



**EU–Russia Cooperation Programme
(Tacis)**

Further information can be obtained from the following:

Contracting authority

European Union — Delegation of the European Commission in Russia
Mario Ronconi
Kadashevskaya nab. 14/1
Moscow 109017
Russian Federation
Tel: +7 095 721 2033
Fax: +7 095 721 2040
E-mail: Mario.Ronconi@cec.eu.int
Website: www.eur.ru

Contractor

KEMA Consulting
Jan Lefers
Utrechtseweg 310
P.O. Box 9035
6800ET Arnhem
The Netherlands
Tel: +31 26 356 2588
Fax: +31 26 442 8713
E-mail: jan.lefers@kema.com
Website: www.kema.nl

**Project ERUS 9902
Rationalisation of the Federal
Wholesale Electricity and Power
Market — "FOREM"**

**Transmission Pricing
and Transmission Price Control**

Konstantin Petrov

August 2004



This project is funded
by the European Union



This project is implemented
by KEMA Consulting

**Project ERUS 9902
Rationalisation of the Federal
Wholesale Electricity and Power
Market – "FOREM"**

***Transmission Pricing
and Transmission Price Control***

Konstantin Petrov

August 2004

What is Tacis?

The Tacis Programme is a European Union initiative for Eastern Europe, the Caucasus and Central Asia, which fosters the development of harmonious and prosperous economic and political links between the European Union and these partner countries. Its aim is to support the partner countries' initiatives to develop societies based on political freedoms and economic prosperity.

Tacis does this by providing grant finance for know-how to support the process of transition to market economies and democratic societies.

Since 1991, Tacis has granted over 2.46 billion euro for the implementation of more than 1,500 projects in the Russian Federation.

Tacis works closely with the partner countries to determine how funds should be spent. This ensures that Tacis funding is relevant to each country's own reform policies and priorities. As part of a broader international effort, Tacis also works closely with other donors and international organizations.

Tacis provides know-how from a wide range of public and private organizations, which allows experience of market economies and democracies to be combined with local knowledge and skills. This know-how is delivered by providing policy advice, consultancy teams, studies and training, by developing and reforming legal and regulatory frameworks, institutions and organizations, and by setting up partnerships, networks, twinnings and pilot projects. Tacis is also a catalyst, unlocking funds from major lenders and investors by providing pre-investment and feasibility studies.

Tacis promotes understanding and appreciation of democracy and a market-oriented social and economic system by cultivating links and lasting relationships between organizations in the partner countries and their counterparts in the European Union.

Table of Contents

1	Preface	5
2	Executive summary	8
3	Introduction	10
4	Status quo	12
4.1	Physical aspects	12
4.2	Commercial aspects	13
5	Transmission network functions	14
5.1	Transmission functions	14
5.2	Transmission service elements	15
6	Pricing principles	17
6.1	Economic efficiency	17
6.2	Cost recovery	18
6.3	Efficient regulation	18
6.4	Simplicity and transparency	19
6.5	Non-discrimination	19
6.6	Trade-off of principles	20
7	Review of the existing pricing arrangements	21
7.1	Network and connection service	21
7.2	Asset valuation and regulatory asset base	21
7.3	Profit allowances and cost of capital	21
7.4	Postage stamp versus locational pricing	23
7.5	Long-term versus short-term locational price signals	23
7.6	Pricing concept	24
7.7	Cost cascading and tariff design	24
7.8	Payment liability	26
7.9	Price control	26

8	Use of network charges	27
8.1	Flat versus locational charges	27
8.2	Marginal versus average cost pricing	27
8.3	Practical solutions	29
8.4	Revenue requirements and cost recovery	37
9	Connection charges	54
9.1	Shallow connection charges	54
9.2	Deep connection charges	56
10	Price control	58
10.1	Rate of return regulation	58
10.2	Yardstick regulation	59
10.3	Performance-based regulation	60
10.4	Cap regulation	61
10.5	Price cap regulation	62
10.6	Revenue cap	65
10.7	Scope of the regulatory caps	70
10.8	Quality of supply	71
11	Recommendations	73
11.1	General design of transmission service charges	73
11.2	Revenue requirements	73
11.3	Transmission use of network charge	74
11.4	Transmission connection charge	75
11.5	Transmission price control	76
11.6	Transmission capacity rights	78
12	List of acronyms and definitions	79
13	References	80
14	Acknowledgements	83
15	Author's biography	84
16	End notes	85

1 Preface

The Project ERUS 9902 — Rationalisation of the Federal Wholesale Electricity and Power Market — 'FOREM' was financed by the Tacis programme and was undertaken by KEMA Consulting.

The overall objective of the project is to support the on-going restructuring and liberalisation process in the electricity sector in Russia with a focus on the Federal Network Company (FNC) and the interfacing problems with the other relevant market participants like System Operator (SO), Administrator of Trade System (ATS), etc. RAO UES of Russia is the main Russian partner in the project.

Recommendations have been given on:

1. Commercial Metering Systems and Metering Rules

Metering equipment to enable an electricity market to function must be accurate, reliable and installed in appropriate locations. A detailed metering analysis was carried out on sample substations and power stations of the Middle Volga region and although the present arrangements conform to the existing rules and regulations, they did not appear to be adequate for the market to operate correctly.

A number of recommendations were developed covering the responsibility for metering, documentation, accuracy, backup metering, delivery point, data collection, metering providers, derogations, data access and metering costs. These recommendations were used in the production of metering rules. The purpose of the metering rules is to specify the minimum mandatory requirements for metering to provide the data for the operation of the market.

2. Improvement of the regulatory regime

Technological Rules (technical codes)

The Electricity Law determines the basic market principles and basic technological rules. However, detailed technological and market rules need to be developed and accepted to govern the interfaces between the different players on the market and to obtain a transparent electricity market. One of the major obstacles is to find mutual consensus of the market and technological rules by the different stakeholders.

Technological rules are necessary to enable network companies to provide a reliable service in an unbundled industry and liberalised power market. The technological rules also allow for non-discriminatory and transparent access to the network. The existing technological rules have been reviewed and several improvements have been identified with respect to precise rules for safety co-ordination, rules for nomenclature, general conditions, rules for system tests, rules for contingency plans, rules for operation liaison and event reporting, data registration code.

Transmission losses, transmission pricing and price regulation

The treatment of transmission losses, in particular pricing and allocation, is one of the most controversial and difficult elements of the restructuring process. A review has been prepared of the present situation in Russia concerning transmission losses (loss classification, time and locational loss dynamics, loss valuation methods, loss reduction mechanisms), transmission pricing and transmission price regulation (assets valuation, establishment of regulatory assets base, regulatory depreciation policy, cost of capital, revenue requirements and efficiency assessments, regulatory price control, cost allocation and tariff design). Recommendations for improvements have been elaborated.

Ad-hoc advice and support

The objective of this subtask was to provide ad-hoc advice and support to the reform process whenever necessary and appropriate. Several tasks have been carried out such as advice on ancillary services, reliability of power supply in liberalised markets, costs of restructuring, background information on international developments of market designs in Asia, Australia and USA, and international treatment of losses.

3. Recommendations of Consumers' access to the market

The aim of open access is to enable producers and consumers to transport electricity through the networks and by doing so, enable the introduction of competition in generation and supply. A precondition to open access regulation is that certain categories of consumers have the right to choose their supplier.

The Russian law regulates consumers' access to the wholesale market and the commercial rules provide criteria for eligible consumers. The law distinguishes between the wholesale and the retail market. While criteria for

wholesale access have been developed, the rules for access on the retail market are not clear.

Recommendations have been given to the retail access arrangements such as:

- eligibility criteria for the retail access;
- information management to handle supplier switching and supporting IT-systems to ensure proper settlement; and
- load profiling that should substitute the demand metering for small consumers.

4. Supply of a Network Configuration Information System

The Network Configuration Information System (NCIS) and related provision of the appropriate hardware and software is to be supplied to the FNC in frame of this project and shall contain data regarding properties of equipment in the electrical network (substations and transmission lines), description of topology of power network facilities and power plants and equivalent data of elements and diagrams of the Russian power grid.

This booklet discusses a number of conceptual and practical issues associated with the design and implementation of transmission pricing and price control regimes in the wholesale market and suggests suitable forms for their application in Russia.

2 Executive Summary

Transmission service charges will apply to the networks owned by the FNC, i.e., for 220 kV lines and higher. Transmission service charges should be made up of transmission connection charge and transmission use of network charge. Short-run signals (in terms of time and location) will be provided by application of nodal pricing on the wholesale market. Transmission loss rentals and congestion rentals (surpluses) resulting from the application of nodal pricing should be used to scale down the transmission use of network charges.*

The transmission connection charge should cover the cost of transmission connection assets (shallow concept) without taking into consideration any reinforcement requirement in the transmission network infrastructure.

The transmission use of network charge will apply to the transmission network infrastructure reflecting its ability to transport energy in a reliable way and will recover the costs incurred by the FNC in order to ensure secure operation and regular maintenance of the network assets (costs related to operation and maintenance (O&M), depreciation and financial return on the core transmission network assets), however, reduced by the surplus resulting from the application of nodal pricing. The transmission use of network charge will be set irrespective of the distance transported but taking into account the entry and exit points on the respective transmission regions and network voltage level. As nodal prices will provide locational signals, imposing secondary signals through locational pricing for the transmission network infrastructure is not of immediate priority. Hence, it is suggested to allocate the transmission infrastructure costs (revenue requirements) of each transmission region to the respective voltage level operated by the FNC in this region using the postage stamp approach and cost cascading among the voltage levels.

If secondary locational long-run signals are still preferred by the Federal Energy Commission (FEC) and/or the FNC, the short-run signals resulting from the application of nodal pricing can be supplemented by locational pricing (e.g., on a zonal basis) for the transmission network infrastructure. In this case, a portion of the revenue requirements should be allocated on a locational basis to the transmission network nodes and aggregated in price zones. It is recommended to allocate only a portion of the total revenue requirements on a locational basis (e.g., 50%) and to pro-rate the remaining part.

* In Russia, currently the loss rentals and congestion rentals are not considered in the transmission pricing regime, but are directly reimbursed to the final consumers through the wholesale market prices.

Two options are suggested for design of the transmission use of network charges: (1) peak-based regime, where the peak usage is appropriately defined so as to minimise the attractiveness of inefficient charge avoidance (e.g., number of peak hours or peak time period); (2) hybrid tariff structure using energy, demand and fixed charges. The definition of peak usage for generators could refer to their rated capacity (maximum net capacity that can be delivered to the transmission network under normal conditions of operation). The definition of peak usage for load entities could refer to their observed load during the settlement period in which system peak load occurs and some number of settlement periods around the peak load or the connected/contracted load.

The transmission price control must be structured in a way that does not discourage certain sources of finance and investment. In order to establish efficiency increase incentives, it is suggested to apply a price control based on revenue cap. Under revenue cap regulation, a restriction is imposed on the growth rate of the regulated revenue. The formula includes indexes of some inflation indicator (e.g., the retail price index (RPI) or the consumer price index (CPI)). In addition, regulated service providers are obliged to reduce their revenues each period according to some assumed productivity growth rate that is to be determined by the regulator.

Cap regulation creates strong incentives for efficiency improvement as any cost saving directly translates into higher profits. However, both theory and practice suggest that, without additional quality regulation measures, these incentives eventually lead to perverse quality degradation. Therefore the FEC should require annual reporting from the FNC on all transmission system incidents that involve loss of supply to customers (including downstream distribution companies). Because such failures will be infrequent, reporting can relate to each individual incident. In addition, the FEC might require that FNC should report on other quality factors for its network, even though these do not relate to loss of supplies. These include the extent to which parts of the transmission network are unavailable for service during the reporting year and the reasons for such unavailability.

As the cap regulation is a new form of price control for Russia, it is recommended to start with a shorter regulatory period, e.g., three years. This will give the FEC the opportunity to balance the incentives and the uncertainties resulting from the lack of experience and reporting of data.

3 Introduction

RAO UES of Russia is in the process of major reorganization and restructuring. The following entities in the new power industry organisation can be identified:

- System Operator (SO)
- Administrator of Trade System (ATS)
- Federal Network Company (FNC)
- Generation Companies (GenCo)
- Regional Network Companies (RNC)
- Supply Companies (SupCo)

The first three entities SO, ATS and FNC are the central entities, where the ATS is responsible for operating the power market, SO is responsible for safe and reliable operation of the power system and FNC is the owner of the transmission system. The Generating Companies are the transmission network users together with the large industrial consumers that are directly connected to the transmission network. The Federal Network Company and the Regional Network Companies are the companies responsible for the 'wires business.'

This structure is illustrated in Figure 1.

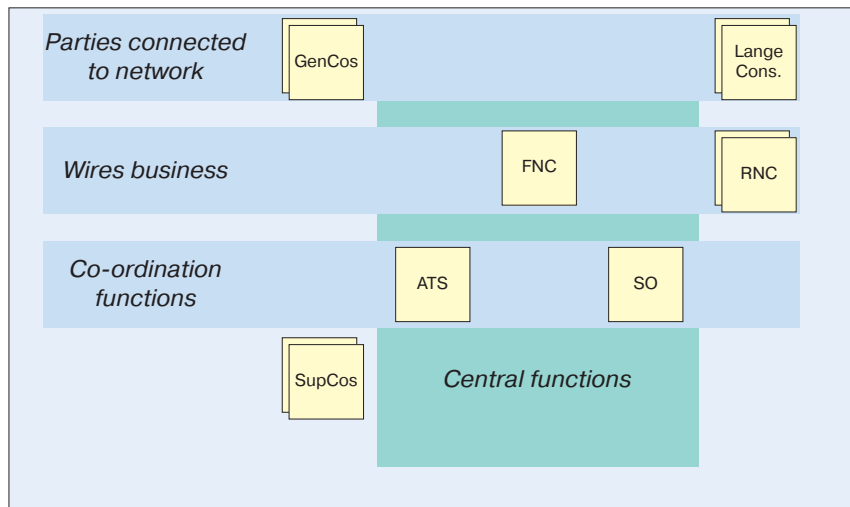


Figure 1. Overview of new structure

From Figure 1, it can be seen that the Russian Federation has adopted a model in which the three central functions are allocated to three different entities. Internationally, this model is known as the 'ISO model.' The ISO model is characterised by the largest extent of unbundling, which means that extra interfaces are created and that more discussion will be needed to lay down the detailed responsibilities of the different entities. These detailed responsibilities and the procedures at the 'interface' between the different entities are normally laid down in technical codes (e.g., Grid Code and Metering Code) and market rules. In the case of the Russian Federation, such documents are called Technological Rules and Commercial Rules.

In the deregulated electricity world, the transportation functions such as transmission and distribution as well as dispatch are still attributed "natural monopoly" features. The transmission system is essential to the electricity sector, and effective physical access to that system has been the key to the development of the sector to date. Similarly, effective and equitable economic access to that system is fundamental to the development of an efficient and competitive electricity market, with pricing policies playing a key role in the provision of effective economic access.

Transmission pricing is one of the most controversial and difficult elements of the restructuring process and has attracted considerable international attention, as new electricity market forms are developed around the world. Thus, the goal of transmission pricing is not merely to recover the costs of the transmission system, but to devise a pricing framework that contributes appropriately to fulfilment of the larger goal of developing an economically efficient electricity sector.

This booklet discusses a number of conceptual and practical issues associated with the design and implementation of transmission pricing and price control regimes and suggests suitable forms for their application in Russia.

4 Status Quo

4.1 Physical Aspects

Six large regional power pools operate synchronously in Russia: namely, the Regional Power Pools (RPP) of Centre, Middle Volga, Urals, North-West, North Caucasus and Siberia.¹ The only regional power pool on the territory of Russia that operates separately is the regional power pool of Far East. A 220 kV transmission line connects the regional power pools (RPP) of Far East to the RPP of Siberia and is utilised depending on the demand/supply in these two power systems. The power system "Yantarenergo" in the Kaliningradskaya province is isolated territorially, however, operates synchronously with the Russian power system.

The Russian power system operates synchronously with the power systems of the Baltic States, Belarus, Trans-Caucasus and Kazakhstan. Further, the Russian power system is connected to the power systems of the former Soviet Republics (now the Independent States) as well as to the power systems of Finland, Norway, Mongolia and China.

At the year of 2001, the installed capacity of the power sector in Russia constituted in total 215.0 GW and electricity production was 863 TWh. In terms of production technologies, 21.3 GW of electricity was generated in nuclear power stations (approximately 10%), 44.3 GW in hydropower plants (approximately 20%) and 149.4 GW in thermal power plants (approximately 70%). Overall, 78 power utilities (AO-energос) operate on the territory of Russia and 69 of them are part of the Russian power system.

The FNC is an open joint stock company, that maintains and develops the unified national (all-Russian) electric transmission network (UNEN). UNEN is a complex of electric network units, belonging on property rights or on other legal grounds to power sector subjects, including substations and transmission lines 330-750 kV and the majority of 220 V, which is a background of functioning of the wholesale market. In accordance with the Russian legislation, all network service providers (FNC and AO-energос) are obliged to get licences for provision of network services including connection and transmission and distribution of electricity through their networks. The transmission network service is a monopoly service. Therefore, the government regulates the activities related to provision of transmission network service in terms of pricing and quality of service. In accordance with the federal law "On Electricity," functioning of FNC and SO is to follow

their common agreement, signed by participation of authorities. An important part of such agreement is to be the problem of control of network subjects. The SO leading operational regimes and conditions is responsible for ensuring the secure operation of the electricity transmission network in Russia.

A fundamental function of the FNC is the transmission network development that reflects the supply/demand balance and is carried out in the most economic and efficient way. The construction of new transmission lines should relieve transmission network constraints and increase its transmission capacity. One part of financial funds for the transmission network development is to be included in the network price (investment part of the tariff), the other part of financing FNC is to get from external investors.

