants).
- Investments in renewable power production projects that are entitled to public support funds (wind power or bio-based power generation) will not necessarily be affected by a residual energy tariff. This depends on the support scheme and levels. A direct investment support scheme based on individual project profitability, for example, will in principle compensate the project owner for the tariff costs as the investment support is set so that the net present value of the project is zero. A feed-in tariff will give the project owner a guaranteed price that can be set so that it also covers the tariff costs.
- Thermal power projects based on gas without CO_2 capture and storage (CCS) are generally assumed to be marginal or even unprofitable given the current power and gas prices in Norway and the rest of the Nordic region. Given that the projects are marginal, a residual energy tariff will lead to investment in other more expensive technologies. Investment in thermal power production without CO_2 capture is not politically possible in Norway currently, but this could of course change.

The costs of different technologies will vary over time as a result of changes to costs of equipment and manpower, support schemes for renewable energy, production specific taxes, and climate policies etc.[26] Network tariffs may of course vary over time and between countries as well. Investors may also have different power price expectations and face different costs of capital.

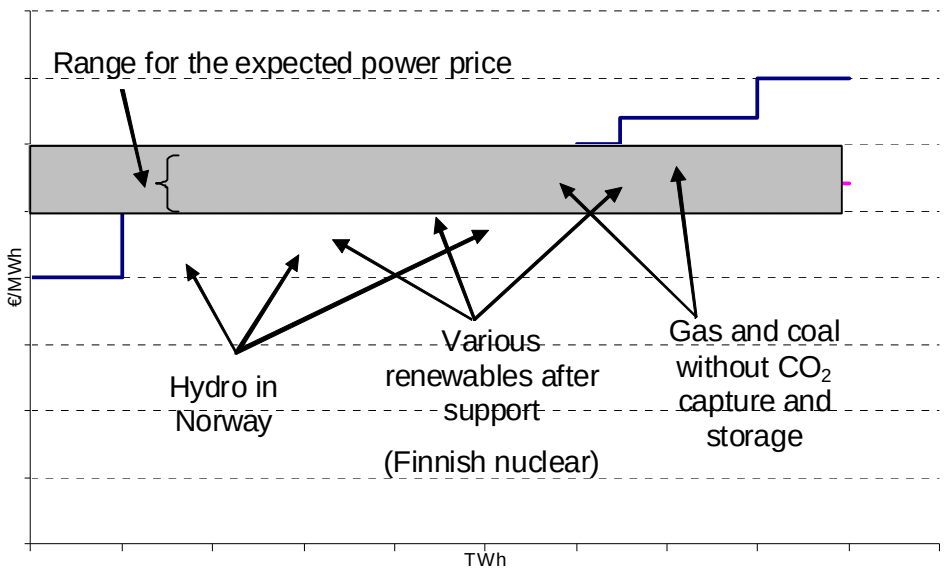
The policy trade-off regarding residual tariffs for generation is thus whether to tax existing plants to the maximum or facilitate investments in new capacity. If it is possible to distinguish between the two, the ideal solution would be to have existing plants (particularly hydro) pay a significant residual tariff and shield new plants from paying tariff costs above the network costs they actually impose on the system.[27]

The figure below shows an example of how the long term supply curve may look for the Nordic power market. The curve is consistent with observed investments the last years in general. For instance, we have seen investment decisions on nuclear power plants in Finland, several hydro power projects in Norway (especially small-scale hydro power), a gas- fired power plant in Norway and conversion to bio in Sweden. Several wind power projects in Norway however have been put on hold because of insufficient support in the current scheme and uncertainty about the future support system.

[26] The costs of new production in the Nordic countries will roughly follow the same pattern but the levels can vary some depending on the access to gas/coal, wind resources etc.

[27] In practice, this could be achieved through a profitability-based residual tariff for generation which ensured that only projects with a profitability above a normal rate of return paid residual tariffs. This has clear parallels to a resource rent tax in the classic economic sense.

Figure 4.2 The long-run supply curve in the Nordic region



Source: Econ Pöyry

4.1.3 Other factors

Regarding quality of supply, an outage or other quality deviations (voltage dips etc.) will lead to a loss for a producer if it causes a situation where he or she can not feed energy into the grid. This loss can in theory be zero for a hydropower producer with regulating ability if the water can be produced at a later hour to the same price (adjusted for the cost of capital or discount rate, depending on the delay). Equally the loss for a thermal power producer may be limited due to the fuel saved. For a hydro power plant without regulating ability and wind power plants, an outage will lead to a loss of income, although this loss is limited to the power price for the hours in question.