4.2 Commercial Aspects

The new wholesale market design in Russia applies nodal pricing and internalises the short-run marginal cost (SRMC) of transmission (cost of transmission losses and transmission constraints). The nodal pricing regime manages congestion and sets prices for transmission losses and transmission constraints through a centralised energy market based on economic dispatch.² The market operator minimises the total system's gain from trade (demand bids and generation offers) subject to transmission and reliability constraints. The price at each node is then set to the incremental offer or bid price of the most expansive unit generated or consumed at that node. These nodal prices become the hourly prices charged to load entities and paid to generators at the respective nodes.

Order N 72-э/3 is the document that deals with transmission pricing. Order N 72-э/3 applies a network service model (point-of-connection model) that allocates the transmission network cost to the connected transmission service users rather than to individual transactions. Each transmission region is priced separately and voltage levels are defined as separate cost pools. Once allocated to the voltage levels, costs of each regional transmission network are cascaded down in order to derive the network cost portion attributable to the customers connected to the respective voltage level. No differentiation by customer types applies. The cost cascading reflects the total customer load connected to the respective voltage level and the transmission use of network charges are denominated in RUB per kW (demand charges).

5 Transmission Network Functions

Transmission service is associated with the ability of the FNC to transport energy via the transmission network in a reliable and efficient way (including network planning and network construction) and to ensure secure and efficient and regular maintenance of the transmission network assets as required in the Technological Rules and the relevant Russian technical standards.

In the new market design in Russia, the transmission function is attributed monopoly features while other functions such as generation and wholesale sales are considered as competitive services.³ These characteristics of network infrastructure seriously limit the extent to which market forces can be introduced in the provision of network services, and in particular investment. Decentralised investment would result in network investment that is insufficient or non-optimal from a wider public policy perspective. In practice, this means that the costs of network O&M and network infrastructure investments will continue to have to be recouped via regulated network charges. The FEC (recently reorganized into Federal Service for Tariffs) is responsible for transmission price regulation.

5.1 Transmission Functions

Transmission networks provide the following services:

- physical transport;
- network security; and
- connection.

The first service addresses the transport of electricity over distances from generation to load centres. For the generators, the ability to transport power to loads is a highly important value of the transmission network. The generators could choose a location near to the customer, thereby lowering the costs incurred in transporting electricity. The same is valid for load entities on the bulk level (e.g., large industrial consumers) that may choose a location geographically nearer to generation sources. The second service is related to the network security and requires the transmission service provider (FNC) to keep additional transmission capacity (over and above that required for the simple transport of energy) in order to be able to cope with losses of circuits due to outages on the transmission network and to enable demand to be met by the available generation capacity under most

circumstances. Transmission connection service is associated with the ability of the transmission service provider (FNC) to provide reliable connection to the transmission network and regular maintenance of the transmission connection assets.

5.2 Transmission Service Elements

Taking into consideration the major transmission functions, one can distinguish the following five elements of transmission service pricing:

- (1) provision of core transmission network facilities (frequently referred to as transmission network service including transport and security);
- (2) provision of transmission connection facilities (frequently referred to as transmission connection service);
- (3) provision of transmission congestion management to relieve transmission constraints;
- (4) provision of energy to cover transmission losses; and
- (5) provision of system services to keep the system balance and to restore this balance after disturbances.

While the first two elements are "must" elements of the transmission service, the latter three elements are of an optional nature and can be excluded. One reason is that the costs resulting from the performance of congestion management or provision of ancillary services⁴ are not directly associated with the network infrastructure costs (O&M cost, depreciation and return on assets) but rather with specific actions of SO and network company. Moreover, including these services in the transmission pricing depends strongly on the specific industry and market model (e.g., wholesale market design, the relationships between the SO and the transmission asset owner, etc.).

The international experience demonstrates a large variety of concepts and solutions. In the old England and Wales Pool, congestion management costs and provision of ancillary services costs were covered via a pool-uplift and were not considered in the transmission pricing. In the Netherlands and Germany, transmission losses should be procured by the transmission SO and are included in transmission charges. In New Zealand and Singapore, transmission losses are handled on the wholesale electricity market and not through transmission pricing.

In Russia, system services, transmission constraints and transmission losses (in particular the load related losses) are dealt with outside the transmission pricing. The new wholesale market design in Russia applies

nodal pricing and internalises the SRMC of transmission, namely, cost of transmission losses and transmission constraints.⁵ The nodal pricing regime manages congestion and sets prices for transmission losses and transmission constraints through a centralised energy market based on economic dispatch.⁶ System service costs will be probably covered through an extra system service charge.⁷

6 Pricing Principles

The transmission system is fundamental to the electricity sector and effective physical access to that system has been the key to the development of the sector to date. Similarly, effective and equitable economic access to that system will be fundamental to the development of an efficient and competitive electricity market in Russia, with pricing policies playing a key role in the provision of effective economic access.

Transmission pricing has attracted considerable international attention, and some controversy, as new electricity market forms are developed around the world. Thus, the goal of transmission pricing is not merely to recover the costs of the transmission system, but to devise a pricing framework that contributes appropriately to fulfilment of the larger goal of developing an economically efficient electricity sector. In particular, given the new market design concept driving the current reform process in Russia, this framework must facilitate fair competition, and give appropriate signals for operational and investment decisions by all parties.

Principles to be followed are listed below and described in more detail in the subsequent sections:

- Economic Efficiency;
- Cost Recovery;
- Efficient Regulation;
- Simplicity and Transparency; and
- Non-Discrimination.

6.1 Economic Efficiency

An effective charging structure will signal to users the costs that they impose on the transmission network and encourage the FNC to utilise its assets optimally (both day to day and longer term with respect to maintenance scheduling and network expansion).

The structure of transmission pricing should encourage economic efficiency by providing economic signals to users for an optimum use of the network, and to the network operator for an optimal operation and expansion of existing assets. While economic efficiency has several dimensions, the most important factor with regards to tariff structures is referred to as allocative efficiency, i.e., the optimised use of the transmission system by individual users.⁸ This implies that prices should

reflect economic costs by signalling the cost of operation and future investments while also reflecting the sunk cost associated with the already existing infrastructure.

By providing appropriate signals to individual users of the incremental cost from their use of the system, profit-maximizing users will behave in a manner that maximizes the net benefits and minimises the cost of the system as a whole. This relates to both short-term operating and consumption (e.g., dispatch, load management) as well as long-term investment (e.g., new plants, plant shut-down) decisions.

6.2 Cost Recovery

Investments in networks are characterised by economies of scale and scope and are lumpy, so that "overbuilding" is often the least cost option. These characteristics of network infrastructure seriously limit the extent to which "market forces" can be introduced in the provision of network services, and in particular investment. Decentralised investment would result in network investment that is insufficient or non-optimal from a wider public policy perspective.

In particular, transmission grids are highly capital-intensive and capital-related costs form the overwhelming part of transmission costs. The economies of scales which are endemic of the transmission technology imply that the SRMC pricing will usually not be sufficient for some grid owning utility to fully cover its cost due to a large fraction of fixed cost.⁹ Therefore the pricing regime should be structured in such a way that the FNC can earn sufficient revenue to cover their efficient O&M and capital costs including reasonable rate of return. The move away from charging only for marginal losses and network congestion entails a move from first to second best pricing which must meet a total revenue requirement. While efficient prices should reflect the costs of the service being provided, cost attribution is problematic in shared electricity networks, because of economies of scale and scope.

6.3 Efficient Regulation

The transmission network will remain a monopoly in the restructured Russian power system and will, therefore, be regulated by the state. The pricing methodology should reflect this need for regulation and should encourage efficient operation by the FNC while keeping a light regulatory burden so far as it is possible but also preferred. The FEC should recognise

the importance to the FNC of recovering sufficient level of O&M and capital costs. However, it is important that the FNC does not incur excessive or unnecessary costs in providing transmission service.

6.4 Simplicity and Transparency

It is important that the pricing rules are clearly understood. Transmission charges should be simple and transparent so that a user can readily determine the charges it will face. Excessively sophisticated approaches may, on first sight, appear to promote efficiency but they will likely appear as some kind of "black box" to users. Under such circumstances, users may not be able to respond to the corresponding economic signals. While the network operator can (and should) make its tariff model and methodology available to all users, they will generally best respond to signals if the effects of user's decisions on its charges are reasonably intuitive. In other words, transmission charges should be predictable.

The primary aim of simplicity and transparency is also related to a number of secondary objectives. First, long-term investment decisions will only be based on any economic signals if the pricing regime is stable over a sufficient period of time. Secondly, a simpler tariff structure will reduce the administrative burden and thus ease implementation. Furthermore, to avoid disputes, the tariff regime needs to be clear and should be based on explicit rules as far as possible. Finally, transparency is an important prerequisite for general acceptance by users and the general public.

6.5 Non-Discrimination

A key element of transmission pricing is the requirement to ensure that a level playing field is created for all transmission service users. This requires the notion of treating all transmission service users equally, irrespective of size, ownership or other factors, i.e., non-discrimination between transmission service users.

In practice, for transmission pricing, this means that all transmission service users should face the same methodology for calculating charges. It does not necessarily mean that all transmission service users will face the same payments, e.g., a large power plant requiring a large and complex connection will not pay the same absolute charge as a smaller plant. They should both, however, have their charges calculated in the same manner.

6.6 Trade-Off of Principles

The application of all pricing principles should be considered in the light of their interdependencies and not with mechanistic prioritising of one of them. There is no universal answer as to how these principles should be ranked in terms of importance. Efficiency and cost recovery are key objectives in transmission pricing but must be balanced against other objectives, including non-discrimination against the market participants, simplicity in implementation and transparency to users, stability to support long-term investment decisions and flexibility to support changing market environments. These objectives represent key criteria for judging the appropriateness of various charges for achieving efficiency and cost recovery.

7 Review of the Existing Pricing Arrangements

7.1 Network and Connection Service

Order N 72-ə/3 provides detailed discussion on the provision of transmission network service. However, little attention seems to have been paid to the transmission connection service. Separate treatment for transmission network and transmission connection service follows economic causality principles, namely to charge the costs to the party that has caused the costs. The pricing regime should follow the classification of transmission service elements and, therefore, distinguish between transmission use of network charges and transmission connection charges.

7.2 Asset Valuation and Regulatory Asset Base

There are not clear statements with respect to the asset valuation and the establishment of regulatory assets base. Arguments about asset valuation vary from insisting ownership rights be recognised, to questioning whether any value should be attributed to sunken investment. The regulators usually endorse particular asset valuation methodology and the asset valuation issues must be considered with regard to the functional adequacy of regulated assets, market assets value, overall profitability of the regulated business and sustainable cash flows of the business as well as equity considerations. The objective of encouraging continuing investment in the regulated transmission service will require the FNC to be provided with an assurance that it will earn a reasonable (risk adjusted) return on its transmission assets. Therefore the treatment of assets and how their value should be adjusted over time reflecting the investment process (but also depreciations) is a crucial issue.

7.3 Profit Allowances and Cost of Capital

Order N 72-ə/3 suggests profit allowance for the regulated transmission service provider that includes three major components:

- (1) for networks of 500 kV and above, the net profit (profit after tax) plus depreciations should be sufficient to fund planned investments;

- (2) for networks of 330 kV and below, the net profit (profit after tax) plus depreciations should be sufficient to cover the debt service (including debt interests and principle); and
- (3) a non-investment profit term that should suffice to cover dividends, reserves, social charges, etc.

Profits belong to shareholders and could be distributed or retained and used for financing of new investments, i.e., it is a decision of shareholders how to finance new investments. Establishment of predetermined liaisons between investment fund requirements and profit (dividend) distribution policy does not seem to be appropriate as it delegates to the regulator functions that naturally belong to the companies' management. The regulatory regime should specify a return on equity (as an element of Weighted Average Cost of Capital (WACC)) and consider taxes either as converting the after-tax in pre-tax rate of return or through explicit taxation allowances in the revenue requirements and after-tax rate of return. The decision on profit (dividend) distribution policy remains in the hands of the company's management that is compatible with the so-called "light-handed" regulatory policy.

The profit mechanism for networks of 500 kV and above implies the entire internal financing of all network investments. The profit is treated as a residual funding source, resulting from the difference between the investment plans and available cumulated depreciation volumes. Differently, the mechanism for networks of 330 kV and below calculates the profit level necessary to cover the debt service requirement resulting from debt funding of investments (net of depreciation). Presumably, this approach strives to ensure reliable funding arrangements for network investments through funds generated internally on the basis of regulatory profit allowances. While this might be desirable from the perspective of keeping a reliable transmission network and sufficient level of investments, such an approach may easily lead to perverse incentives to the financial management and financial discipline of the FNC, e.g., excessive spending of depreciation funds and increasing the requirements to residual profit component (at the time point of assets replacement).

Ideally, investments in asset replacement should be financed through cumulated depreciations (valued at replacement cost to catch up asset price increase resulting from inflation) and investments in network extension should be funded by new external financing sources. Funding investments through profits that constitute part of the revenue requirements means that customers are forced to participate in the financing process. Once constructed, however, the assets should be depreciated over their life and the annual depreciation allowances should be included in the transmission network charges. Hence, the approach applied by Order N 72-э/3

implies double counting and charges transmission service users twice: via the residual profit component and the depreciation allowance.

There are no incentives for minimising cost of capital, i.e., optimising capital structure as the Order N 72-э/3 determines residually the required profit level and limits the decision freedom of the companies' financial management with respect to their capital structure. It is understandable that the FEC strives to ensure financial viability of the regulated network service providers and reliable service delivery. However, in keeping with an incentive-based approach, FEC may need to establish forward-looking benchmarks for calculation of the WACC based on the yardstick of an efficiently financed business in Russia. This would involve making assumptions regarding a normal, efficient capital structure, independent of the specific ownership and financing arrangements for the regulated network service provider. Accordingly, the WACC should be calculated on the basis of an efficient standard equity/debt ratio.

7.4 Postage Stamp Versus Locational Pricing

Order N 72-э/3 applies a network service model (point-of-connection model) that allocates the transmission network cost to the connected transmission service users rather than to individual transactions. Order N 72-э/3 suggests cost allocation based on a postage stamp approach differentiated by transmission region and voltage level. Postage stamp tariffs have been applied in a number of countries with a varying level of inclusion of cost components. Depending on the size of the transmission network area, this approach might provide no incentives for proper location, since postage stamps will not differ in dependence of zone of connection nor of the distance involved in a transaction.

7.5 Long-Term Versus Short-Term Locational Price Signals

The new wholesale market design in Russia establishes short-term market signals through different prices at the transmission network nodes. The nodal pricing regime manages congestion and sets prices for transmission losses and transmission constraints through a centralised energy market based on economic dispatch. Additionally, long-run signals could be established by locational differentiation of the transmission network cost across the system. The transmission service users pay transmission use of network charges only for their connection point to the transmission net-

work, irrespective of the transaction concluded. This transmission use of network charge, however, will differ across the system, depending on where the transmission service user is connected to the transmission network.

7.6 Pricing Concept

Order N 72-э/3 applies average cost pricing. Under considerations of cost recovery, the proposed average cost pricing may appear preferable. The average cost pricing reflects the cost of the existing transmission network assets. Such a pricing scheme, however, could fail to achieve allocation efficiency. In periods with excessive network capacity, transmission use of network charges would be inefficiently high, while in periods when the transmission network is heavily loaded prices would be inefficiently low. This may lead to an excessive network utilization that would make an inefficient extension of the network necessary. An alternative solution could be to apply methods based on long-run marginal cost (LRMC). The purpose of transmission pricing methodologies based on LRMC is to signal to the transmission service users the costs of network expansion due to a marginal (incremental) increase in generation or demand.

7.7 Cost Cascading and Tariff Design

Order N 72-э/3 cascades the transmission cost onto network voltage levels in proportion to the total connected load. Further, Order N 72-э/3 denominates transmission use of network charges in RUB per kW (demand charges). The arguments in favour of peak-based charging and/or cost allocation methods arise because transmission networks are dimensioned on the basis of reliably serving peak system demand and most of their costs are fixed. The exact form of a demand charge can vary. The peak demand can be defined as the single highest period of demand or supply over a number of years, within each year, or the average demand/supply over a small number of periods within the year. Alternatively, the peak demand can be defined as the contribution to system peak demand or regional peak demand, the customer's own peak, connected load or the peak for each supply point servicing a customer.

Simple aggregation of connected customer load for the purposes of cost allocation as suggested by Order N 72-э/3, however, may fail to reflect the

actual use of the transmission network and the coincidence of the time occurrence of peak demand at each voltage level.

Using only the connected customer load but not the measured demand would decouple completely the chargeable basis (connected customer load) from the actual usage of transmission network service (measured demand) and entirely ignore the real time load flow and actual use of network. Therefore, sometimes combinations between measured demand and connected customer load are used (e.g., in the Netherlands). Further, imposing demand charges on the basis of a single measured demand value would make the total payment for transmission use of network charges dependent solely on this demand value and would expose the transmission service users to stochastic demand fluctuations. In order to avoid these adverse effects, additional chargeable demand values could be considered for the purposes of payments for transmission use of network charges.

Pure demand charges (even using more than one peak period to determine the payment liabilities) could create substantial differences for the payment of transmission service users having different load profiles. The effective payments by transmission service user with higher load factors could be significantly lower than those payments by transmission service users with lower load factors. In contrast to demand charges, an energy charge encourages under-consumption of electricity but is likely to have some positive impact on the energy market.

Energy charges might be desirable to meet some equity objectives as large customers pay a greater share of energy charge revenues than smaller customers. Energy charges achieve these equity objectives and are less likely to result in distortions in the energy market, since relative energy costs in different demand periods are not distorted. Energy charges might be more appropriate than demand charges, where there is excess network capacity and high demand charges would provide a perverse economic signal to restrict the economic usage, whereas on the contrary higher network utilisation should be encouraged.

Finally, fixed charges — appropriately applied — are less distortional than variable charges, are simple to apply, could be differentiated by size of customer, and would reduce the taxation impact on consumption. Fixed charges are the first best solution for sunken cost recovery under marginal cost pricing. However, imposing high fixed charges without any link to demand or energy will certainly face significant opposition from the customers.

7.8 Payment Liability

Order N 72-э/3 allocates the payment liability solely to load entities. Secure transmission network benefits both load entities and generators. Without a reliable transmission network, the generators (connected to the transmission network) are not able to feed in electricity into the transmission network. On the other hand, reliability is also important for load entities. A measurement for the value of reliability for load entities is the value of lost load. It could be argued that the foregone revenues of generators not allowed to generate electricity due to an unreliable transmission network are lower in relation to the value of lost load. However, such arguments would not change the principle that a fair allocation of payment liability will need to consider generation and load entities.

7.9 Price Control

The major ideas of Order N 226 follow the logic of rate of return regulation. The document leaves open options for application of alternative methods; however, it does not elaborate further details concerning these alternative methods. Lack of transparency and clarity on timing and conditions for application of alternatives might increase significantly the regulatory risk in the Russian power sector. Under conventional rate of return regulation rates of regulated service, providers are reviewed on a regular basis and have to be adjusted (process called "regulatory review") to lower levels if cost savings have been achieved in the interim (since the last review). Regulated service providers thus only benefit from cost savings to the extent that regulatory reaction to cost savings is lagged (this lagged reaction is called regulatory lag). If the regulatory lag is short — one or two years — incentives for cost savings are suppressed almost completely. In such a framework, regulated service providers are said to operate dynamically inefficiently. Additionally, the rate of return regulation creates overcapitalization (overinvestment in fixed assets) incentives and potentials for increase of capital costs. More advanced forms of price regulation (incentive regulation) have been developed and applied in other countries. These methods of regulation decouple totally or partially the allowed revenue/prices from the actual costs of the regulated service provider and provide efficiency increase incentives via retaining the achieved cost saving in the companies for a certain time period.

8 Use of Network Charges

8.1 Flat Versus Locational Charges

The new Russian wholesale market design is based on nodal pricing and internalises in this way the SRMC of transmission (marginal transmission losses and congestion).¹⁰ Short-term market signals are established through different prices at the transmission network nodes.¹¹

Additionally, long-run signals could be established by locational differentiation of transmission network costs across the system. In this case, transmission service users would pay transmission use of network charges only for their connection point to the transmission network, irrespective of the transaction concluded. This transmission use of network charge, however, may be different across the system, depending on where the transmission service user is connected to the transmission network.

Examples of locational pricing for the network infrastructure components could be found in the practice of England & Wales and Australia. The transmission use of network charge in England & Wales reflects the cost of installing, operating and maintaining the transmission system. It is calculated on the basis of the Investment Cost Related Pricing (ICRP) Method developed by National Grid Company (NGC) and using a LRMC pricing concept. Investment Cost Related Pricing builds upon a complete new network alongside existing rights of way and ignores real power flows. Instead, it bases its methodology on the transportation problem, a linear optimisation problem of the operations research literature. The use of system charge presents a zonal fee (14 zones exist), payable according to the location in the grid, rather than to the type of transaction. In Australia, the Cost-Reflective Network Pricing (CRNP) has been developed in preparation of the National Electricity Market (NEM). The CRNP method allocates approximately 50% of the annual revenue requirements of the transmission companies involved on a locational basis. The remainder is allocated on a postage stamp rate basis.

8.2 Marginal Versus Average Cost Pricing

The approach of using marginal cost based prices as signals for efficient network extension attempts to replicate the outcome on the com-

petitive markets. In these markets, producers sell at the competitive market price whenever it is equal to or greater than their SRMC of production. Short-run marginal costs of transmission are defined as the additional costs arising when one additional kWh is demanded and the installed capacity remains constant. Differently, LRMC take into consideration also the transmission network investments when one additional kWh is demanded.

Short-run marginal cost pricing typically implies that the marginal price for a non-congested transmission network is equal to the marginal transmission losses. Whenever the network is congested, SRMC will increase by consideration of opportunity cost. If a power system had a perfect network with no losses, or any kind of constraints limitations, at a given instant in time, the spot prices at all nodes of the system would be the same; i.e., there would be no spatial differentiation of spot prices. In any real network, the effect of losses and network constraints creates differences among the nodes and consumers pay, or generators are paid, different prices, depending on their locations. Those parties should be allowed access that can make most efficient use of the network in periods when network constraints bind. This will typically be the suppliers that can generate/procure electricity at lowest variable cost at the time of transmission constraints and that try to contract with the customers with the highest willingness to pay. The underlying principle of SRMC pricing method is that spot prices between two nodes should not differ for more than the SRMC of transporting the electricity.

An alternative solution could be to apply methods based on LRMC. The purpose of transmission pricing methodologies based on LRMC is to signal to the transmission service users the costs of network expansion due to a marginal (incremental) increase in generation or demand. The difference from the SRMC pricing concept is that in this case, the network assets are not considered static but rather dynamic, i.e., new investment resulting from incremental demand or generation is taken into account.

The average cost pricing reflects the cost of the existing transmission network assets. Under consideration of cost recovery, the average cost pricing may appear preferable. Such a pricing scheme, however, could fail to achieve allocative efficiency: in periods with excessive network capacity, transmission use of network charges would be inefficiently high, while in periods when the transmission network is heavily loaded prices would be inefficiently low. This may lead to an excessive network utilisation that would make inefficient network extensions necessary.

8.3 Practical Solutions

8.3.1 Contract Path

The contract path method requires the nomination of a particular grid supply and receipt point for a bilateral transaction and of a path between these two nodes. It addresses both the amount of contracted capacity to transport as well as the distance associated with the transport over the contract path. Power flows, however, in accordance with the physical laws, taking always the way of least impedance and not along a contract path, thus causing loop flows. The contract path method will thus not represent the physics of the system. It refers to the cost components of the particular contract path. The contract path method can be based on average cost (valued with rolled-in embedded costs or replacement costs) or marginal (incremental) costs. The contract path may allocate a share of the costs of the contract path to the transaction or draws on the incremental costs of providing the transaction.

8.3.2 Postage Stamps

The postage stamp approach represents the opposite to the contract path method. Cost components of a number of a particular asset category or all categories are allocated to all customers on a pro-rata share. The allocation is normally done based on the individual share in the coincidental peak. The approach is used in a number of countries with a varying level of inclusion of cost components.

8.3.3 MW-Miles Method

MW-miles methods could be used for pure bilateral transactions but also for multilateral transactions. To calculate the MW-miles, the power flows over the circuits have to be multiplied with the electrical distance of the circuits. All products are summed up to the overall number of MW-miles transported (total MW-miles transported). In case of a bilateral transaction, two load flow calculations will be necessary, without and with the transaction. The incremental effect of the transaction is estimated as a difference between the product of load flows and the affected lines without and with the analysed transaction (incremental MW-miles). The ratio between the incremental and total MW-miles can be used to allocate the transmission cost to the individual transaction.

A major problem in applying the MW-mile method on multilateral transactions (no direct relationship between a generating unit and a load can be identified) is the allocation of an increment of load at one bus to the increase in generation at the various buses and vice versa for an increment of generation. The MW-mile approach calculates the MW-miles between two nodes for a particular circuit. It does not provide immediately a method or a tool to allocate the MW-miles to a particular node.

8.3.4 Location Dependent Pricing

In the context of the point-to-point service approach, the grid charges according to this solution will reflect the contract path and depend on the distance of electricity transport. In the network service-pricing scheme, locational differentiation of transmission fees is created by different nodal (or zonal) prices across the system. The transmission users pay a transmission fee only for their connection point to the grid, irrespective of the transaction concluded. This transmission fee, however, may be different across the system, depending on where the transmission user is connected to the network. Locational pricing is applied in the form of SRMC or LRMIC pricing.

Short-Run Versus Long-Run Price Signals

To the extent that generators and large customers are less insulated from market signals, it should be recognised that SRMC may provide a relatively stable signal, even in the longer term. Generators and large customers receive some locational signals via nodal prices in the energy market. Other things being equal, the nodal prices are likely to encourage:

- large customers to locate in areas where nodal charges are low, and hence, indirectly, network augmentation costs are reduced; and
- generators to locate near nodes where nodal prices are high and hence where generation at such nodes is likely to reduce losses, congestion and associated network investment.

On the other hand, reliance on pure SRMC charging to signal and recover the costs of the network raises inter-temporal equity issues, given the long life of network assets. Long-run marginal cost charges are likely to be more equitable in this regard. Under full SRMC charging (full nodal pricing), network users would pay for investment via accumulated future losses and congestion rentals. Even if SRMC were sufficient to recoup the costs of investment, it is unclear whether those

parties paying for short-term losses are the same as those benefiting from network expansion in the longer term. Therefore, additional locational signals could be established through locational pricing of transmission network infrastructure. However, one should be aware that a transmission pricing design based on a locational concept is characterised by significant complexity that will naturally limit transparency and practicability of the transmission pricing approach, and ultimately could hazard the successful implementation. Moreover such locational pricing will create geographical price differences (supplementary to the differences resulting from nodal pricing) and may face significant social and political resistance. If the FEC and/or FNC prefer to consider additional locational pricing for transmission infrastructure in order to establish long-term signals it is recommended to use locational pricing only for a portion of revenue requirements. Depending on the methodology applied, this portion will be either ex-ante determined (approach used in Australia) or will result from explicit modelling of LRMIC (approach used in UK).

Below, some methods for locational transmission pricing used in different countries are described, namely:

- Investment Cost Related Pricing applied in England and Wales (zonal pricing based on LRMIC);
- Cost-Reflective Network Pricing applied in Australia (zonal pricing based on LRMIC);
- Pricing based on marginal transmission losses applied in Norway (zonal pricing based on SRMC); and
- Full nodal pricing applied in New Zealand, Singapore (nodal pricing where transmission constraints and losses are handled on the market).

Investment Cost Related Pricing

The use-of-system charge is set by NGC for making available its transmission system for the bulk transportation of electricity. The use-of-system charge reflects the cost of installing, operating and maintaining the transmission system. It is calculated on the basis of the ICRP method developed by NGC and using the LRMIC pricing concept. The charges are levied on both generators and demands but the charge rate varies for each. A rate is determined and set for generation or demand in each area. These prices vary with location and are derived based on an analysis of capacity requirements. In the north of England, significant amounts of coal-fired generation have traditionally been available while most of the demand is in

the south. Therefore, NGC charges so as to discourage generation in the north of England while encouraging demand there. Thus, the generation charge is relatively high while the demand charge is relatively low. Conversely, in the south of England, the generation charges are low, and in some cases negative, so as to encourage generation at locations where the economic incentives would not otherwise encourage generation. In these same locations, demand charges are very high so as to discourage the excess off-takes. This location signal is required because the spot prices in the UK pool do not vary with location, and hence cannot convey a signal as to the relative value of new generation capacity or load in different locations.

Investment Cost Related Pricing builds upon a complete new network alongside existing rights of way and ignores real power flows. Instead, it bases its methodology on the transportation problem, a linear optimisation problem of the operations research literature. The concept of ICRP was introduced by the NGC in England and Wales in 1993/94. It is applied to use of system charges for the transmission network (275–400 kV). Investment Cost Related Pricing represents a quasi-LRMC approach intended to provide locational signals based on the cost of network expansions. The concept comprises three steps and can be summarised as follows:

- (1) In a preliminary step, NGC derives the "expansion constant" as an annuitised value of investments in new transmission capacity per MW and km;¹²
- (2) The second step represents the core of the methodology. Investment Cost Related Pricing assumes the building of a complete new network alongside existing rights of way. In addition, the method ignores real power flows but creates a new set of optimal flows by solving a linear optimisation problem (transportation problem), based on the "triad demand" of each load connection point and the registered capacity for each generator;¹³ and
- (3) The third step determines the incremental cost for 1 MW of additional generation or load at each network node. For this purpose, incremental costs are derived starting from an arbitrarily chosen slack node with an assigned incremental cost of zero. For the sake of simplicity, geographically adjacent nodes with similar incremental cost are finally aggregated into zones.¹⁴

The ICRP thus derives the quasi-LRMC for 1 MW of network capacity at system peak load for each node of the network. The resulting incremental charges are higher (lower) for generation (load) in the north and actually

negative for generation in some southern zones.¹⁵ As revenues from this locational tariff element amount to less than a quarter of total cost, they are supplemented by a uniform postage stamp that is supposed to cover the security cost of the network.

The major advantage of the ICRP is that it is focused on the explicit determination of long-run average incremental costs (an approximation of LRMC) of the network, i.e., the method aims to provide a realistic picture of the LRMC using a direct approach for their calculation. Moreover, the method provides rather stable figures regarding the transmission charges and, thus, does not lead to unnecessary price fluctuations, because the results depend only to a limited degree on the actual system conditions and costs.

On the other hand, the method uses a substantial approximation of the network topology for the purposes of the LRMC derivation. Further, the method opens some contentiousness over data because of the approximate assumptions in the generation dispatch, an increase in demand of 1 MW at node i is assumed to be supplied out of the reference node and for an increase of generation at node i vice versa. The ICRP ignores the physical properties in the existing capacity of the transmission network but does preserve the geography of the system rights of way. Flows are represented by directed flows of material as on a rail, road or pipeline and ignore more or less the differences related to the nature of electric power. Finally, the ICRP explicitly ignores the existence of scale economies (assuming instead that the costs of new construction are totally linear in the length of the line and the megawatts of capacity) and lumpiness (assuming instead that additional transmission capacity can be built in any size increment desired).¹⁶

Cost-Reflective Network Pricing

The CRNP cost allocation is a nodal based method and it requires all of the costs to be attributed to the links ("lines") between the nodes on the network. Some assets are wholly dedicated to providing network service to a participant or group at a single point. These assets can be allocated to the participant directly at the node. Other assets throughout the network are shared among the users and the relative use by each participant must be determined. The shared network costs are allocated amongst users based on the marginal change in network element current flows as a result of an increment of user load at each bus. The generation source for each load is defined using the "electrical distance" as a measure of the capability of a generator to supply each load point.

Using this approach, a greater proportion of load at a particular location is deemed to be supplied by generators that are electrically closer than those that are electrically remote. The generator to load allocation is carried out according to relative fault contributions by each generator to a three-phase fault at each load point.

The following critical points have been pointed out by NECA's¹⁷ review on the CRNP transmission pricing approach. It can lead to perverse pricing signals because it seeks to reflect total, rather than marginal, costs to customers. This means that CRNP takes no account of the level of spare capacity on the system in setting prices. Therefore, if the system is at full capacity, CRNP will produce a lower unit price compared to a situation where there is spare capacity. Such pricing signals are the opposite of those that one would expect to see in a competitive market. Secondly, it allocates costs on the basis of load flows in an attempt to identify the users of particular assets on the system. Some of these load flows may, however, be subject to significant changes on a periodic basis. Such changes can create volatile use of system charges as the cost allocation methodology reallocates charges between customer groups.¹⁸

Pricing Based on Marginal Losses

Both previous methods represent different variations of LRMC pricing. The use of marginal losses, on the other hand, serves as a proxy for SRMC pricing.

Marginal losses are typically calculated for different nodes (or zones) of the network by means of a power flow analysis. In the case of Norway, these losses are then expressed as a percentage of load or generation at each node, the so-called loss coefficient. However, as marginal losses depend on the loading of the network, they may change considerably over time. For this reason, loss coefficients in Norway are differentiated between high- and low-load hours,¹⁹ and are updated several times a year (every eight weeks). Loss coefficients may be both positive and negative, but are subject to a maximum cap of $\pm 10\%$. All losses (i.e., the product of load or generation and the corresponding loss factor) are valued at the NordPool spot price for the relevant hour. Finally, energy-related payments for marginal losses only represent one part of the Norwegian tariff system, while the residual network costs are recovered by means of other tariff elements.

Apart from locational marginal pricing for wholesale spot markets, the use of marginal losses represents the second dominant application of SRMC

pricing. While the approaches discussed above focus on investment decisions, this model primarily aims at short-term, i.e., operational or dispatch decisions. However, provided that marginal losses remain stable over an extended period of time, this method will simultaneously influence investment decisions as well. In summary, marginal loss pricing therefore clearly has the ability to improve short-run allocative efficiency but it may not necessarily provide correct long-term signals.

To achieve the method's full benefits, loss coefficients would basically have to be calculated in real-time. In practice, this is not possible which explains the decision by Statnett to provide estimated data for predetermined time periods. Due to the variation of losses over time, this adjustment causes a deviation between the standardised loss coefficients and actual losses. For this reason, pricing signals from marginal loss coefficients in Norway will not generally be truly identical to the system's SRMC, thereby somewhat reducing the scope for achieving allocative efficiency. The same basically holds true for the limitation of marginal losses to $\pm 10\%$, which will provide insufficient (i.e., too small) pricing signals at all places where marginal losses are capped.

A similar dilemma can be observed with regards to the price of marginal losses. In theory, these would also have to be determined in real-time. In Norway, hourly spot prices from NordPool arguably serve as a good proxy. But this approach depends on the existence of an (liquid) electricity spot market. An alternative approach has been implemented in Sweden, where the energy price of losses is determined for one year in advance (separately for four time periods), based on the bulk-purchasing rate that Svenska Kraftnat has to pay to its suppliers. But this again introduces strong elements of average cost pricing and reduces the scope for allocative efficiency.

Like all other methods of marginal cost pricing, this approach does not allow the full recovery of network cost by the marginal element.²⁰ Nevertheless, marginal loss payments actually lead to an income in excess of the total cost of network losses, called transmission loss rentals,²¹ allowing a slight reduction in payments for the residual cost.

Compared to the approaches for LRMC pricing discussed above, marginal loss charges will be much more volatile. This does not harm short-run allocative efficiency as it will provide adequate signals for dispatch decisions. But the lack of stability can be detrimental with regards to long-term investment decisions, especially when marginal losses fluctuate considerably or, in the worst case, may be both positive and negative at the same location over time.²² Under these circumstances, a user may no longer be able to assess the impact of marginal losses on his long-term cost.