Different kinds of power production may have different effects on the power system and the grid costs. The uncertainty of the production from wind power for example may increase the need for balancing power and ancillary services. In today's system, the Nordic spot market is cleared 12-36 hours before the actual production hour, which will leave a not insignificant uncertainty about the wind strength hour for hour for the actual production hours. Even though the quality of wind prognosis is improving, there is still significant uncertainty left, for example on the timing of wind speed changes. The same is true for hydro power without regulating ability, but the uncertainty is less as the variations are slower than for wind.

Even if certain production technologies should increase the costs of the total system, it is not given that this should have consequences for them through the design of residual tariffs or other tariffs. In part, increased investment in physical assets for providing i.e. ancillary services for some technologies will be reflected through customer-specific connection charges or investment contributions. Other measures such as the methodology for congestion management and technical standards regarding quality of supply can also in part substantially reduce the system costs. In Norway, the possibility of coordinating the production of wind power parks with flexible hydropower with reservoirs will also reduce the need for balancing power caused by the uncertainty of the wind production.

4.2 Power-intensive industry

The consumption of the Norwegian power-intensive industry and the wood-processing industry varies, but has been in the 35-40 TWh range annually since the end of the 1990s.[28] The power-intensive industry operates in a competitive international market. High energy costs in Norway may lead to temporary shut-downs of industrial activity in the short run and emigration of industry in the long run. Many industrial companies have long-term power contracts so that they are not necessarily exposed to the current power market price, but such contracts do not necessarily include grid tariffs.[29] A residual tariff based on energy can therefore affect the companies' decisions to uphold their consumption within Norway, even if they have long-term power contracts. Similarly, residual tariffs on maximum or installed capacity will have an impact on the long-term profitability of connecting to – or remain connected to – the grid. Residual tariffs based on energy will also influence the short-run use of the grid.

It is not the subject of this report to do a study of the socially optimal power prices including grid tariffs to the industry as such, but to evaluate how residual tariffs that in principle do not reflect costs directly attributable to the industrial activity affect the decisions of the industry in the short and long term. We are in other words discussing the potential to "tax" the industry beyond what they pay through the power price and variable grid tariffs that reflect marginal losses.

We consider first the impact of residual tariffs on the decision to maintain the current activity level, that is whether the existing industry will remain connected to the grid or not. Whether a shut-down due to residual tariffs will constitute a welfare loss to Norwegian society depends on a number of factors:

- The alternative value of the industry's capital, manpower and raw materials - that is, whether the input factors would render a higher yield in other sectors in the Norwegian economy.
- Readjustment costs caused by a shut-down (i.e. unemployment benefits, investments in new infrastructure or industrial activities etc.).
- Possible need for new grid investments to take into account the changed power flow and demand for transmission in the long run (ref also Statnett, 2002).
- Global effects on CO_2 emissions.

The effects of higher residual tariffs may be complicated in the short term. Contractual obligations can for example make it more profitable to maintain production in spite of higher energy costs if a stop would mean that the obligations must be met in other ways, such as buying in from other producers to a higher cost. Higher energy costs due to increased residual tariffs will therefore not necessarily lead to immediate consequences for the decision to maintain production. In the longer run, however, the production will be shut down when the residual tariffs increase so much that the operating margin

[28] In official statistics the term power-intensive industry is used for companies that are involved in production of chemical raw materials, iron, steel and ferro-alloys, aluminium and other metals. We use the term power-intensive industry about both the companies described in the official statistics and the wood-processing industry.

[29] This is at least true for the historic contracts based on terms decided by the Norwegian parliament that are being phased out (the last contract will end in 2011) even though the companies holding such contracts have the opportunity to sell power back to the market on certain conditions. New contracts based on commercial terms will often be purely financial contracts with full freedom to respond to the current power market prices.

becomes negative (and is expected to stay negative for a longer time). There will be substantial differences between companies and industries with regard to price sensitivity, but the general conclusion is the same.

The argument above is true for a closed economy. It is strengthened if we consider the possibilities to move the production to another country. In the long run a company may earn profits in Norway, but if the profits increase by moving the production to another country with lower energy costs, it will be profitable to shut down in Norway anyway. It is in other words the opportunity cost of producing in other countries that limits the possibilities of covering the residual income need from the industry. From a cost recovery perspective, it is better from a Norwegian economic viewpoint that power-intensive industry contributes with a small tariff rather than nothing at all.

High residual tariffs that make current industrial production in Norway unprofitable will of course also tend to make investments in new activity in Norway within power-intensive industries unprofitable. Whether this is desirable for new industrial plants is another question. This depends on whether the newcomer has to pay the full economic costs of his or her electricity consumption. Hence, having different residual tariffs for existing and new industry may in fact be economically efficient. In practice, it may be difficult to implement such a system (cf. the discussion on power generation above).