Full Nodal Pricing

The nodal pricing manages congestion and sets transmission prices through a centralised energy market based on economic dispatch. The basic idea of the nodal pricing approach is to organize the market as a pool in which generators (and ideally loads) submit hourly bids for node specific injection and withdrawals of power to the pool operator with full co-ordination and price setting authority. The pool operator minimises the total system's gain from trade (demand bids less generation offers) subject to transmission and reliability constraints. The price at each node is then set to the incremental offer or bid price of the most expensive unit generated or consumed at that node. These nodal prices become the hourly prices charged to loads and paid to generators at the respective nodes. When there is no congestion and losses, all nodal prices are in theory identical.

As prices rise, suppliers will provide more if the price exceeds the cost of providing an additional unit, while consumers will consume less as the price rises above the benefit gained from the last unit being consumed. Dispatchers can use this mechanism to ensure that supply and demand are matched most efficiently. Full dynamic SRMC-based spot pricing promotes efficiency by signalling the true marginal cost of producing and transporting power across the network. In this case, the spot price at a particular location will equal the cost which would actually be incurred by supplying one more unit of demand at that location, taking account of the necessary adjustments in generation patterns to re-dispatch generators to avoid interregional constraints, make up losses and maintain system security.

The nodal pricing approach is also able to accommodate bilateral transactions.²³ Such a transaction is then subject to an ex-post transmission charge that equals the opportunity cost of the transaction, i.e., the cost difference of selling the power to the pool at the injection node price and buying it back at the withdrawal node price. Thus, the transmission charge between any pair of nodes is set ex-post to the nodal price difference between the nodes. The cost of transmission, therefore, varies between each pair of locations and is only known to trading parties after the fact.

This regime provides an explicit and accurate signal to market participants of the behaviour required to achieve an efficient dispatch and accounts explicitly for the available transmission capacity. Therefore, if transmission flows between two regions reach the system's flow capacity; the locational price in the receiving region will rise relative to the price in the sending region, which encourages more expensive generators in the receiving

region to provide for additional local demand. A revenue (congestion surplus) is earned on the constrained transmission capacity because the transmitted power is sold in the receiving region at a higher price than that for which it is purchased in the sending region.

8.4 Revenue Requirements and Cost Recovery

8.4.1 Establishment of Revenue Requirements

Revenue requirements are equivalent to the justified costs that should be allowed to be recovered through selling of the regulated transmission service. The terms "economically justified" and "eligible cost" are widely used by the regulatory authorities. Eligible costs should include the reasonable efficient O&M and capital costs (including depreciation and return on assets). The O&M costs are often considered the "cash outlay" costs of an infrastructure business. Recovery of these costs does not provide any return to the infrastructure owner, as they are paid out in the form of salaries, ongoing O&M costs, emergency service costs, etc. These costs allow the business to provide and maintain service. On the other hand, the inclusion of capital costs in the revenue requirement formula recognises the owner's investment in the regulated utility and the capital-intensive nature of network infrastructure businesses.

The regulation should recognise the importance of recovering sufficient level of O&M and capital cost. However, it is important that the regulated network service providers do not incur excessive or unnecessary costs in providing their services. It is in the interests of all concerned that the FNC and interested parties are able to examine the level of current and forecast costs, and are able to compare those costs with other similar entities (abroad).²⁴

Establishment of revenue requirements (including capital and O&M costs) is a topic that should be discussed for the purposes of network pricing design (cost allocation and tariff structure) and network price control concept (price regulation).

Asset Valuation

The inclusion of capital costs in the revenue requirement formula recognizes the owner's investment in the regulated utility and the capital-intensive nature of network infrastructure businesses. Failure to include adequate

capital related costs as part of the revenue requirement of the regulated business risks a reduction in investment in the industry. This could ultimately lead to reductions in cost coverage and quality service levels, hence to a reduction of the security of supply in the medium and long term. Fundamental to the measurement of capital costs in the revenue requirement is an assessment of the regulated business' capital investment and the establishment of a regulatory asset base (asset value that is used for the calculation of return on assets). Arguments about asset valuation vary from insisting ownership rights be recognised, to questioning whether any value should be attributed to sunken investment. The regulators usually endorse particular asset valuation methodology. These asset valuation issues must be considered with regard to the functional adequacy of regulated assets, market assets value, overall profitability of the regulated business and sustainable cash flows of the business as well as equity considerations.

The review process did not find explicit statements in the Russian documents of how the opening asset value has been computed for the purposes of price determination. Based on these investigations, one might assume that the existing assets are valued on a historic cost basis. The historical cost methodology values assets at their original purchase price. This approach has the advantage that it is administratively efficient and can be easily audited because the data should be available from financial statements. Further, it is relatively inexpensive since it does not require the explicit computation of regulatory asset values and it is objective because it relies on actual data rather than judgements. However, the method exhibits some disadvantages. Historic costs may understate asset prices in times of inflation and overstate asset prices in times of technological change. Secondly, it may lead to unstable prices (e.g., prices may rise when new, more expensive assets replace existing assets). Thirdly, data may be inadequate (especially for assets that have been acquired a long time ago). Historic costs are generally applied in the regulatory price control in the USA. Also, a number of regulatory authorities in Europe applied current book values from financial statements of the regulated service providers for asset valuation purposes in the regulatory accounting.

On the presumption that some compensation for asset or capital devaluation through inflation must be conveyed to the investors in order to allow for its long-run operation (new investments have to be paid for in current values and must be partly financed through revenues in the past), inflation compensation can in principle be achieved through the inclusion of asset replacement costs in the regulatory schemes. The replacement cost

methodology calculates the cost of replacing an asset with another asset (not necessarily the same) that will provide the same services and capacity as the existing asset.

One interpretation of depreciated replacement costs is that it is a valuation methodology consistent with the price charged by an efficient new entrant into an industry. Hence, it is consistent with the price that would prevail in the industry in a long-run equilibrium. A second interpretation is that replacement costs reflect the price that a firm with a certain service requirement would pay for existing assets in preference to replicating the assets.

The main economic principle for assessing the economic value of any assets is that their value to investors is equal to the net present value of the expected future cash flows generated by those assets. The practical difficulty in making this assessment for regulated monopoly businesses is that the future revenue derived from the assets is itself determined by the regulator — hence the issue of circularity associated with the use of discounted future cash streams as a methodology to value sunk assets. This potential circularity could be eliminated by the use of a replacement cost approach. The value of a network is the sum of the depreciated replacement cost of the assets that would be used if the system were notionally reconfigured so as to minimise the forward looking costs of service delivery.²⁵

The application of top-down calculations based on the discount cash flow approach would be hardly possible because of lack of a stable price starting point in Russia.²⁶ Moreover, the application of final consumer prices as a starting point for discounted cash calculation could lead to low assets value if these prices are not cost-reflective. Obviously, this will be the logical outcome since the revenue streams that could be earned with the particular assets will depend on the pre-determined starting prices used in the calculation. Further, even if there is a political will to correct the starting point of calculation, the determination of the normative level of these starting prices will be associated with a number of methodological difficulties mostly related to the "circularity" problem. The "circularity" problem results from the fact that on one side, the regulatory assets value depends on the revenues (prices) as they determine the cash inflow in the discount cash flow calculation. On the other side, the common logic of the regulatory process requires the regulatory asset value as an input in order to determine the revenues and prices.²⁷

All the issues described above necessitate the application of proxy approaches for determination of regulatory asset base. Generally, one will

need to compare the usefulness of the historic and replacement cost approaches. The application of historic costs for the establishment of the opening value of the regulatory asset base is also unlikely to be appropriate due to the strong inflation figures in the past. If in Russia the assets have been revaluated (in many of the Central and Eastern European countries, the assets have been revaluated), the current balance sheet values reflect more the replacement costs (at the time of revaluation) than the historic costs.

The application of replacement costs will require special attention. As already mentioned, the replacement cost methodology calculates the cost of replacing an asset with another asset (not necessarily the same) that will provide the same services and capacity as the existing asset at the time of valuation. The application of replacement costs should be checked by FEC with respect to the following major issues:

- Is FEC competent to impose revaluation of assets for regulatory purposes and does it contradict with the national statutory accounting?
- Is it feasible to revalue the assets?
- If price should be raised after the assets revaluation is it politically and socially acceptable?

Depreciation

To an accountant, the term 'depreciation' means a systematic allocation of the cost of an asset to the accounting periods in which the asset provides benefits to the entity. This allocation is designed to mirror the consumption of the service potential or economic benefits associated with an asset over its useful life, resulting from both use and obsolescence. The purpose of provisions in accounting is to ensure that the cost of the flow of services provided by capital assets is met in the price of these services, and additionally to build up funds for the replacement of these assets.²⁸

In the companies' accounting, a number of approaches are used for constructing the depreciation schedule. These include:

- Nominal (or historic cost) accounting method — usually straight line or diminishing balance reductions²⁹ in the original cost over time;
- Current cost accounting method (based on replacement cost estimates) — again straight line or diminishing balance depreciation based on the age of the asset is applied; and
- More flexible arrangements whereby depreciation is adjusted to complement other components of return so that the revenue stream mirrors the behaviour of an annuity.

While the assets valuation aspects have been discussed in the previous section, here an explanation will be given on the establishment of depreciation profiles. In traditional regulatory frameworks, straight-line depreciation is the norm and the use of straight-line depreciation by FEC would be consistent with those frameworks.³⁰ In common with the traditional approach used in valuation, this approach calculates the write-down of the gross asset value to obtain the depreciated asset value, by assuming a linear relationship between accumulated depreciation and the age of the asset relative to its expected economic life. This approach amounts to assuming that the economic depreciation of an asset is equivalent to straight-line depreciation of that asset. It makes sense that the approach to depreciation taken within the cost valuation methodology matches the one utilised within the regulatory framework.

Return on Assets

Regulated service providers compete for finance with companies operating in competitive markets and thus have to accept these conditions. Equity and debt finance will only be available to utilities that agree to credit conditions given to firms that operate in competitive industries and have a comparable credit ranking. Equity finance will only be available if profitability (consisting of a dividend component and market value growth component) can be expected that covers the risk-free rate of interest (i.e., yield of long-term credible government bonds) and a risk premium (where risks and risk premium can be expected to lie significantly below corresponding values in competitive industries, provided that the political and regulatory environment is predictable and stable). Investors will be interested in engaging in the regulated industry, if projects allow them to meet financial requirements. These financial requirements are measured against the benchmark of earnings to be made in other product markets, in industries in other countries or in the international capital markets.

Given the capital-intensive nature of electricity network businesses, the return on capital component of the regulated revenue could account for 50% or more of annual aggregate revenue. As relatively small changes to the rate of return can have a significant impact on the total revenue requirement and ultimately on end-user prices, it is important that the regulator sets the rate of return at a level that reflects a commercial return for the regulated businesses. Setting a rate of return below the cost of funds in the market could make continued investment in developing the network difficult or unattractive for the owner. This would cre-

ate pressure for the regulated service providers to reduce maintenance and capital expenditure below optimum levels and undermine the quality of service offered to users. Conversely, if the rate of return was set too high by the regulator, the regulated businesses would earn a return in excess of their cost of capital. This would distort price signals to consumers and investors, resulting in a misallocation of resources and sub-optimal economic outcomes.

The WACC is a commonly used method for determining a return on an asset base. Weighted Average Cost of Capital is determined in the regulatory regimes as the weighted average of the cost of each individual component of the capital structure weighted by its share.

$$RR = WACC_t = EP_t \times ROE_t + DP_t \times DI_t$$

where:

WACC — weighted average cost of capital for period t ;

RR (%) — rate of return for period t ;

ROE (%) — return on equity for period t ;

EP_t — proportion of capital comprising equity and equal to (value of equity)/(value of equity + value of debt) estimated at the end of period t ;

DI_t (%) — cost of debt for period t ; and

DP_t — proportion of capital comprising debt and equal to (value of debt)/(value of equity + value of debt) estimated at the end of period t .

Traditional industry use of WACC is to determine the value of the cash flows resulting from an investment to assess its profitability. It is a common practice in industry to determine the after-tax cash flows and apply a nominal after-tax WACC discount factor to those cash flows in order to develop a present value. In the regulatory environment, WACC is applied to an asset base in order to determine the cash flows a business will receive. A number of regulatory authorities have determined that a pre-tax real WACC should be applied to an inflated value of asset base in order to determine ongoing returns. The use of pre-tax WACC is used due to the reversal usage of the WACC formula and the tax affairs being viewed as a matter for the business and not the regulator. Some regulatory offices use after-tax WACC and consider taxes explicitly in a separate position similarly to O&M cost.

Operation and Maintenance Costs

Operation and maintenance costs are the costs incurred by the FNC in maintaining and operating the transmission network assets to the technical standards used in Russia and the requirements of the Technological Rules. Operation and maintenance costs will cover those cost elements that are spent for accounting purposes.

8.4.2 Treatment of Transmission Losses and Congestion Surplus

As a result of the application of nodal pricing, in which generators are paid at nodal prices of generation buses and consumers are charged at marginal costs of load buses, a surplus (transmission loss and congestion surplus) is collected for the owners of the transmission system.

Settlement is performed on the basis of the actual amounts of electricity generated by generators, and actual amounts of energy consumed by loads, as measured by meters distributed throughout the network. Because nodal prices are used in the dispatch and settlements processes, the amount paid by purchasers of electricity (loads) through the settlements process is normally higher than the amount paid to generators. In other words, the settlements process normally results in net income to the market operator (ATS) referred to as the "surplus." A number of solutions are conceivable.

Reduce Transmission Network Charges

The first solution is to use the surplus to reduce transmission charges, and hence is indirectly returned to end consumers. This mechanism for distributing the surplus is transparent as it will be part of the FNC's regulated revenue, can be universally applied, and is equitable, as consumers will receive a benefit via reduced transmission charges. It is also efficient in that it will have a minimal impact on the market while retaining economic signals. However, this approach has been criticised in some countries for adding complexity to the market design.

Return the Surplus Directly to Market Participants

An alternative suggestion is that the surplus be immediately returned to spot market participants. Accordingly, the surplus would be distributed to certain retailers and/or generators in proportion to the extent and nature of their participation (energy, capacity) in the spot market at the time that the surplus was accumulated. The rationale for this

approach is that giving the surplus to market participants will allow them to commercially trade the surplus in the market. It could be argued that this is a more viable option than delivering the surplus to the FNC who is a non-commercial player. It is not directly obvious why generators or retailers have a legitimate claim to what is essentially a revenue related to transmission service functions. Nor is it apparent where the commercial incentive for generators and retailers to pass the surplus on to end use consumers would come from. Returning the surplus to market participants may be a simple option, but given the above concerns over incentives it is doubtful whether this option meets the efficiency or transparency objective. In addition, the assignment of the surplus to generators or retailers raises a number of equity issues that would need to be resolved. Finally, such approach would probably dilute locational signals resulting from the nodal pricing. On the other hand, the latter could be a desirable effect at least for the demand side where most probably the nodal pricing would face some social or political resistance.

Use the Surplus to Relieve Network Constraints

The accumulation of a surplus provides a signal that it may be economically viable to augment the networks and hence remove constraints and this could be another solution. However, this option raises a number of issues, such as who becomes the owner of the new transmission lines and is thus entitled to the remaining and future surplus, as well as the regulated income. Moreover, augmentation of the transmission network may not be always the optimal solution.

Having listed above the properties of the different methods, it is preferred to apply the first method where the surplus is used to reduce the transmission charges. If the surplus was to be paid to generators, or refunded to purchasers, then the knowledge that these payments or refunds were to be made could be taken into account when developing bidding strategies, thereby diluting the effect of the marginal pricing concept on which the market is based. With the first option, the FNC can use the settlements surplus it receives to reduce the charges it imposes for the use of the transmission network. These reduced prices should ultimately filter through to reduced electricity prices for consumers. The Australian Federal Regulator ACCC in its determination of December 10, 1997 stated that the proposal to distribute the settlements surplus to end use consumers via reducing network charges is the most transparent, equitable and efficient distribution of the surplus.

8.4.3 Forms of Cost Recovery

In systems that present increasing returns to scale, such as the transmission business, this surplus is not sufficient to finance the system operation and development.³¹ The collected surplus covers a small part of the revenue requirements, the percentage varying depending on the system (see Table 1). A supplementary mechanism for cost recovery is necessary. For example, in South America, a two part tariff is used with the second part, in the form of transmission tolls, being added to the marginal cost income to fully finance the system. These tolls, or wheeling rates, are charged to users of the system, definition of users and allocation of charges varying from country to country.

Table 1. Cost recovery under SRMC pricing

Country	Pricing Concept	Cost Recovery
Australia	Long-Run Average Cost	50% (predefined, no explicit calculation of LRMC)
New Zealand	SRMC	10%
Norway	SRMC	17.2%
Chile	SRMC	10%
U.S. Estimate	SRMC	5%–20%
Bolivia	SRMC	3.6%

Source: *Read, Transmission Pricing in New Zealand, 1997; Glende & Westre, Transmission Pricing in Norway; Rudnick, Presentation: Latin American Experience in the Restructuring of Electric Power; 1998. Federal Energy Regulatory Commission, The Recovery of Fixed Transmission Costs, December 7, 1997, NARUC-DOE National Electricity Forum; Powerlink; OFGEM and NGC Reports; ACCC and IPART Reports. Normative Solutions.*

The ideal cost recovery regime would involve a "perfect tax" which, in order to avoid distorting incentives in the electricity sector, should be levied on some basis other than electricity consumption. Thus, it could be argued, the efficiency of the electricity sector could be enhanced if the government were simply to write off the value of existing assets, so that pure marginal costs could prevail. This is obviously unrealistic, but the argument is also flawed, in two respects. First, this solution could only be implemented by raising general taxation, which will increase distortions elsewhere in the economy.

Second, the precedent could have unfortunate repercussions with regard to future expectations of cost recovery, and price levels, in the electricity sector.

Economic theory suggests that normally only exactly one system of relative prices between all goods leads to maximum efficiency and it occurs when prices are equal to marginal cost. This rule will be labelled "allocatively efficient pricing" and provides the best possible outcome that maximizes aggregate consumer welfare. However, in the case of decreasing long-run average cost function (typical case in the electricity transmission networks), with prices equal to marginal costs, revenues fall short of total costs by an amount equal to the fixed costs. Therefore if the regulated service provider is required to be self-financing, it is necessary to choose a different set of higher prices. The Ramsey prices³² are the "second best" prices that succeed in raising enough revenue to cover total cost with the smallest possible sacrifice in consumer welfare. The general result of Ramsey pricing rules is that the departure from the marginal cost pricing (best solution) should be inversely proportional to the price elasticity of demand of the relevant product. The principal idea is that the demand structure should not be heavily distorted by the allocation of the remaining cost.³³ Ramsey prices are applied in a limited manner as the price elasticity of demand is not known.

Locational Versus Non-Locational Cost Recovery

On one hand, it could be argued that sunk cost recovery should be designed so as not to send any locational signals since there is no economic point in sending locational signals, based on historical costs, to parties who have already located. The signals they require are already conveyed by the locational spot prices and by the associated new investment regime. On the other hand, the application of locational pricing for the existing transmission is frequently justified by the need to establish long-term signals. This argument is also supported by the fact that such locational pricing models aim either to calculate explicitly or approximately the level of LRMC of transmission through partial allocation of sunk cost on a locational basis.

Determining the appropriate balance of these charges should be driven by the following considerations:³⁴

- longer-run cost signalling objectives;
- cost attribution objectives, but these require an investigation of the relationship between LRMC and average network costs and may lead to different balances across regions;
- the extent to which greater reliance on locational charging is likely to lead to network bypasses; and

- equity (or cost averaging) objectives across the network resulting from social goals or other political constraints.

Non-Locational Post Stamps

This method is a further simplification of zonal pricing where the zones are defined solely on the basis of voltage level without using further criteria for locational differentiations resulting from load flow and allocation of generation and demand. According to this method, the revenue requirements are allocated to each voltage level (but not to node) using a cost cascading approach; i.e., the cost of a certain voltage level is allocated to the cost of the lower voltage level and the consumers connected to this voltage level. The transmission tariffs are differentiated by voltage level (additionally could include differentiation for types of customers and time of use).

To ensure that costs are properly allocated, the costs should be classified into categories that reflect the main factors determining the level of overall costs of providing the relevant service. The cost allocation procedure should provide an estimate of the cost associated with the delivery of the products provided by the regulated service provider and should establish a basis for pricing these products. Generally, the cost allocation process includes the following steps:

- definition of cost categories;
- definition of cost pools (cost centres);
- allocation of costs of each category to cost centres;
- definition of products delivered by the regulated service provider; and
- allocation of cost attributed to each cost centre to the relevant products.

This is also the process followed by the Order N 72-э/3 where each transmission region is priced separately and voltage levels are defined as separate cost pools (cost centres). Once allocated to the voltage levels, costs of each regional transmission network are cascaded downward in the system in order to derive the network cost to be allocated to customers connected to the respective voltage level.³⁵

Locational Methods

Use of Distance Capacity Related Approach

For any off-take or injection, the resulting MW-km are calculated. For off-takes it is assumed that the generation has been injected at the reference node and for injections vice versa. The MW-km of each injection and each

load over all nodes are summed up to calculate an aggregate MW-km utilisation of the network. The revenue requirements are divided through the aggregate MW-km in order to obtain a specific MW-km price. The MW-km of all injections and off-takes are valued with this specific MW-km price and then divided through the relevant usage quantity (registered capacity or peak load) in order to obtain a capacity related price for generation and demand at all nodes.

Load Flow Based Participation Factors

The algorithm of average participations is simple and robust but it uses a very simplified representation of the network (only the first Kirchhoff's law). However, this simplification may be justified, since the criteria to allocate the complementary charge are also fairly loose and mostly of a qualitative nature.³⁶ The method deals separately with injections to and withdrawals from the network. First, the responsibility of the historical flow in each line and at a given time is fully allocated to the injections at the nodes. Next, the flows are completely allocated again to the withdrawals from the network. In this way, there is the option of assigning different global weights to injections and withdrawals when computing the final coefficients of allocation of flows to generators and loads. Differently from the average participation factors, the incremental participation factors are defined by means of a sensitivity analysis that relates a variation in the injected/ejected power in a busbar to a variation in the flow through a certain transmission line. It considers that a change in the injected/ejected power in any bus is absorbed by a similar negative change in the reference bus, while injections from generators and loads in other buses, as well as losses, are maintained constant.³⁷

Use of Fault Level Contribution Matrix Approach

Similar to the method of incremental participation factors, this method is also based on a sensitivity analysis. It uses the 'electrical distance' represented through the impedance of the circuits in order to pair generation and load. A way to pair generation and loads using the "electrical distance" is the application of Fault Level Contribution Matrix. By faulting each bus, the fault current supplied by each generating unit is determined depending on the impedance of the circuits. After pairing the generation and load, a full system load matrix is constructed, showing how the generation nodes are allocated to the load nodes. The matrix is multiplied with a sensitivity matrix that gives the information on how the circuit loading changes in case of changes in different network nodes. The figures obtained are used to allocate the circuit cost to the various nodes.

8.4.4 Cost Allocation under Time Dependence Aspects

Primary responsibility for the total network capacity required lies with those consumers who make demands for energy at times that coincide with the demands made by many other consumers. The degree of coincidence dictates the level of the peak demand on the network and the sharpness, or otherwise, of the peak obtained when demand is plotted against time. Because the time at which the peak demand occurs may change from day to day and month to month, it could be argued that it is better to allocate network costs over a number of hours rather than to a single hour. This is particularly the case if the actual peak observed tends to be broad rather than narrow in time. Further, the cost allocation may not be proportional to the energy delivered during these periods because the probability of not meeting the load in the demand peak periods is greater than in the demand off-peak periods. Therefore, the proper economic pricing signals would be ensured when larger portions of the revenue requirements are allocated to the hours of peak demand.

The time relationship of the network costs is not as immediately obvious as the variation with time of the cost (prices) of the generation. However, the overall capacity and, hence, capital and maintenance cost of the networks is dictated by the peak loading to which they are designed. This, in turn, depends on the coincidence in time of the demands made by individual consumers. When this is taken into account, responsibility for the costs of these networks can be allocated to particular time periods. Development of tariff structures includes reaching a decision on the periods at which time of use tariffs will be made available to the network users — viz time of day and season of year. It is also necessary to define which hours of the day are to be included in each period and if the periods should change in time or duration from season to season.³⁸

The time of use tariffs vary with the time of delivery and allocate transmission network cost (revenue requirements) to the different periods of time. The allocation rules could use the probability of not meeting the maximum demand in the relevant voltage level or simple criteria based on the load profiles on each voltage level. In a simple case, it could be found that there is not a reasonable expectation of demand exceeding the available network capacity during any off-peak period (e.g., night, weekend, etc.) However, as to the remaining hours the probability of not keeping the demand supply balance is higher, hence a greater portion of the network cost could be allocated to those hours.³⁹ The time of use tariff would require time-dependent metering to register the energy or demand served in the tariff time periods.

8.4.5 Payment Liability

Secure transmission network benefits both the load serving entities/direct retail customers and the power plants. Without a reliable transmission grid, the generation units (connected to the transmission grid) are not able to feed in electricity into the transmission grid. On the other hand, reliability is also important for load-serving entities and directly connected customers. A measurement for the value of reliability for load customers is the value of lost load. It could be argued that the foregone revenues of generators not allowed to generate electricity due to an unreliable transmission network are lower in relation to the value of lost load. Another argument to allocate a higher portion of transmission service payments to load could be justified by the fact that load is less elastic and pricing under the aspects of allocative efficiency should consider allocation according to the inverse price elasticity of demand for transmission service.

Also on the balancing market (superimposed on the day-ahead market) where the generators might be allowed to react using incremental/decremental bids when the transmission system precludes them from generating. This supports the proposition that the generators benefit from availability of a secure transmission network. It would be inappropriate for generators to be recompensed in such a way, were the transmission system not planned and operated to defined security standards and were generators not to bear part of the burden of having a secure transmission network. On the other side, imposing payments on the generators for full compensation of revenue requirements on a locational basis would not bring additional incentives in term of locational efficiency, as part of transmission network costs (assets and non-assets related costs) is related to a function that might not exhibit a locational nature and provides equivalent benefits to all transmission service users without any differentiation on their location.⁴⁰ From this point of view, advantages of joint payment liabilities might be justified with the fact that the forwarding of costs incurred by the generators for payments of those costs of the transmission network would reflect the individual bidding strategies of the generators in the wholesale market and accordingly their willingness to reduce some of the expected production surpluses. The effect might be limited as the proportional payments would be more or less passed through in the bids of the generators. However, recovery of sunken transmission cost through the energy market prices will certainly have an adverse effect on the efficiency signals delivered by the wholesale market prices.

It is proposed that both generation and load are liable for transmission service, the prevailing part to be charged to the load nodes (e.g., 25% to generators and 75% to the load). However, this should only be applied if the FEC and/or FNC deem it appropriate to introduce locational pricing for the transmission network infrastructure. If this is not the case and the postage stamp approach is used, it is more meaningful to assign the costs entirely to the load entities.

8.4.6 Transmission Tariff Design

The Order N 72-э/3 details the transmission use of network charges in RUB/kW. Economic theory suggests that charges would be better based on peak usage, rather than energy.⁴¹ Cost allocation according to the peak demand seems to reflect the nature of the transmission network costs (most of them are fixed) as the transmission network is designed to meet the system peak.⁴² Also from an economic point of view capacity has no value unless it is fully used. Thus, at off-peak times, no one who would be prepared to pay for more capacity since none is required, while at peak-times capacity is more likely to be fully utilised, and users are likely to be prepared to ration capacity. However, using only one peak period makes the whole payment liability dependent on one single value and exposes the service users to rather stochastic characteristic of demand. Moreover, a user may substitute other forms of supply (e.g., via embedded generation) to lower peak usage and hence charges. This may be inefficient, though, as the user is incurring costs and using potentially expensive resources instead of using an asset for which the costs are the sunk.⁴³ In order to avoid these negative effects additional system peak loads are considered in the transmission price design (e.g., load triad used by NGC in UK).

Using only a demand charge (even when using more than one peak period to determine the payment liabilities) could create a substantial difference in the payment of transmission service users having different profiles. The effective payments of customers with higher load factors will be much lower than the customers with lower load factors. Therefore for equity reasons, the tariffs could consider the introduction of energy-dependent charges and fixed charges⁴⁴ and the related non-demand criteria in the process of tariff design development. Variable energy-based charges achieve these equity objectives and could result in limited distortions in the energy market, since relative energy costs in different demand periods are not distorted. These charges might be more appropriate than peak-based charges, where there is an excess network capacity.

However, when the charges are based on energy, then the user faces a network cost that depends upon its throughput of the network, even though throughput may have no impact on the level of costs to be recovered. Consequently, the user sees a higher cost per unit of delivered energy than it would if charges were truly fixed. As a result, the user may have an incentive to reduce energy consumption below the level implied by SRMC alone. This would be inefficient, as reducing the usage of the transmission assets produces no saving in the cost of the asset.

We suggest two tariff design options: (1) peak-based regime, where the peak usage is appropriately defined so as to minimise the attractiveness of inefficient charge avoidance; and (2) combined approach using energy, demand and fixed charge. The formulas below are designed for joint payment liability, i.e., load and generators pay for transmission service but could be simply rewritten only for load.

Option 1: Demand-Based Charges

The charges reflecting the residual costs will be calculated as:

$$UNC_g = \frac{CA_g}{RC_g} \text{ (RUB/kW/year)}$$

$$UNC_l = \frac{CA_l}{AL_l} \text{ (RUB/kW/year)}$$

where:

- CA_g — revenue requirements to be recovered through transmission use of network charge from generators;
- CA_l — revenue requirements to be recovered through transmission use of network charge from load;
- UNC_g — transmission use of network charge for generation;
- UNC_l — transmission use of network charge for load;
- RC_g — total rated generation capacity; and
- AL_l — system peak load or average of a number of system peak loads (for pre-defined assessment period⁴⁵).

Option 2: Multi-Component Tariff Design

The total costs allocated to nodes comprise three component prices:⁴⁶

$$CA_g \rightarrow CA_g^{(fixed)} \text{ (e.g., 25\%); } CA_g^{(demand)} \text{ (e.g., 50\%); } CA_g^{(energy)} \text{ (e.g., 25\%)}$$

$$CA_l \rightarrow CA_l^{(fixed)} \text{ (e.g., 25\%); } CA_l^{(demand)} \text{ (e.g., 50\%); } CA_l^{(energy)} \text{ (e.g., 25\%)}$$

The demand, energy and fixed transmission use of network charges are derived as follows:

Generation Nodes

$$UNC_g^{(demand)} = \frac{CA_g^{(demand)}}{RC_g} \text{ (RUB/kW/year)}$$

$$UNC_g^{(energy)} = \frac{CA_g^{(energy)}}{AL_g} \text{ (RUB/kWh/year)}$$

$$UNC_g^{(fixed)} = \frac{CA_g^{(fixed)}}{AL_g} \text{ (RUB/year)}$$

Load Nodes

$$UNC_l^{(demand)} = \frac{CA_l^{(demand)}}{RC_l} \text{ (RUB/kW/year)}$$

$$UNC_l^{(energy)} = \frac{CA_l^{(energy)}}{AL_l} \text{ (RUB/kWh/year)}$$

$$UNC_l^{(fixed)} = \frac{CA_l^{(fixed)}}{AL_l} \text{ (RUB/year)}$$

9 Connection Charges

Separation of transmission network service from transmission connection service follows the economic causality principles, namely to charge costs to the party that has caused these costs. Accordingly, the transmission pricing should distinguish between transmission use of network charges and transmission connection charges. The transmission connection charge should cover only the cost of the transmission connection assets and should be charged directly to the respective connected party (generators and load entities, called transmission connection service users). On the other hand the transmission use of network charge should recover the costs (including depreciation, O&M costs and return on the transmission network assets⁴⁷) necessary for secure transport of power.

Connection charges are levied when a user first connects to the system or requests a substantial change to its capacity connection. This applies to both generation and demand, although the treatment of these two types of connections may vary. Connection charges should cover the cost of providing the assets necessary to make the connection, including a reasonable rate of return on those investments, plus, possibly, ongoing maintenance costs. In general, these charges should be designed to recover in full from the user, often as an up-front, one-off, charge, the relevant costs of making the connection.

In deciding on the relevant connection charges that should properly be charged to the connected party, a fundamental problem is related to the split of assets between the "connection" and the "core transmission network." Generally, two extreme positions can be considered. In the first, known as shallow charging, the connected party is required to fund only the assets specifically required to connect it to the system and for the specific benefits of this particular user. The second, known as deep charging, levies all the costs consequential on making the connection on the connected party, including the costs of incremental investment in the wider system, whether or not these relate to the network local to the connection point. The advantages and disadvantages of these approaches are considered below.

9.1 Shallow Connection Charges

Under this approach, the connected parties pay for only those assets located in the immediate vicinity of their point of connection. These are the

assets required to make a physical connection from the existing system to the connected party. Any costs associated with the network augmentation are regarded as resulting from general demand or generation growth and are recovered from all users through transmission charges.

There is inevitably scope for defining shallow connection assets more or less narrowly. For example:

- a "very shallow" approach would exclude from connection assets all spurs between the generator or customer and the main system. It would therefore include all substation assets (mainly, transformers, switchgear, and busbars) and exclude all overhead lines or cables, other than those which constitute substation assets;
- a "less shallow" approach might include all single and multiple spurs that serve to connect the generator or customer or groups of generators and customers to the core transmission network, in addition to substation assets;
- an intermediate position might be one in which "generation only spurs" are included in the definition of connection assets, to give potential generators adequate locational signals. Spurs connecting customer exit points to the main system would be classified as "core transmission network" assets, on the grounds that the case for locational signals is stronger for generators than for customers.

Shallow connection charges tend to minimise the charges required to connect to the system, thus encouraging the development of demand and generation connections. Although this requires a precise definition of assets belonging to the "connection" as opposed to the "core transmission network," it is possible to create a set of clear and explicit rules to cover most situations.⁴⁸ Consequently, shallow charging makes it easier for users to estimate connection themselves and thus tends to be more transparent.

The usual critic against shallow connection is that connected parties receive limited locational signals about the full incremental cost (not only the cost of direct connection assets but also the costs of augmentation of the core transmission network) resulting from their connection. Charging the costs of augmentation of the core transmission network via the general transmission tariffs may encourage inefficient entry of new users as the cost caused by their connection will be recovered by the existing transmission service users. Finally, shallow connection charges socialise the risk of stranded network assets that may occur if a transmission service user leaves the system.⁴⁹

9.2 Deep Connection Charges

In this approach, the connected parties pay for full incremental costs incurred by the transmission network service provider as a result of their connection, including investments at places remote from the connection site. Examples of this are where the system needs reinforcement at a higher voltage level than that of the connection as a direct result of the connection, or where equipment has to be upgraded, perhaps to increase transmission capacity between different parts of the network, or to cater for increased fault current levels. In other words, the connected party is required to pay for all the transmission assets which would not be required if the particular user did not exist, including the cost of reinforcements at remote sites.

Therefore, deep connection charges tend to be higher than under a shallow connection policy. On the contrary, deep charging tends to reduce the level of transmission charges for all users (excluding the connected party). The major advantage of deep connection charges is that they provide strong locational signals for new connections. This helps to improve allocative efficiency, although many customers, especially domestic consumers, will generally be unresponsive to locational signals. Nevertheless, it is fair to say that deep charging inhibits the same positive properties as locational transmission charges.

The main drawback of a deep connection charging policy is that these charges are difficult and arbitrary to apply in practice. While it is relatively easy to define the required assets in the immediate vicinity of the connection, this is not the case for reinforcements throughout the system. Most importantly, this relates to problems caused by the significant and "lumpy" nature of reinforcement that can be driven by a small increase in connection capacity. For example, if a new connection requires network reinforcement and the capacity upgrade, because of lumpiness, exceeds the new load, significant redundant capacity will occur, which is not immediately required by new or existing users. In consequence, the resulting costs may be much higher than would be required to meet the new connected load.