In sum, this means that the basis for recovering the residual income need from the existing industry is limited upwards to the operating margin in the companies before residual tariffs in the short run and in the long run to the opportunity cost of producing in other countries. This basis will vary over time and between companies and industries, but it is unlikely be very high. For new industry, the "tax base" depends on the profitability of establishing new activities in one location compared to others and whether the overall tariff system reflects the full grid costs of the new consumption.

Other factors

The power-intensive industry has relatively high requirements regarding the quality of supply even if there is certain variation between companies and industries. Even a short outage or voltage dip can in several industries lead to halt in the production because processes are interrupted and need to be restarted. In an aluminium plant the consequences of a shorter outage can be limited and it is also possible to produce with a reduced capacity for a period. But the costs can also be substantial and outages above a couple of hours can have very large economic consequences. However, the value of quality of supply is still limited by the inherent profitability of the industrial production.

The industry consumption is also characterised by being relatively more flexible to that of other customer groups, both in the short and long term. The ability of the industry to respond to short term price signals caused by a scarcity of capacity or energy has a considerable value for the power system. This is a characteristic that is rewarded through the short term price signals (power prices, marginal losses and capacity fees), possible arrangements for energy and capacity options sold to the system operator and similar market-based measures.[30] With respect to the residual tariffs, the short-term flexibility should only be relevant to the extent that such characteristics are not reflected in other price signals. In other words, if the industry is rewarded for its flexibility

[30] For instance, it is possible for the Norwegian system operator to procure balancing services through an option contract with power-intensive industry.

through ordinary tariffs and other market-based instruments, the residual tariffs should not be affected.

4.3 The petroleum sector

The petroleum sector is a very important part of the Norwegian economy and it also accounts for a considerable consumption of energy, both on the Norwegian continental shelf and onshore. So far most of the sector's energy demand has been covered by gas turbines on the platforms themselves, but some offshore installations also use electricity from the onshore grid through high-voltage sea cables. There is however some onshore electricity consumption from petroleum activities. Most of the onshore consumption is related to onshore installations such as gas processing plants, refineries and one LNG plant.

The petroleum sector's power consumption onshore is modest in comparison to the power-intensive industry (around 10 per cent), but it is increasing. The energy consumption per extracted unit of oil or gas will for example normally increase during the life time of a field. There is also a certain political pressure to reduce the CO_2-emissions from the Norwegian Continental Shelf through electrification of offshore installations, even though the costs of this can be very high.

The power consumption of the petroleum sector is to a large extent characterised by it being connected to natural resources that are located in specific places and cannot be moved. There are often certain degrees of freedom to choose where to bring ashore gas pipelines and where to place other onshore installations, including the exact point of connection to the grid, but it is normally no alternative to go outside of Norway. The choice between power supply from the onshore grid and supply from gas turbines on the platforms is however an important parameter within the limits set by the overall politics. For onshore facilities such as installations to bring ashore gas it is normal to connect to the mainland grid, even though it is also in these cases an alternative to construct an installation-specific power plant instead (which is not connected to the ordinary onshore grid).

The willingness to pay for grid connection in a petroleum project depends on the profitability of the underlying project. For marginally profitable projects, any residual grid tariff could overturn the project or lead to the choice of a different energy supply solution if electrification from shore is not mandatory. For projects with a profitability above a normal rate of return, there will however be a willingness to pay or a basis to cover the residual income need of the grid.

We can illustrate the effect of a residual tariff on a petroleum project with an example:

- A project has an investment cost of €5 billion and an assumed life time of 15 years. The project will consume 1 TWh annually to be supplied from the onshore grid.
- The investor's required rate of return is 10 per cent real before tax.
- Assume that the project yields a return of 15 per cent real before tax, and before having to pay any residual tariffs. This corresponds to a net income after deduction of costs to energy, operation and maintenance with a net present value of €7.8

billion. This implies a substantial excess return above a normal rate, which is typical for many petroleum projects.[31]

- Assume then that the project needs to pay a residual tariff of 5 €/MWh, that is, €5 million per year. Given a required rate of return of 10 per cent this will have a net present value of approximately €38 million. The net present value of the whole project is then reduced to about €7.4 billion and the return of the project is reduced to 14.9 per cent. If the tariff is doubled to 10 €/MWh, this will reduce the return to 14.8 per cent.

A residual tariff will in this way reduce the yield of a project, but the effect is quite modest in our example. As long as there is a excess return of at least 0.2 per cent real before tax, the project will be able to carry a residual tariff of at least 10 €/MWh.

The example illustrates that there is a substantial willingness to pay for grid connection in petroleum projects beyond what follows from power prices and short-term marginal costs. Note that we have not taken into consideration the alternative cost of choosing an installation-specific power generation or dedicated power plants without connection to the onshore grid. If such supply is technically and politically possible, it can limit the willingness to pay for grid connection in comparison to our example.