In the sense that some remote reinforcements can be argued to be of benefit to a great number of users, deep connection charges may not be cost-reflective. In order to avoid discrimination, this may require an allocation of these costs, which will be arbitrary by nature, to the connected party and the existing users. Deep connection charges are therefore neither transparent nor predictable for users.

Deep connection charging may discourage new connections because of the higher level of initial charges that connected party will need to pay.

Deep connection charging requires an explicit redistribution mechanism among the connected parties. If the first user is charged the full deep connection costs the spare capacity gain will bring benefits to the later connected parties. These later connected parties should reimburse part of the full deep connection costs to the first user.

Similarly, if applied for the first time, deep connection charging will probably be discriminatory between existing network service users and new entrants. While it could consistently be applied to all new connections, this would be impossible for existing users, given the historic nature of the system. In other words, it would be impossible to determine for each existing connection what remote reinforcement costs were necessary in the past to accommodate those connections and to derive appropriate charges in that case.

10 Price Control

The introduction of economic regulation is caused by the need to monitor and control the activities of companies in markets where full and fair competition cannot be relied upon to protect consumers' and other interests. Generally, if full and fair competition exists then there is no need for regulation in order to protect the interests of consumers. However, where competition is not possible (in case of market failures, e.g., natural monopoly, public goods, etc.), economic regulation is required, generally to prevent the abuse of market power by dominant or monopolistic industry participants.

Regulation in the line-bound energy industries can be defined as the action of state authorities to limit (profit maximizing monopolistic) utilities in the use of one of their strategic parameters such as price, quality, quantity, investments (e.g., capital intensity, plant type). In this sense, regulation is used to prevent abuse of market power by incumbent monopolistic utilities. Regulation represents a codified, continuous monitoring of utilities and control of certain parameters if parameter thresholds (e.g., prices, profits, revenues, etc.) are exceeded.⁵⁰ Usually, regulation allows for a "fair rate of return" so that strategic management decisions such as investment in regulated businesses/industries are evaluated similarly as in competitive industries: no undue profits are to be earned, but profitability of regulated businesses must be such as to allow for continuous operation of the business, including replacement investments.

10.1 Rate of Return Regulation

Under the rate of return regulation regime, costs of service provision are passed through onto customers. In addition, utilities are allowed a "fair" rate of return on the rate base (either equity or total fixed assets) that is also levied upon customers. This regulation method is synonymously labelled "cost of service" or "cost-plus" regulation (with the allowed rate of return representing the "plus" element).

Explicit per unit revenues (average prices) are derived by averaging forecast cost (plus "rate of return") data over expected units to be supplied:

$$P_t = \frac{TC_t + ROR_t \cdot RB_t}{Q_t}$$

where:

- P — price/average revenue;
- TC — total cost;
- ROR — (allowed) rate of return;
- RB — rate base;
- Q — quantity sold; and
- t — time index.

Rate of return regulation is flexible and responsive to the regulatory authority's wishes concerning service quality and other matters, because the regulatory authority can allow or disallow any costs it chooses. On the other hand, rate of return regulation is complex and could be potentially distortional, because it requires the regulatory authority to judge and "second-guess" detailed business decisions taken by the monopolist.⁵¹ Moreover, under conventional "rate of return" regulation, rates of regulated network service providers are reviewed on a regular basis and have to be adjusted (so-called "regulatory review") to lower levels if cost savings have been achieved in the interim (since the last review). Network service providers thus only benefit from cost savings to the extent that regulatory reaction to cost savings is lagged (so-called "regulatory lag"). If the "regulatory lag" is short — one or two years — incentives for cost savings are suppressed almost completely. In such a framework, network service providers are said to operate dynamically inefficiently. Additionally, the "rate of return" regulation creates over-capitalization (over-investment in fixed assets) incentives (Averch–Johnson effect) and potentials for increases in capital costs.

Numerous regulatory methodologies have been developed to counteract the deficiencies of "rate of return" regulation.⁵² All these alternative methodologies focus on the establishment of incentive mechanisms from "rate of return" to "profit sharing," or further to "cap" regulation, or again through comparative approaches such as "yardstick" regulation rather than individual assessment of network service providers.

10.2 Yardstick Regulation

Yardstick regulation is not based on an assessment of the cost position of individual utilities but upon a comparison of prices or cost positions and cost determinants between firms. Under a "yardstick" mechanism based on price information, utilities are not allowed to charge higher prices than some statistical mean price that is calculated over all utilities unless this was justified by "special conditions."⁵³

This form of price control was suggested by Shleifer (1985). Each in a group of comparable regional monopolists has a price cap determined by the average cost of the others in the group. In this model, operating costs C_i of firm i depend on the amount of effort expended by the firm $C = C(e)$. For $n \geq 2$ identical firms, the regulator defines for each of them the allowable cost:

$$C_i = \frac{1}{n-1} \sum_{j \neq i} C_j$$

Under a cost-based "yardstick" approach, utilities could apply for certain rates according to some standard regulation method (e.g., "rate of return"). The regulatory authority then compares detailed cost components between utilities and disallows any parts of cost components that appear excessive upon comparison with other utilities.

10.3 Performance-Based Regulation

Performance-based regulation (PBR) was introduced as an alternative to cost-of-service regulation in the United States' electricity sector in the late 1980s and early 1990s. The primary objective of PBR is to inject competitive market incentives into monopoly markets and weaken the link between costs and rates. Performance-based regulation builds on "rate-of-return" regulation. In addition, selected "performance" targets are set for the utility in quantitative but non-monetary terms. If performance targets are achieved, these will come along with cost savings. The cost savings will not fully (or they will not immediately) have to be passed through to customers. Thus utilities are, for example, allowed a higher rate of return if performance targets were previously met. Potential categories of performance targets are: construction cost, transportation losses, system reliability (measured as problems in voltage or pressure stability, outages, time to restore supply upon outages), customer satisfaction (e.g., measured as numbers of customer complaints). It should be emphasised that the philosophy and major ideas beyond PBR coincides to some extent with the term "incentive regulation" that is frequently used in the European regulatory theory and practice. On the other hand, incentive regulation is usually associated with forms of regulatory cap controls that generally rely upon extended regulatory lag and fixing of efficiency increase requirements, factors that are not directly considered in the basics of PBR.

The use of performance indicators gives the companies incentives to increase their operating and investment efficiency and reduce cost. If a company is able to achieve better performance than the target set, it can retain some of the earnings resulting from cost savings. On the other hand, PBR requires individual negotiations between utilities and regulatory authorities in the process of selecting target categories and target levels (high regulatory burden). Such an approach is only feasible if regulators face a limited number of regulatees or if target schemes can be blue-printed for a number of regulatees.

10.4 Cap Regulation

Given that regulation is seen as a set of instruments to mimic market conduct, "cap" regulation can be used to reproduce competition. In competitive industries, intermediate profits can be earned if product or process improvements give an "innovator" some competitive advantage over his rivals. This advantage can only be sustained as long as competitors do not imitate the innovation (this may be restricted by patent legislation). In competitive capital-intensive industries, this intermediate advantage is usually

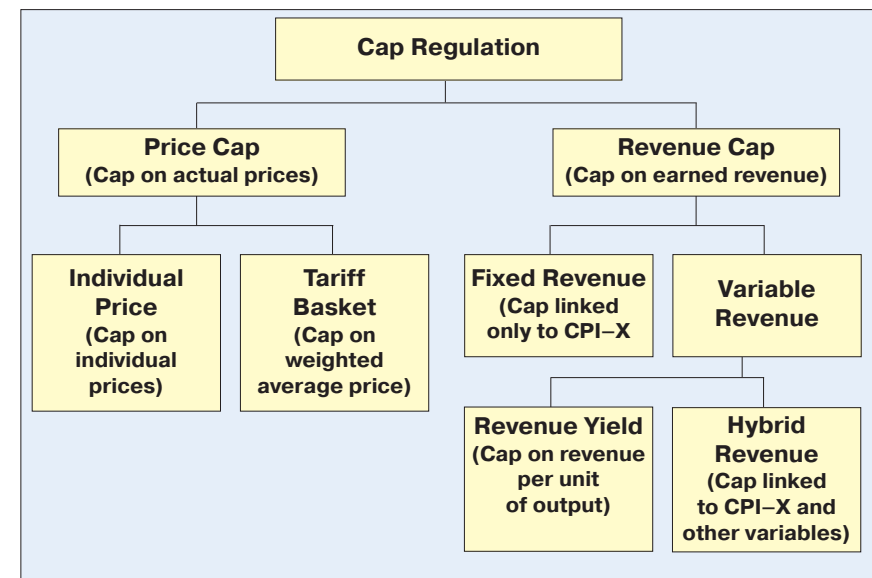


Figure 2. Forms of cap regulation

sustainable over several years. In analogy, regulated utilities are allowed to maintain "intermediate profits" gained through exceeding the productivity improvement target for periods of between three to five years by instituting a corresponding lag between "regulatory reviews."

A number of forms of cap regulation exist (see Figure 2). They are separately analysed in the following subsections.

10.5 Price Cap Regulation

Under price cap regulation, a restriction is imposed on the growth rate of average revenues or particular prices.⁵⁴ Usually, average revenues are indexed to some inflation indicator (e.g., the RPI or the CPI). In addition, regulated utilities are obliged to reduce their prices each period according to some assumed productivity growth rate (so-called "X-factor") that is to be determined by the regulator.

The starting level and the development path of prices are fixed such that under normal conditions, the utility will earn a "fair" rate of return in each period. This implies the imposition of a time path of price ceilings for a period usually extending over several years according to the following formula:

$$P_t = (1 + \text{CPI}_t - X_t)P_{t-1}$$

where:

P — price cap based on projected data for the regulated services for year t price;

CPI — annual percentage change in the Consumer Price Index;

X — productivity growth rate; and

t — time index.

Prices may be controlled individually (individual price caps) or grouped into "baskets" (tariff basket price caps).

10.5.1 Individual Price Caps

With individual price caps, the regulator sets the upper limit for each individual price. This is the most direct form of price control but its application is limited to situations where the number of services provided is small and stable and costs are easily identifiable.

Under a price cap, if the actual quantity demanded is greater or less than that forecast, then revenue will increase or decrease, respectively. The effect of this change in revenue on the firm's profits, and rate of return on capital, will depend on how its costs change with the change in volume. If the average cost of producing the service does not change with a change in the volume sold, then under a price cap, total costs will move in proportion with total revenue and there will be no change in the firm's profit and return on capital. On the contrary, if the average cost falls with a decrease in demand, then under a price cap, total revenue will increase more than total costs, if the volume sold increases and there will be an increase in the firm's rate of return on capital. The opposite will happen if there is a fall in demand.

Hence, the appropriateness of the price cap, as opposed to other forms of the cap, for a utility will depend on the nature of its cost function.

10.5.2 Tariff Basket Price Caps

The alternative is an average price control, or "tariff basket." With these, prices are grouped into baskets on the basis of the services to which they apply. A representative weighted average price for the basket is calculated. An upper limit or cap is then applied to the weighted average price. This allows for some variation between the changes that can be made to individual prices within the basket provided that the weighted average price for the basket is within the regulated cap.

The weights are typically based on either the revenue or quantity shares of each service in the basket. These weights may be fixed at the start of the regulatory period and then held constant throughout the period, or alternatively, reset at suitable intervals.

Under a tariff basket, the limit on allowed price increases is expressed in terms of a ratio of 'notional' revenues. The tariffs proposed by the regulated service providers should satisfy the tariff basket constraint if this ratio is less than a given cap. This cap is, in turn, determined on the basis of a CPI-X formula. The form of tariff basket is set out in the equation below:

$$1 + \text{CPI}_t - X_t \geq \frac{\sum_{i=1}^n \sum_{j=1}^m p_{ij}^{t+1} q_{ij}^N}{\sum_{i=1}^n \sum_{j=1}^m p_{ij}^t q_{ij}^Q}$$

where the regulated service provider has n tariff categories, which each have up to m components;

- p_{ij}^t — price being charged in year t for component j of tariff i ;
- p_{ij}^{t+1} — proposed price for component j of tariff i in the coming year $t + 1$;
- q_{ij}^N and q_{ij}^O — forecasted or historic quantities of component j of tariff i that will be used in the price control (N indicates that quantities apply for the new year and O that quantities apply for the old year);
- CPI_t — annual percentage change in the Consumer Price Index; and
- X_t — percentage figure determined by the regulator and reflects generally the productivity improvement requirements.

Under the presented tariff basket formula, tariffs are approved on the basis of an allowed ratio between two notional revenues: one based on proposed prices and the other on current prices. However, the actual revenue earned by the regulated network business will depend upon the actual tariffs applying to the actual quantities sold of each of the charging parameters. If the regulated network business sells an additional unit, of whatever charging parameter (kWh, kW or customer number), the marginal revenue it earns will be equal to the tariff applying to that extra unit. As a result of this link between marginal revenue and tariff structure, the proposed tariff basket form of control provides the regulated service providers with an incentive to set their tariff structures to reflect the underlying cost structure, in order to minimise profit risk.

Proponents of a tariff basket argue that it can give a less intrusive form of regulation with lower costs of regulation and strong incentives for efficiency. However, controls, like these that are based on existing price structures, can have the effect of limiting price or product development. As explained above for a single price cap, the appropriateness of a tariff basket as a form of price control will depend on the extent to which costs move in proportion to a change in volume. For example, if costs move in proportion to changes in volume, then there will in general⁵⁵ be no change in the utility's rate of return, if there is change in volume from that forecast, assuming there is no factor compensating for the changes. However, if the costs do not move in proportion to volume,⁵⁶ the utility will face a risk of a fall in rate of return if volumes decrease and the benefit of a rise in return if volumes increase.

10.6 Revenue Cap

Under "revenue cap" regulation, a restriction is imposed on the growth rate of total revenues.⁵⁷ Total revenues can either be fixed in nominal terms over the regulatory period, or a mechanism similar to that used under "price cap" regulation can be used: total revenues are indexed to some inflation indicator (e.g., the RPI or the CPI). In addition, the regulated companies are obliged to reduce their total revenues each period according to some assumed productivity growth rate. In the simplest form, this implies the imposition of a time path of total revenue ceilings for a period usually extending over several years according to the following formula:

$$R_t = (1 + CPI_t - X_t)R_{t-1}$$

where:

- R_t — revenue cap based on projected data for the regulated services for year t ;⁵⁸
- CPI_t — annual percentage change in the CPI;
- X_t — productivity growth rate; and
- t — time index.

With price caps, the limit is applied to the actual price or an average of the actual prices charged. With a revenue cap, the upper limit is applied to the revenue earned. Like price caps, revenue caps can be applied to segments of the regulated business rather than to the businesses a whole. For example, separate revenue caps may be established for household and business customers, or for different categories of service.

The different types of revenue cap are explained in the following subsections.

10.6.1 Fixed Revenue Cap

Under a fixed revenue cap, allowed revenues are set equal to target revenues for the duration of the review period and are adjusted each year only for general price inflation and any assumed productivity increases. This is a relatively simple form of control, with low implementation and administration costs. It provides the regulated service providers with guaranteed revenue, regardless of the volume or level of service, it is called upon to provide to customers. If in any year, customer demand varies from the level expected at the time that the target revenues were set, the actual level of revenue collected will either exceed or fall short of

the cap. A correction mechanism will then return the surplus to customers or make up the shortfall.

Under this control method, the impact of unexpected variations in costs is borne by the regulated service providers, at least until the next review. In the circumstances where this may create an incentive for more prudent cost risk management, it accords with the general objectives of incentive regulation. However, when variations in costs are the result of unexpected changes in customer demand, there are potentials for perverse incentives to be created, with the risk of unintended consequences. During the establishment of revenue targets, regulated service providers will have an incentive to inflate their estimates of demand growth, as a way of minimising the risk of higher-than-expected growth depressing their profits. Once revenue targets have been set, there will be an incentive for regulated service providers to minimise all the costs depending on the units transported and number of consumers. This may lead to postponing new connections and/or network infrastructure investments resulting in unsatisfied customer demand as well as a general decline in service levels. There is some potential for the difference between actual revenues and the level warranted by service volumes to increase progressively towards the end of the control period, adding to the risk of demands for regulatory intervention and interperiod price adjustment problems. Unless it can be argued that the probable demand forecast error for the network service cost drivers will be small, fixed revenue caps appear to carry material financial risks for regulated service providers.⁶⁰ As a consequence, there is a likelihood of introducing regulatory distortions into the provision of services.

10.6.2 Variable Revenue Cap

(1) Revenue Yield Cap

Using this kind of cap, the regulator calculates the maximum average revenue by dividing the total revenue by total output of the regulated service provider. It sets a cap on the maximum average revenue each regulated service provider is allowed to earn per unit of output (that is, its maximum average charge per kWh of electricity distributed) for the first year of the regulatory period. It also determines a CPI-X formula for adjusting this average revenue per unit in each subsequent year of the period, and can set limitations on the amount by which some or all individual prices (within the tariff system) can increase. The regulated service provider is able to rebalance prices (increase or decrease individual prices), so long as the weighted average of its distribution prices does not exceed the set maxi-

imum average charge per kWh, and it complies with any limitations imposed by the regulator. The level of the revenue yield control is set on the basis of the benchmark revenue established by the regulator, together with a forecast of the expected volume of output. If the volume of output turns out as expected, the regulated service provider will receive 100% of its revenue. Where regulated service providers sell in excess of the volume expected, their revenue will be in excess of the benchmark revenue. Conversely, if they sell less than the expected volume, their allowed revenue falls below the level anticipated when the price control was set.

Under a revenue yield cap, there is an inherent incentive to expand volumes and this might be in conflict with demand management initiatives. The regulated service provider will have a strong disincentive to participate in such programs, as it reduces the revenue they are allowed to earn under the price control. Regulated service providers may therefore fail to consider demand management initiatives, even when it would be economically efficient to do so.

Under the revenue yield approach, the allowed average revenue need not be systematically related to costs, since it is based on a single measure of quantities to proxy services that in reality have many different dimensions of output and unit costs. As output increases for one service, allowed revenues increase proportionately, whereas costs may not. This upholds the incentive for the regulated service provider to expand volumes by the cheapest means possible.⁶¹

Finally, under the revenue yield approach allowed revenue is closely tied to the quantity transported, and therefore the regulated service provider bears considerable revenue risk if load growth is significantly different from forecast. Costs, on the other hand, are unlikely to match movements in load so closely, given that network services tend to be characterised by a high proportion of fixed costs. The regulated service provider is therefore likely to bear a significant amount of financial risk, associated with the volatility in profits.

(2) Hybrid Revenue Cap

Effectively, in an average revenue yield cap regime, the profit increase due to volume effects is a result of lagged regulatory adjustment. An automatic adjustment mechanism may then prevent excessive profits from being earned between regulatory reviews.⁶² In a pure revenue cap regime, distortions could occur when customer-driven changes in costs cannot be recovered by the regulated service providers.

Given the weakness of revenue yield cap and pure revenue cap, hybrid revenue cap regulation schemes have been developed. Under such regimes,

the allowed revenues are linked by a predetermined formula to another variable or group of variables (in addition to CPI-X), for example, level of demand, selected cost drivers or performance measures. Allowed revenues are therefore not known in advance, but are determined by the formula once the values of the independent variables are known.⁶³ In practice, regulators typically use a fixed revenue component combined with annual revenue drivers that may be based on service volumes, cost or performance-related variables. For example, customer numbers, energy consumption, energy demand, length of network lines and system losses are among the variables included in hybrid cap control formulas used by regulators. The major advantage of such price control schemes is that they sustain the link between cost and regulated revenues during the regulatory period and mitigate the risk exposure for regulated service providers.

An example of a variable revenue cap formula is given below:

$$R_t = R_{t-1} \left(\alpha \frac{Q_t}{Q_{t-1}} + \beta \frac{C_t}{C_{t-1}} \delta \right) (1 + \text{CPI}_t - X_t)$$

where:

R_t — revenue cap based on projected data for the regulated services for year t ;

α — "quantity weighting factor." This represents the proportion of total required revenue that is assumed to vary with a change to the quantity of units distributed in the case of the network business. (In more specific terms, it equals the proportion of total required revenue that is allowed, with regard to the revenue cap, to increase by the same percentage as the percentage change to the annual quantity of units distributed/sold.);

Q_t — projected quantity of units distributed/sold in year t (kWh);

β — "Customer weighting factor." This represents the proportion of total required revenue that is assumed to vary with a change to the number of customers of the business concerned. (In more specific terms, it equals the proportion of total required revenue that is allowed, with regard to the revenue cap, to increase by the same percentage as the percentage change to the number of customers.);

C_t — forecast number of customers in year t ;

δ — "Residual Factor." This represents the proportion of total required revenue that it is assumed does not vary with a change to the quantity of units sold or the number of customers. In more specific terms, $\delta = 1 - \alpha - \beta$;

CPI_t — annual change in the CPI; and

X_t — productivity growth rate.

The hybrid price control forms reduce the economic distortions resulting from pure forms of revenue and price cap control. The weakened link between volume and allowed revenue implies a reduced incentive on the regulated service provider to bias prices in favour of low marginal cost services in order to expand volumes. It also lessens the bias against participation in demand management programmes as an alternative to network investment. However, by retaining a link between allowed revenue and output, a hybrid control avoids the disincentive present under a pure revenue cap to encourage new connections and, potentially, to artificially discourage load growth.

In the absence of competition in the Russian electricity networks, the FEC should protect customers by acting as a proxy for normal competitive forces. A good regulatory regime should provide companies with similar opportunities and incentives to those they would face in a competitive market. However, in mimicking market forces, the FEC also needs to balance this duty with the need to protect the interests of network owners, in order to ensure that the Transmission Operator (FNC) earns a reasonable rate of return on their efficient investments.

To carry out this role, the FEC needs to monitor the companies' performance. It is preferred to adopt a light-handed approach to doing this. One of the reasons to propose this approach is that it will provide incentives for the FNC to focus its attention on delivering the best service to customers, rather than debating costs with the regulator. The second important reason is that it is inappropriate for the regulator to micro-manage the network service providers. The external position of the regulator, the lack of operational business competence and the information asymmetry between regulated service providers and regulator predetermine the unsuccessful outcome of regulatory micro-management policy.

The CPI-X framework includes a number of price control options. The international experiences demonstrate no clear trends towards one preferred solution (detailed discussions are contained later in this report). The current price control over NGC in England & Wales is based on pure revenue cap (no revenue driver is included). National Grid Company preferred this method and argued that other form of variable revenue caps (with explicit adjustment of allowed revenue) would boost demand and introduce an undesired bias against demand side management. Pure revenue caps are applied also by ACCC (the federal regulator in Australia) to the transmission service providers in Victoria and New South Wales.

Differently, the Norwegian regulator applies a variable revenue cap (revenue driver is energy) to regulate Statnet. The Netherlands applies tariff basket approach for TenneT, however, with predetermined fixed output (quantity) levels. Similarly, the Energy Agency of Slovenia imposed at the beginning of 2003 a tariff basket approach on the Slovenian TSO ELES.

One option would be to apply revenue cap for transmission. Revenue cap is a form of economic regulation with low regulatory involvement; i.e., it supports the light-handed regulation approach. One additional important advantage of revenue caps is that they support revenue stability; i.e., the revenue does not fluctuate with changes of demand. For the last reason, revenue caps are sometimes preferred by the practitioners for the purposes of transmission price control. The second option is to use a tariff basket cap. The advantages of the tariff basket include: strengthening links between marginal revenue and tariff structure; keeping the link between revenue and costs; minimising complexity and absence of forecasts and correction factors.

Both of the discussed options (revenue cap and tariff basket) will give the regulated FNC the freedom to manage the structure and level of their prices. This freedom has advantages, but it also brings the risk that the company may change the prices in a way that disadvantages some consumer groups. In comparison with the distribution network businesses, the risk of tariff rebalancing is lower as the FNC provides transmission service only to large users (either distribution companies) or large industrial users.⁶⁴ If the FEC feels that such risk could occur in the first regulatory period, subsidiary limitations of annual increase of the network charge components could be imposed in order to avoid negative imbalance effects.⁶⁵

10.7 Scope of the Regulatory Caps

The philosophy proposed requires that the FEC regulates the price basket or revenue, based on the premise that the regulated charges will bring sufficient revenues to cover the costs of provision of transmission network service (i.e., on revenue/prices of transmission use of network charges) including the efficient O&M and capital costs⁶⁷ of the regulated transmission network business for provision of the transmission network service plus the cost of the purchase of energy to cover transmission network losses. The transmission connection service (connection charges) will be regulated as an excluded service⁶⁸ (i.e., outside of the proposed caps) and based on the cost of connection including reasonable rate of return on

transmission connection assets. All charges collected for provision of non-regulated services provided by the FNC are not subject to regulatory price control and will be excluded from the revenue cap.⁶⁹

10.8 Quality of Supply

All types of price regulation face a fundamental problem. In the case of rate-of-return regulation, utilities are generally free to define their own investments and quality levels. In line with economic theory, this tends to create incentives for overinvestment in both assets and quality. Not surprisingly, many regulatory regimes therefore focus on preventing this type of inefficiency and to avoid sub-optimally high levels of investment. Simple types of cap regulation, on the other hand, may allow a regulated company to reduce its cost by reducing its quality of supply by cutting investment, maintenance, or personnel with the aim of increasing profits. Consequently, price regulation may thus also provide incentives for underinvestment in electricity networks.

Quality of service is just as important to consumers as prices. If standards of service fall but prices remain the same, consumers are effectively suffering an increase in prices. In competitive industries, dissatisfied customers will then either demand lower prices or switch suppliers. Likewise, investors will be less willing to invest if they believe that companies are investing too much or too little in service standards. A monopolistic firm, on the other hand, may try and collect the allowed revenue while reducing product and/or service quality.

Removing the link between prices and costs creates strong incentives for efficiency improvement as any cost savings directly translate into higher profits. However, both theory and practice suggest that, without additional quality regulation measures, these incentives eventually lead to perverse quality degradation. Even though this may cause additional cost, e.g., for network users, the network operator may still find it more profitable to cut costs at the expense of quality. Thus, under price/revenue cap systems, inclusion of quality regulation elements is imperative. Different authors have recognised that strong cost reduction incentives (driven by the profit motive) result in substandard quality levels. In the telecoms industry, it is worth noting that in 1996, the Oregon Public Utility Commission terminated its price cap regulation plan for US West due to quality decline.

More theoretically, the economic problem is that from the company's point of view, quality is an external effect. This means that the company's private choices regarding cost and quality do not necessarily lead to the most

desirable outcomes from a social welfare standpoint. Regulatory authorities must therefore act to protect consumers' interests by devising incentives to reach socially optimal outcomes. In line with economic theory this will be achieved if the externality is "internalised," i.e., any outcomes of the company's cost and quality choices are made a function of the resulting societal losses.

In short, price regulation generally has to be accompanied by some kind of regulation of quality of supply (for transmission and distribution networks), with the aim of both avoiding distorted and excessive investment and of preventing a decrease in quality and performance standards. Federal Energy Commission should provide regulated service providers with incentives to increase efficiency and reduce costs, in general to at least maintain present levels of quality and, where applicable, aim at a better cost/quality trade-off.

At present, companies only monitor average network performance, using criteria like frequency and duration of interruptions. Generally, such monitoring of average network performance is appropriate under present circumstances. It is the most widely adopted technique for monitoring supply reliability and enables companies' performance to be compared with each other, and with performance in other countries. Federal Energy Commission should require annual reporting from FNC on all transmission system incidents that involve loss of supply to customers (including downstream distribution companies). Because such failures will be infrequent, reporting can relate to each individual incident. In addition, FEC might require that FNC should report on other quality factors for its network, even though these do not relate to loss of supplies. These include the extent to which parts of the transmission network are unavailable for service during the reporting year, and the reasons for such unavailability. Federal Network Company should also monitor voltage and frequency on its system.

11 Recommendations

11.1 General Design of Transmission Service Charges

- Transmission service charges should apply to the networks owned by the FNC, i.e., for 220 kV lines and higher.
- Transmission service charges should be made up of transmission connection charge and transmission use of network charge.
- Short-run signals (in terms of time and location) will be provided by application of nodal pricing on the wholesale market.
- Transmission loss rentals and congestion rentals (surpluses) resulting from the application of nodal pricing should be used to scale down the transmission use of network charges.

11.2 Revenue Requirements

- The calculation methodology should be designed to remunerate the FNC's reasonable costs and to provide an adequate return on assets to ensure that capital remains in the industry and that new capital can be attracted to the sector as and when it is necessary.
- The revenue requirements of the FNC should cover:
 - Operation and maintenance costs: these are the costs incurred by the FNC in operating and maintaining the core transmission network to the Russian technical standards and the requirements of the Technological Rules;
 - Depreciation costs: these are the costs related to depreciation of the investment in the core transmission network assets (excluding transmission connection assets) during their life in order to accumulate sufficient capital for asset replacement (transmission connection will be discussed later in this document); and
 - Return on assets: the FNC will be allowed to earn a return on its assets.
- Due to the fact that the FNC is a monopoly business, its revenue/prices will be regulated by the FEC.

11.3 Transmission Use of Network Charge

- The transmission use of network charge will apply to the transmission network infrastructure reflecting its ability to transport energy in a reliable way and will recover the costs incurred by the FNC in order to ensure secure operation and regular maintenance of the network assets (costs related to O&M, depreciation and financial return on the core transmission network assets), however, reduced by the surplus resulting from the application of nodal pricing;
- The transmission use of network charge is set irrespective of the distance transported but taking into account the entry and exit points on the respective transmission regions and network voltage level;
- As nodal prices will provide locational signals, imposing secondary signals through locational pricing for the transmission network infrastructure is not of immediate priority. Hence, it is suggested to allocate the transmission infrastructure costs (revenue requirements) of each transmission region to the respective voltage level operated by the FNC in this region using the postage stamp approach and cost cascading among the voltage levels;
- If secondary locational long-run signals are still preferred by the FEC and/or the FNC, the short-run signals resulting from the application of nodal pricing can be supplemented by locational pricing (e.g., on a zonal basis) for the transmission network infrastructure. In this case, a portion of the revenue requirements should be allocated on a locational basis to the transmission network nodes and aggregated in price zones;
- If locational pricing for the transmission network infrastructure is adopted, it is not recommended to allocate the total revenue requirements on a locational basis but rather to define a percentage of the total revenue requirements (e.g., 50%) that should be allocated on a locational basis. The remaining part will be charged pro-rata to the transmission service users;
- Should the FEC and/or the FNC prefer to introduce locational pricing for the transmission network infrastructure, it is suggested allocating the revenue requirements jointly to the generation and load nodes, the prevailing part to be charged to the load nodes (e.g., 25% to the generators and 75% to the load). If the revenue requirements are allocated by means of a postage stamp concept, it is reasonable to allocate the payment liability entirely to the load entities;

- All transmission service users located in the same transmission region and connected to the same voltage level (or located in the same price zone within the transmission region if locational pricing applies) will pay the same transmission use of network charges;
- In order to improve efficiency signals, the transmission use of network charge can be differentiated by time of use (time of day or season of year) whereas the allocation rules could use the probability not to meet the maximum demand in the transmission network or simple criteria allocating the transmission infrastructure cost to time periods where the demand is higher than a predefined threshold;
- Two options for design of the transmission use of network charges are suggested: (1) peak-based regime, where the peak usage is appropriately defined so as to minimise the attractiveness of inefficient charge avoidance (e.g., number of peak hours or peak time period); and (2) hybrid tariff structure using energy, demand and fixed charges;
- The definition of peak usage for generators could refer to their rated capacity (maximum net capacity that can be delivered to the transmission network under normal conditions of operation). The definition of peak usage for load entities could refer to their observed load during the settlement period in which system peak load occurs and some number of settlement periods around the peak load or the connected/contracted load.

11.4 Transmission Connection Charge

- The transmission connection charge should cover the cost of transmission connection assets (shallow concept) without taking into consideration any reinforcement requirement in the transmission network infrastructure;
- The transmission connection charge should cover the costs of spurs between generating stations and the core transmission network;
- The transmission connection charge should be paid by each transmission service user connected to the federal transmission network and should reflect the individual cost of the transmission connection assets;
- The transmission connection charge should be set to recover the reasonable connection costs (including capital and O&M costs) including a return on the transmission connection assets. The transmission connection charges will be subject to regulatory control;

- The transmission connection charge should be denominated in RUB/year and paid by each transmission service user connected to the transmission network;
- The transmission connection charge could include a termination payment (in case that transmission service users connected to the transmission network wish to be disconnected) that should cover the cost of transmission connection assets still not covered till the date of termination and any additional costs resulting from the disconnection.

11.5 Transmission Price Control

- The regulatory framework must be structured in a way that does not discourage certain sources of finance and investment. The analysis of the credibility of the political and economic environment and the credibility of the regulatory framework are closely interlinked. Investors will only be interested in an engagement in the power industry if there is a transparent regulatory process that is credible and time-consistent. This credibility is largely dependent upon the perceived independence of the regulator;
 - Under conventional rate of return regulation rates of regulated service, providers are reviewed on a regular basis and have to be adjusted (process called "regulatory review") to lower levels if cost savings have been achieved in the interim (since the last review). Regulated service providers thus only benefit from cost savings to the extent that regulatory reaction to cost savings is lagged (this lagged reaction is called regulatory lag). Additionally the rate of return regulation creates overcapitalization (overinvestment in fixed assets) incentives and potentials for increase of capital costs;
 - Under cap regulation, a restriction is imposed on the growth rate of prices (basket of prices or individual prices) or revenue. The regulatory cap formulas include also indexes of some inflation indicator (e.g., the RPI or the CPI). In addition, regulated service providers are obliged to reduce their prices (price cap) or revenues (revenue cap) each period according to some assumed productivity growth rate that is to be determined by the regulator;
 - Decoupling price from costs under cap regulation creates strong incentives for efficiency improvement as any cost saving directly translates into higher profits. However, both theory and practice suggest that, without additional quality regulation measures, these incentives eventually lead to perverse quality degradation;
- A number of alternative options for cap control were reviewed and two were preselected: revenue cap and tariff basket cap. Both of them are widely used in international practice. Revenue cap is a form of economic regulation with low regulatory involvement where a limitation is imposed on the revenue earned for provision of regulated services. Under tariff basket cap regulation, a limitation is imposed on a basket composed by the weighted (by means of quantities) tariff components charged for provision of regulated services. The important advantage of the revenue caps is that they support revenue stability; i.e., the revenue does not fluctuate with changes of demand. For this reason, the revenue caps are frequently preferred for the purposes of transmission price control. The advantages of the tariff basket cap include: strengthening links between marginal revenue and tariff structure; keeping the link between revenue and costs; minimising complexity and absence of forecast and correction factors. From the two pre-selected options, the revenue cap is preferred;
 - The philosophy proposed requires that the FEC regulates transmission activities, based on the premise that the regulated charges will bring sufficient revenues to cover the costs of provision of transmission network service (i.e., on revenue/prices of transmission use of network charges) including the O&M and capital costs of the regulated transmission network business of the FNC. The transmission connection charges will be regulated as an excluded service (i.e., outside of the proposed cap) and based on the cost of connection including a reasonable return on the transmission connection assets. All charges collected for provision of non-regulated services provided by the FNC, if any, are not subject to regulatory price control and will be excluded from the regulatory cap imposed on the transmission network service provision;
 - Partial decoupling of revenue/price from costs under cap regulation creates strong incentives for efficiency improvement as any cost saving directly translates into higher profits. However, both theory and practice suggest that, without additional quality regulation measures, these incentives may eventually lead to quality degradation. Therefore supplementary measures for quality control will be necessary. The FEC should require annual reporting from the FNC on all transmission system incidents that involve loss of supply to customers (including downstream distribution companies). Because such failures will be infrequent, reporting can relate to each individual incident. In addition, FEC might require that FNC should report on other quality factors for its network, even though these do not relate to loss of supplies. These

include the extent to which parts of the transmission network are unavailable for service during the reporting year and the reasons for such unavailability;

- As the cap regulation is a new form of price control for Russia, it is recommended to start with a shorter regulatory period, e.g., three years. This will give the FEC the opportunity to balance the incentives and the uncertainties resulting from the lack of experience and reporting of data. The regulatory framework must be structured in a way that does not discourage certain sources of finance and investment. The analysis of the credibility of the political and economic environment and the credibility of regulatory framework are closely interlinked. Investors will only be interested in an engagement in the power industry if there is a transparent regulatory process that is credible and time-consistent. This credibility is largely dependent upon the perceived independence of the regulator. Lack of transparency and clarity on timing and conditions for application of alternatives might increase significantly the regulatory risk in the Russian power sector. The creation of a credible and time-consistent energy regulation is an important factor for the future successful market reform evolution and attraction of investor capital in the Russian power industry.