The willingness to pay for *remaining* connected to the grid similarly depends on the project economy, but now from the point of view of an existing activity.

The possibilities to cover the residual income need from the petroleum sector will then depend on the project economy, which again depend on the prices of petroleum products and the costs of extracting oil and gas. With high oil and gas prices the “tax base” for residual grid costs from the petroleum sector is substantial, despite the quite limited power consumption of the sector.

Other factors

The power consumption of the petroleum sector will be similar to that of the industry with respect to requirements to quality of supply. For example, gas compressors, refineries or other installations can stop even because of short-term outages or voltage dip. However for this sector as well, the willingness to pay for quality of supply will depend on the value of the products to be sold (oil and gas after deduction of the production costs).

4.4 Households, small businesses and the public sector

So-called “general consumption” account for roughly 70 TWh of the Norwegian net consumption of electricity. A significant part is used for heating of water and buildings.[32] General consumption is defined to be the consumption of households, holiday houses, agriculture, public sector, smaller industries, services and more.

Their willingness to pay for grid connection is generally high per kWh and probably also increasing. This is because grid connection is essential for electricity-specific

[31] There are examples from the Norwegian Continental Shelf with significantly higher returns than this.

[32] For households, this amounts to about half of the their total electricity consumption according to estimates from Statistics Norway.

consumption such as lighting and data and communication equipment among other. Many of the grid customers also use only electric heating and have no or poor access to alternative sources of heating. This is also supported by estimations made on the value of quality of supply for this customer group (which again is related to the willingness to pay to be connected to the grid), as seen from the value of lost load found in surveys carried out by the Norwegian research institutions SINTEF and SNF. Also, the mere quantity of the electricity consumption in the sector gives a substantial base towards covering the residual income need.

Residual tariffs may however still affect the decisions of these consumers in several ways:

- A residual energy tariff per kWh of consumption will lead to a lower grid utilisation in the short run than what is economically beneficial to society. The price sensitivity of the demand of the general consumption is however small, especially in the very short term, so the negative consequences of this are not necessarily very big.
- In the long run, a residual energy tariff will give incentives to choose alternative sources of heating to electricity above what follows from power prices, tariffs that reflect short-term price signals, customer-specific connection charges, fees and potential support schemes for alternative heating. This can lead to welfare losses if the residual tariff causes a more expensive alternative solution to be chosen.
- Tariffs based on installed or maximum capacity will also give incentives to invest in alternative heating solutions due to the higher cost of electric heating. This may again cause welfare loss if the alternative solutions are more expensive in total than electricity.
- Residual tariffs per customer or per meter will to a small or no degree affect the grid utilisation in the short run or the choice of heating solution in the long run.

On average, the consumers in the distribution grid will have a very high willingness to pay for connecting to the grid, and staying connected. There will nevertheless be significant differences between different end-users within this general category. In particular, small businesses may have a higher sensitivity to changes in the network costs than households or the public sector.

4.5 Summary – the demand for transmission

In this chapter, we have discussed important characteristics of the demand for transmission and distribution from a Norwegian perspective, both short-term use of the grid and the long-term willingness to pay for connection or remaining connected, for a total of four main groups of grid customers. In the table below, we sum up the most important characteristics of the willingness to pay for connection and the price sensitivity for remaining connected in the long run, which are particularly important for the design of optimal residual tariffs. Both in the power-intensive industry and power generation there are variations between companies and power production technologies with regard to price sensitivity. The large share of installed hydro power production in Norway means that power production must be seen as less price sensitive overall than the industry with regard to maintaining their connection to the grid.

Table 4.1 Characteristics of the long-run demand for transmission and distribution for different groups of Norwegian network customers

Customer group	Willingness to pay for connection	Willingness to pay for remaining connected	Notes	Tariff base
Generation	Limited by the power price, low to medium in general but high for the most profitable projects	High for existing hydro, medium for other technologies	Varies between technologies and over time	122 TWh
Power-intensive industry	Low - limited by the opportunity cost of producing elsewhere	Low - limited by the opportunity cost of establishing new activities elsewhere	Varies between technologies and over time	35-40 TWh
Petroleum sector	Medium - limited by the opportunity cost of other energy supply options and the profitability of the petroleum activities	Medium - limited by the value of maintaining production or switching energy supply, but likely to be significant	Varies with the oil price	4-5 TWh
General consumption	Very high	Very high	Differences between households, businesses etc., but high on average and less variable than for other groups	70 TWh

5 The current Norwegian tariff scheme

In previous chapters, we have described important insights from economic theory on optimal tariff design along with important characteristics of the demand for transmission and distribution of electricity, and highlighted certain elements of the current Norwegian tariff scheme. In this chapter we give a more complete description of the system and discuss the economic characteristics in light of the theoretical analysis and the description of the demand for transmission and distribution. At a general level, the Norwegian system employs a two-part tariff which approximates locational marginal pricing, combined with widespread use of project-specific connection charges or investment contributions, and a distribution of residual costs which is fairly similar to a Ramsey principle.

5.1 Tariffs that give short-term signals

5.1.1 The energy part of the tariff

According to the current regulation the so-called energy part of the tariff or variable tariff is to be set to reflect the marginal losses of the grid as a main rule. This applies to all grid levels. While the marginal losses are integrated in the power prices in some power markets, the marginal losses in Norway is an explicit tariff element which varies with the hour of the day and location in the grid. The grid loss will vary depending on where in the grid the production feed-in or consumption occurs. The energy part is meant to reflect the system load inflicted on the grid by a given customer with the transmission of one more kWh and reflect the marginal loss costs inflicted by a given customer by the feed-in or withdrawal of power at a specific location. Because of this several marginal loss fees are defined for each point of connection in the central grid.

The energy tariff per MWh is equal to the Nord Pool system price (NOK/MWh) multiplied with the marginal loss rate at the connection point. Marginal loss rates in the central grid are updated every week, and there are different rates at daytime (between 6 a.m. and 22 p.m.) and night time/in the weekend. The rates are set symmetrically for production and consumption in a given point, so that a marginal loss rate of 5 percent for production is paralleled by a minus 5 percent rate for consumption. The marginal loss rate has an administrative limit of +/- 10 percent even though the actual percentage can be higher in given cases. This can be seen as an ex ante approach to nodal pricing.

In the regional grid the energy part is principally designed in the same way as in the central grid but the time-resolution is lower than the one practised by Statnett.

In the distribution grid a time-differentiated energy tariff must be offered to all customers with mandatory meter reading several times per year, which reflects assumed differences in losses between seasons due to system load (i.e., the load is typically higher in winter due to heating demand).

5.1.2 Price areas/capacity fees

As described briefly in previous chapters, *price areas* are used for managing congestion in the Nordic power market: The market is divided into areas based on (structural long-term) capacity constraints in the grid and the prices will then vary between the areas

depending on whether there is a surplus or deficit in the areas. The area prices can also be seen as a pragmatic approach to nodal pricing, and function as a capacity fee.

Higher prices in a deficit area stimulate increased production and lower consumption. In the short run this will be the case for already existing installations but in the long run it will also affect localisation decisions: It is as a basis more profitable to invest in new production in a high-price area and consumption in a low-price area. Different area prices signals where it is most profitable to invest, even if the signal is not as precise as with nodal pricing.

For short-term/temporary congestions within the price areas *counter-trade* or *special regulation* is used: Generators in surplus areas are paid not to produce when congestion arises, while generators in deficit areas are paid to increase their production.

Grid capacity fees are not used in the regional or distribution grid today, but it is in principle allowed to use such fees in these grid levels as well according to the current Norwegian regulation.

5.2 Residual tariff parts

5.2.1 Generation

Power producers pay a residual tariff for based on a ten-year average of the production in MWh (for 2008 the calculation basis is the average production from 1997-2006). The rate is 0.64 €/MWh. In certain regions of Norway there is a differentiated tariff for new production that under certain conditions are given a lower residual tariff (see description below). The rate of the residual tariff is set for the central grid and is normative for the lower grid levels as well, so that the residual tariff cost is the same for all production independent on the grid level it is connected to. The level of the residual tariff can vary within a range agreed by the Nordic Transmission System Operators in 2000 in order to harmonise the residual tariffs for generators (approximately 0.2-0.7 €/MWh).

Differentiated tariff for new production due to special grid conditions

In general, the Norwegian Energy Act with regulations allows for differentiated tariffs to grid customers due to relevant grid conditions and on objective and non-discriminating terms. This applies to generators as well, and since 2005 it has been possible to receive a reduced residual tariff for production feed-in (0.11 €/MWh) for new production in selected regions. Statnett decides which regions where this special tariff is to be valid and to what extent, and enters into contracts with producers who apply for the reduced tariff if the new production plant fulfil the terms set to receive it. The tariff is set annually and is valid in 15 years after a contract has been entered into. It is valid for all new production in an area and there is no discrimination between technologies. When there are sufficient investment projects for new production in the area with the special conditions, the special tariff will no longer be offered to new plants (these will then be faced with the normal residual tariff rate).

In 2007 there were two such areas: One in Central Norway with an allowed maximum production volume to offer the differentiated tariff of 9 TWh and the other in Western Norway with a maximum volume of 3 TWh.

Consumption

The residual tariff part for consumption connected to the central or regional grid is based on the measured peak capacity in reference hours. Choice of reference hours is to be unpredictable as far as possible.