11.6 Transmission Capacity Rights

If transmission capacity rights will be applied in Russia, the following should be taken into consideration:

- The design of the transmission capacity rights should reflect the detailed features and properties of the wholesale market;
- The conceptual framework for transmission capacity rights should be developed and co-ordinated with evolution of the wholesale electricity market in order to mitigate the risks in the transition period;
- The existing system users should be given the opportunity to receive transmission capacity rights. They could receive transmission capacity rights "free of charge," however, following strictly defined rules and procedures (including definition of volume and lengths of rights);
- In the case of new market participants (who expect significant transmission congestion), transmission capacity contracts could be obtained from the holder(s) of existing transmission congestion contracts if there are such for sale between the two points. In the case of new investment in the transmission networks, new transmission capacity rights will be issued.

12 List of Acronyms and Definitions

ACCC	Australian Customer and Competition Commission
ATS	Administrator of Trade System
CEE	Central and Eastern European Countries
CPI	Consumer Price Index
CRNP	Cost-Reflective Network Pricing
FEC	Federal Energy Commission
FNC	Federal Network Company
FOREM	Federal Market for Energy and Capacity
GenCo	Generation Company
ICRP	Investment Cost Related Pricing
IPART	Independent Pricing and Regulatory Tribunal
LRMC	Long-Run Marginal Cost
NCIS	Network Configuration Information System
NECA	Australian National Electricity Code Administrator
NEM	National Electricity Market
NGC	National Grid Company
O&M	Operation and Maintenance
PBR	Performance-Based Regulation
RNC	Regional Network Company
RPI	Retail Price Index
RPP	Regional Power Pools
SO	System Operator
SRMC	Short-Run Marginal Cost
SupCo	Supply Company
TSO	Transmission System Operator
WACC	Weighted Average Cost of Capital

13 References

Arbeitsgruppe Netznutzung Strom der Kartellbehörden des Bundes und der Länder (2001): Reichweite der kartellrechtlichen Eingriffsformen für die Überprüfung der Höhe der Entgelte für die Nutzung der Stromnetze.

Baumol, W. J. / Bradford, D. (1970): Optimal Departure from Marginal Cost Pricing, in: American Economic Review, Vol. 60, No.3.

Averch, H. / Johnson, L. (1962): Behavior of the Firm under Regulatory Constraint, in: American Economic Review, Vol. 52.

Blankart, B. Ch. (1980): Ökonomie der öffentlichen Unternehmen, eine institutionelle Analyse der Staatswirtschaft, München.

Blanchard, O. J. / Fischer, S. (1989): Lectures on Macroeconomics, MIT Press, Cambridge MA, P. 478–488.

Brealey, R. / Myers, S. (1991): Principles of Corporate Finance, Fourth Edition, McGraw-Hill, Inc., New York.

Buchner, J. (1997): A Framework of Transmission Pricing in the Dutch Electricity Supply Industry — Issues, Options, Suggestions, Project for the Dutch Economic Ministry, KEMA–ECC Europe B. V., Maastricht, January.

Conference Proceeding of the VNIE Conference on Network Losses and Measures for Reduction of Network Losses (November 2002).

Dorofeev V.V. / Mirolyobov V.A. (2002): New Experience in Creation of Competitive Power Market in Russia and Consequences for the Development of the United Power System in Russia, CIGRE.

DTe (2000): Guidelines for Price Caps Regulation in the Dutch Electricity Sector in the period from 2000 to 2003.

Eggen, E. / Buchner, J. / Petrov, K. / Giesbertz, P. / Fredriksen, O. / Ajodhia V. (2001): Micro and Macro Benchmarking Methods in Conference Proceedings of Distributech Conference, Berlin, November.

Federal Law on State Regulation of Electricity and Heat Prices in Russian Federation (Федеральный закон "О государственном регулировании тарифов на электрическую и тепловую энергию в Российской Федерации").

Haubrich, H.-J. / Swoboda, P. (1998): Grundlagen für Systemnutzungs-/tarif/-regelungen im Bereich der öffentlichen Elektrizitätsnetze, Wissenschaftliches Gutachten im Auftrag des Bundesministeriums für wirtschaftliche Angelegenheiten der Republik Österreich.

IPART, Independent Pricing and Regulatory Tribunal of New South Wales (1999): Regulation of Network Service Providers, Price Control Issues and Options, Discussion Paper 34, March.

Jamaasb, T. / Pollit, M. (2000): Benchmarking and Regulation of Electricity Distribution and Transmission Utilities: Lessons from International Experience, University of Cambridge, Draft Paper.

Joskow, P. / Schmalensee, R. (1986): Incentive Regulation for Electric Utilities, in: Yale Journal of Regulation, Vol. 4, No.1.

Kahn, A. E. (1971): The Economics of Regulation: Principles and Institutions, Vol. I, II, John Wiley and Sons, Inc., New York.

Liston, C. (1994): Price-Cap Versus Rate-of-Return Regulation, in: Journal of Regulatory Economics.

Order N 226 of the Russian Government "Pricing of heat and electrical energy" from 02.04.2002 (Постановление Правительства РФ "Ценообразование в отношении электрической и тепловой энергии" от 2 апреля 2002 г. № 226).

Order N 49-э/8 of Federal Energy Commission (FEC) on "Methodological guidelines on establishment of electricity and heat prices on retail market" from 31.06.2002 (Постановление ФЭК РФ N 49-э/8 "Об утверждении методических указаний по расчету регулируемых тарифов и цен на электрическую (тепловую) энергию на розничном (потребительском) рынке от 31 июля 2002 г.").

Order N 72-э/3 of Federal Energy Commission (FEC) on "Methodological guidelines on establishment of transmission service pricing for national electricity network" from 23.10.2002 (Методические указания по расчету размера платы за услуги по передаче электрической энергии по единой национальной электрической сети, утверждены Постановлением Федеральной энергетической комиссии Российской Федерации от 23 октября 2002 г. N 72-э/3).

Perez-Arriaga, I. J. / Rubio, F. J. (1998): Marginal Pricing of Transmission Services: An Analysis of Cost Recovery, in: Electricity Transmission Pricing and Technology, edited by M. Einhorn, R. Siddiqi.

Petrov, K. (2001): Establishment of Cap Regulation in Electricity Networks — 1st Balkan Power Conference, Bled Slovenia, September.

Petrov, K. (2001): Advance Methods for Transmission Pricing — Conference Power-Gen for Asia, Kuala Lumpur Malaysia, September.

Ramsey, F. (1927): A Contribution to the Theory of Taxation, in: Economic Journal, Vol. 37.

Rudnick, H. / Palma, R. / Fernandez, J. (1995): Marginal Pricing and Supplement Cost Allocation in Transmission Open Access, in: IEEE Transactions on Power Systems, Vol. 10, No.2, May.

Tutton, T. (1999): Rolling Forward the Regulatory Asset Base in the UK Electricity and Gas Industries, Report of PriceWaterhouseCoopers.

Sharkey, W. (1982): The Theory of Natural Monopoly, Murray Hill, New Jersey.

Fischer, R. / Serra, P. (2000): Regulating the Electricity Sector in Latin America, Economia, Fall.

Vickers, J. / Yarrow, G. (1991): The British Electricity Experiment, in: M. Einhorn (Hrsg.): From Regulation to Competition: New Frontiers in Electricity Markets, Kluwert Academic, Dordrecht et al.: P. 39–71.

Wholesale Market Design in Russia (2002): Document Developed by "Carana."

14 Acknowledgements

In carrying out this project and the production of this brochure, KEMA Consulting are indebted for the help and assistance given by the Electric Power Research Institute of Russia, VNIIE, who were partners in the project and in particular to Valery Voronitzky. They also acknowledge the assistance of FNC, and in particular from Roman Berdnikov, and many other organizations that make up the Russian Power Sector and the staff in these organisations. Lastly to the European Union-Delegation of the European Commission in Russia who not only funded the project and this brochure under their Tacis initiative but also for the support and help given by their team in Moscow.

15 Author's biography

Dr. Konstantin Petrov is an energy economist and electrical engineer with a major expertise in the area of market design, regulatory economics and pricing. He received his Ph.D. degree from the Institute of Energy Economics at the University of Cologne (Germany). His major assignments in the areas of restructuring and regulation include: pricing and economic regulation, development of market/industry design as well as applied modelling in energy sector. Dr. Petrov has been involved in a number of projects in more than 15 countries. From 2002 to 2004, he advised the Federal Network Company of Russia on a number of issues concerning transmission network access pricing, transmission price controls, wholesale market monitoring and regulatory interfaces.

16 End notes

- 1 See Dorofeev V.V / V.A. Mirolyobov (2002): New Experience in Creation of Competitive Power Market in Russia and Consequences for the Development of the United Power System in Russia, CIGRE.
- 2 The Russian wholesale market design does not exclude bilateral transactions. With changing supply and demand conditions, generators and customers will see fluctuations in short-run prices. When demand is high, more expensive generation will be employed, raising the equilibrium market prices. When transmission constraints bind, congestion costs will change prices at different locations. Even without transmission congestion constraints, the spot market price can be volatile. This volatility in prices presents its own risks for both generators and customers, and there will be a natural interest in long-term mechanisms to mitigate or share this risk. These realities have been recognised in the new wholesale electricity market design in Russia by allowing bilateral transactions. Such transactions would need to consider then the SRMC of transmission through an ex-post charges that equals the opportunity cost of the transaction, i.e., the cost difference of selling the power to the pool at the injection node price and buying it back at the withdrawal node price. Thus, the charge for SRMC of transmission between any pair of nodes is equal to the nodal price difference between the nodes.
- 3 Generally, investments in networks are characterised by economies of scale and scope and are lumpy, so that "overbuilding" is often the least cost option. Loop flows arising from Kirchoff's Law in effect result in (positive or negative) externalities, since particular flows will have knock-on effects elsewhere in the network. The implications of the above features for network pricing are that: (1) charges that are based on marginal costs will not recover the fixed costs of the network; and (2) it is often not possible to attribute specific assets to particular users, since electricity flows follow multiple routes along the path of least resistance, according to Kirchoff's Law.
- 4 System services are necessary for the assurance of a safe, reliable and efficient transport of electricity. The prevention and remedying of large-scale outages by keeping reserve and balancing power and maintaining black start facilities that are required for the start-up of the facility after a total or partial outage, should also be regarded as falling under this heading. System services are managed and provided by the system operator. In order to provide these services, the SO needs to purchase ancillary services that include: frequency response, voltage and reactive power control and black start. The financial effect for the SO is that it will incur some costs of purchasing of these services and will need to cover them via a separate system service charge from the parties using the system services. It should be noted that the SO will incur cost not only for purchasing ancillary services and payments for system constraints but also their own costs for system control. The sys-

tem control costs comprise assets cost related to the control centre (hardware, software), communication devices, system protection equipment and also the running costs for the operation of this equipment.

- 5 In a uniform pricing approach, the system/market operator would determine a settlement price for energy based on the dispatch that would have occurred in the absence of congestion and the need for re-dispatch. The price is based on the bid of the marginal unit that would have been dispatched under that assumption. Settlements for generators and loads participating in the spot market are based on this uniform price. In addition, generators that are constrained on with bids higher than the uniform price are paid their bid prices. Generators that are constrained down/off are either paid nothing or an "opportunity cost" based on the difference between their bids and the uniform price (old pool of UK). All additional costs incurred by the system/market in the form of payments to constrained on or constrained off generators are charged to all loads as an "uplift." In a locational marginal pricing approach, embedded in the new market design in Russia, the system/market operator determines market-clearing prices at each location based on the marginal cost of serving load at that location, given the dispatch, grid constraints and participant bids. The effects of congestion, including the costs of the redispatch to relieve congestion, will be reflected in the prices determined for each location. Differences in locational prices between one location and another define the opportunity cost of transmission, and hence the appropriate basis for congestion charges. Spot traders in a congested area will thus pay or be paid at the higher locational prices for their area, which includes a congestion charge component, while spot traders in an uncongested area will pay or be paid at their lower locational prices. To ensure comparable treatment between spot traders and bilateral traders, bilateral traders are charged for congestion based on the differences in the locational prices between the point of delivery and the point of receipt.
- 6 As a result from nodal pricing, transmission loss rentals and congestion rentals occur. The transmission loss rentals result from the fact that the transmission losses are quadratic function of the power flow and the marginal transmission losses (the first derivation) exceeds the average transmission losses. Congestion surplus is earned on the constrained transmission capacity because the transmitted power is sold in the receiving region at a higher price than that for which it is purchased in the sending region. The treatment of these rentals (called also surpluses) is discussed later in this document.
- 7 In order to judge the cost-reflectivity of the tariff, it must be noted that the system service tariff should reflect the costs incurred by the SO for procurement of ancillary services. As those costs could have different nature (fixed or variable, e.g., fixed payments for contracting reserve and variable payments for utilization of reserve), a simple kWh-charging base on the total market volume (or total energy consumption) may not be completely cost reflective. Also variations in consumption or production (instead of total energy demand or production) would incur higher cost for ancillary services. Regarding payment liability, the first option is that the generators are not (directly) charged. The same

approach is often proposed for transmission charges, only connection charges are paid by generators. The usual argument used in the debate concerning payment liabilities is that it is not necessary to charge the generators, as the costs will be forwarded to the consumers. However, the disadvantage is that this approach would not give price signals to the generators for efficient use of the system. The generators also take advantage of the system service and the generators do also cause costs for the provision of the system service. In some cases, the generators can cause more costs for the provision of system services than consumers do. This is illustrated by the fact that the total amount of required operational reserve is set equal to the largest generating unit. Consumers and small generators should therefore pay less than large generators. From an economic point of view, there is no good reason to assign costs to other parties than those that cause them as it could create negative pricing effects and distort market signals.

- 8 The other dimensions are referred to as productive efficiency (optimising the cost of building the network) and dynamic efficiency (optimum timing of investment decisions by the user and network operator). These are again more relevant for regulating the allowed revenue of the transmission network service provider.
- 9 In theory, transmission investment should be undertaken, until it is more costly than savings in generation costs, in terms of losses and congestion. In practice, the decision whether or not to invest is taken on the basis of reliability criteria, not according to expectations about future electricity supply costs. Indeed, in most electricity spot markets, the price of electricity is capped, and the cost of unserved energy is never revealed in the market. That is, investment criteria focussing on the future costs of energy cannot operate efficiently in the first place. There is therefore no reason why investment that is undertaken on the basis of strict reliability criteria (and an underlying notion of the value of lost load) should be consistent with signals arising from the energy market. The short-fall between those rentals, which should theoretically accumulate, and actual funds can be interpreted as indicating a discrepancy between reliability targets which are administratively determined and those which would emerge in a more market-driven outcome.
- 10 The underlying principle of SRMC nodal pricing method is that spot prices between two nodes should not differ for more than the SRMC of transporting the electricity. If a power system had a perfect network, with no losses or any kind of constraints, at a given instant in time, the spot prices at all nodes of the system would be the same, i.e., there would be no spatial differentiation of spot prices. In any real network, the effect of losses and network constraints creates differences among the nodes and consumers pay or generators are paid different prices, depending on their locations. The basic theory of real-time or spot-market pricing of electricity was developed by Vickery and Schweppe et al. As set forth by Schweppe et al., the optimal price for electricity is differentiated in space and time and accounts for the variable costs of producing any electricity at the time it is used and any added requirements to compensate for whatever

transmission losses accompany the supply and delivery of the electricity used, and any generating or transmission capacity limitations that might influence the availability of supply as a function of time. See: Vickery, W. (1971): Responsive Pricing and Public Utility Services, Bell Journal of Economics and Management Science, Vol. 2, pp. 337–346; Schweppe, F. / Caramanis, M. / Tabors, R. / Bohn, R. (1988): Spot Pricing of Electricity, Kluwer Academic Publ., Boston, MA.

- 11 In the framework of full nodal pricing, transmission constraints and transmission losses are integrated in the market price calculated for each node in the transmission network.
- 12 To reflect the different cost of lines and cables, cable lengths are adjusted (increased) in relation to the cost difference between lines and cables. All investment costs are valued at replacement cost.
- 13 Generator ratings are uniformly scaled down so that total installed generation equals total demand. Triad demand is defined as the half-hour during the year with the highest demand for electricity and the next two highest half-hours, which are at least 10 days apart from the highest half-hour and from each other.
- 14 There are currently 15 generation zones and 12 demand zones.
- 15 However, these bonuses are not paid out unless there is a guaranteed availability.
- 16 Some additional critical points raised by Ofgem are as follows: (1) the charging basis is hard to justify. Either the security charge should be differentiated by location or the expansion constant should be revised upwards; otherwise, there is a danger that NGC's charges artificially stimulate the demand for more transmission lines; (2) the suppliers pay 75% and generators 25%. An adjustment towards increase of generators' contribution