In the central grid the residual tariff is settled based on the customer's average total consumption at peak hour over the last five years per connection point. This is again adjusted downward with interruptible consumption for the customer and further corrected with an adjustment factor per point, the so-called k-factor. The k-factor is set to reflect the total consumption in the point, available capacity during winter (production), total consumption of power-intensive industry in the point and an own adjustment factor for power-intensive industry. Power-intensive industry is in this context defined to be industry with a consumption of more than 15 MW at peak load and minimum 7000 load hours (exception from the limit of 7000 hours can be applied for if a company is very close to the limit or below it in certain years). This contributes both to reduce the tariff costs of the power-intensive industry and reward joint-location of production and consumption in the central grid.[33]

In the distribution grid a fixed part per customer or meter is set which is to cover at minimum customer-specific costs for households. Several distribution grid companies operate with differentiated fixed parts based on capacity/fuse size or house type. For businesses it is normal to use a capacity-based residual tariff part. Also most grid companies use the so-called energy tariff (see above) to a substantial degree to cover the residual income need. In 2006, the energy tariff per kWh of grid use amounted to 55 per cent of the total revenues in the distribution grid. This is clearly higher than the cost of losses and the income from a theoretically optial marginal losses tariff.

Both in the distribution grid and the regional grid the residual tariff parts for different kinds of consumption can vary between areas because of differences in historic grid costs (the cost base) for the network companies.

5.3 Long-term price signals

Tariffs that give short term price signals also give long term signals on the value of new production and transmission capacity and location of new power consumption. In addition, the residual tariff parts will give certain long term investment signals even though they are designed to be as neutral as possible:

- The differentiated tariff for new power production due to special grid conditions in certain areas is explicitly designed to give an extra signal of the value of new production in that given area.
- The k-factor model makes joint localisation of power production and consumption in the central grid more favourable than to choice separate location.

[33] Interruptible consumption shall in accordance with the current regulations be offered a lower grid tariff through a lower fixed tariff part (that is only to cover customer-specific costs for the customers who have agreements of instantaneous disconnection and equipment for remote disconnection). The discount is explained by the value for the grid of flexible consumption, but is proposed to be removed in connection with the Norwegian Government's strategy on bio-energy (that is, it will no longer be mandatory).

- The mark-up on the energy tariff in the distribution grid to cover the residual income need makes it more profitable to choose alternative heating to electricity compared to a situation where an energy tariff only reflects marginal losses.

In addition there are two other rules that also give long term price signals: Customer specific connection charges or investment contributions, and the rules on production-related grid installations.

5.3.1 Connection charges and investment contributions

The network companies may use a connection charge for new connections to the grid or upgrades of existing connections. This charge must be general (i.e. a fixed rate per connection regardless of actual costs) and be applied to all new connections in the relevant grid, but it may be differentiated according to fuse size.

In accordance with the current regulations, the grid companies can also decide to impose a customer-specific connection charge or investment contribution to cover the costs of new connections or reinforcement of the grid to existing customers. This opportunity to charge the customer-specific fee is valid for both production and consumption. The investment contribution is always based on the investment costs in each case, minus any general connection charge paid by the customer. Investments financed by such contributions are not included in the cost base for the normal tariffs of the grid companies.

As a main principle it is not possible to impose an investment contribution in meshed grids other than in very special cases. This goes for both production and consumption. This means that this instrument normally cannot be used in the central and regional grid apart from genuine radial connections (lines that go from the meshed grid to points for production feed-in or consumption without connection to other points in the grid) or investments that obviously are customer-specific (for example new transformer capacity at a generation site).

The investment contribution in this way gives signals about the costs of investments that are customer-specific, but to a lesser extent the costs of reinforcements that also benefit other customers (i.e. investments in the meshed grid).

5.3.2 Production-related grid installations

For so-called production-related grid installations the producers are obliged to cover all costs. Such investments will in other words not be included in the grid companies' cost base for grid tariffs. Production-related grid installations are power lines and other installations which main function is transmission of power from a production plant to the closest point of connection in the existing grid.

In a social cost-benefit perspective the rules for production-related grid installations and customer specific connection charges are largely concurrent. In the current regulation the most important difference is that the connection charge only includes the investment cost while the fee to production-related grid installations also include other costs such as operations and maintenance.

5.3.3 Overview

The table below gives an overview of the main elements of the Norwegian tariff system.

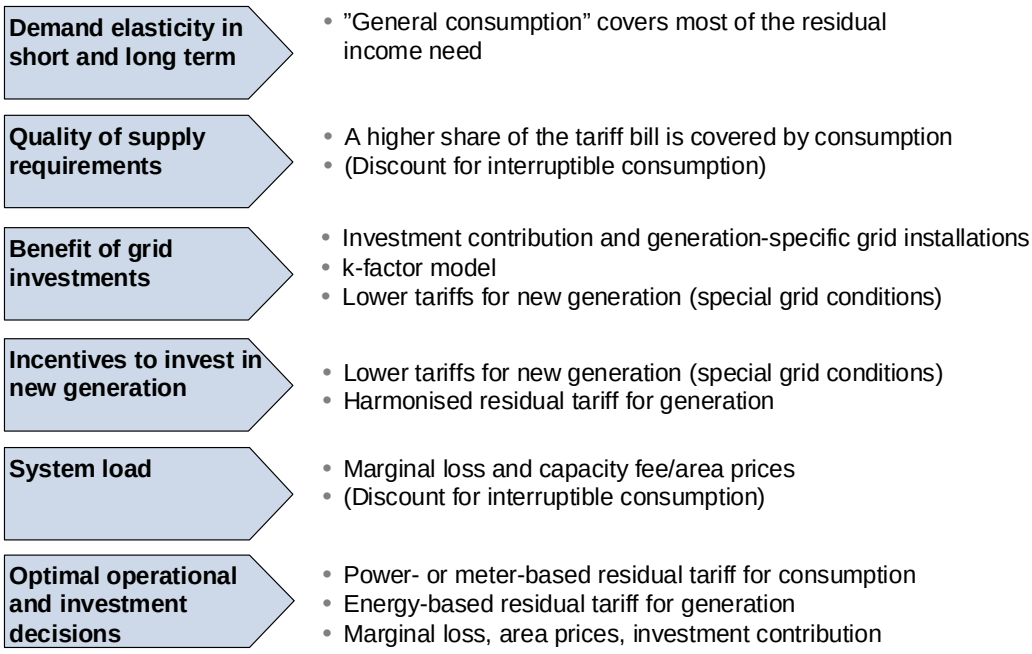
Figure 5.1 Overview of the Norwegian tariff system

		Central grid	Regional grid	Distribution grid
Generation	**Non-residual tariffs**			
	Marginal losses	Yes Differentiated by time and location	Yes Differentiated by time and location	Yes Differentiated by time and location
	Price areas	Yes		
	Investment contribution	Yes	Yes	Yes
	Cost of generation-specific grid installations	Yes	Yes	Yes
	Residual tariffs			
	Energy-based tariff for production feed-in	Yes Reduced tariff in certain areas	Yes Reduced tariff in certain areas	Yes Reduced tariff in certain areas
Consumption	**Non-residual tariffs**			
	Marginal losses	Yes Differentiated by time and location	Yes Differentiated by time and location	Yes Differentiated by time and location (consumption > 8000 kWh)
	Price areas	Yes		
	Investment contribution	Yes	Yes	Yes
	Residual tariffs			
	Maximum power-consumption based tariff	Yes Maximum power consumption adjusted with k-factor and differentiated fees for interruptible consumption	Yes	Yes
	Energy-based tariff			Yes
	Fixed tariff based per customer or meter			Yes

5.4 Economic consequences of the current tariff scheme

It is clear from the above discussion that the current Norwegian tariff scheme reflects many of the concerns that we discussed in the previous chapters and that the scheme is in line with the main purpose of the Norwegian Energy Act, which requires economically efficient tariffs. The main principles of the scheme are repeated in the figure below.

Figure 5.2 Characteristics of the Norwegian tariff system



The basic regulation builds to a large extent on the economic principles we described in chapter 2 and 3. For instance, the overall system leads to general consumption covering around 90 per cent of the residual costs (as a rough estimate).

The developers of new power production and consumption will to a large extent be met with the costs they inflict on the grid in both the short and long term. However, it is possible to make some improvements. For instance, it is possible to go further in the direction of nodal pricing in the central grid (that is, to establish more price areas of a smaller geographical size and to increase the frequency of the calculation of the marginal loss rates). It is in principle also desirable to go further in the use of customer-specific connection charges or investment contributions, and study the practical possibilities of using such charges to a greater extent in the meshed grid.

The need for coordinating investments in production, grid and consumption also raise some fundamental challenges to both the current regulation and the possible improvements. The coordination challenge will be relevant in many contexts. At the same time as the production technology mix is changing in Norway and Europe, there are also substantial changes to consumption. Implementing investments in the power system also takes a lot of time. Both the planning and construction phases are time-consuming processes. Statnett for example estimates that it will take about ten years to implement the grid extensions necessary to transport large amounts of new wind power from the north of Norway to the south. Grid investments are also increasingly controversial for environmental reasons, especially overhead lines in the central grid.

There are two challenges with regard to the consumption that are particularly relevant:

- The petroleum sector's increasing demand for power from the mainland especially to onshore installations is a challenge. A substantial growth in the size of 4 TWh on an annual basis is likely to be realised the following years, something that entails close to a doubling of the sector's power consumption from the onshore grid. Additional consumption growth can also be possible depending on the activity level on the Norwegian Continental Shelf and in the onshore installations,

and especially if authorities require increased use of electricity from onshore in the future. Parts of this increase in demand will likely come in regions with weak security of supply.

- The power-intensive industry can rapidly be expanded but it is equally likely that it can be reduced. In the first case we might get similar challenges as with the power consumption of the petroleum sector, while we in the second case can experience that certain areas will get large power surpluses (the county of Sogn and Fjordane is an example of this).

In both cases it is desirable that new production and new consumption receive the correct location signals. This will only partly be the case in today's system, because of the challenge caused by of lumpy investments and the limited room for investment contributions in the meshed transmission grid.

With regard to power generation there are also some more specific coordination challenges related to the market design as such. For instance, problems may arise if a marginal developer of a power plant triggers the need for extensive investments in the grid that he or she does not have the willingness to pay to cover, but which can be sensible to implement to prepare for further investments in power generation in the future. Another problem occurs when a new power plant releases the need for reinforcement in other areas in the grid where it is not possible or desirable to cover the costs with customer-specific connection charges or investment contributions. This means that the producer will not see the full economic consequences of his or her generation.

Optimal coordination of generation, grid and consumption is difficult to achieve fully with only general economic tools. It can be necessary with situation-specific tools. Customer-specific connection charges and differentiated tariffs for new power production due to special grid conditions are examples of such tools, but they will only solve parts of the challenges. For instance, it is questionable whether the discount on the residual tariff for new production in certain areas is sufficiently high to make all projects that are beneficial for society profitable also for the developer.[34]

Another issue is that any petroleum activity which fulfils the requirements to be defined as a power-intensive industry based on capacity and load hours may end up getting the same residual tariffs as this industry, however without this being in accordance with their long-term willingness to pay. A similar possible weakness is the geographical differentiation of residual tariffs for consumption based on the differences in historic grid costs (at the regional and distribution level).

On the whole, the current scheme is to a large degree built on economically correct principles, both with respect to the allocation of the grid costs and price signals. The allocation of grid costs should however be seen in accordance with the benefit of the grid investments over time for the grid customers. Changes in demand and generation will make it necessary to re-evaluate the overall allocation of network costs as the grid is being developed. With respect to price signals it is the need for coordination between

[34] Principally there is nothing wrong with setting a negative tariff for new production to increase the strength of the price signal. A challenge will be to decide the correct tariff level that will vary depending on the local grid situation and the available technologies for power production. To differentiate the tariff on a case-to-case basis will also bring along problems of strategic behaviour. This however can be limited by using auctions or similar tools to reveal the true cost of possible new production in a region.

production, grid and consumption that is particularly challenging. The residual tariffs can be used as a tool that can contribute to increased coordination through more precise price signals, but it is also important to be aware that such changes is not necessarily a very precise tool. Changes in the residual tariffs must anyhow be made based on a closer evaluation of a) the costs and benefits to society from changes in the tariff structure and b) if there are other tools that can contribute to solving the challenges of coordination in a more efficient way (increased use of customer specific connection charges, the process of applying for concession and other similar qualitative tools).

References

Bruneekreft, G., K. Neuhoff and D. Newbery (2005): "Electricity transmission: An overview of the current debate", *Utilities Policy* 13, 73-93

ECON (2006): *Tariffering av energimålte kunder i distribusjonsnettet.* Econ-report nr. 2006-002, ECON Analyse. (*Network tariffs for small customers in the distribution grid.* In Norwegian.)

Econ Pöyry (2008): *Fordeling av tariffinntekter mellom produksjon og forbruk.* Econ-report 2008-082, Econ Pöyry. (*Allocation of tariff revenues between generation and consumption.* In Norwegian.)

Joskow, P. and J. Tirole (2005): "Merchant transmission investment", *Journal of Industrial Economics*, Volume LIII June 2005, No. 2.

Joskow, P.L. (2005): *Regulation of Natural Monopolies.* Massachusetts Institute of Technology. Prepared for the Handbook of Law and Economics, A.M. Polinsky and S. Shavell, editors. Elsevier, B.V.

Laffont, J.J. and J. Tirole (1993): *A Theory of Incentives in Procurement and Regulation.* Cambridge, Massachusetts: MIT Press.

Schweppe, Fred C., Caramanis, Michael C., Tabors, Richard D. and Bohn, Roger E. (1988): *Spot pricing of electricity.* Boston/Dordrecht/London: Kluwer Academic Publishers.

Stoft, S., C. Webber and R. Wiser (1997): *Transmission Pricing and Renewables: Issues, Options and Recommendations*. Environmental Energy Tecnologies Division, Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley.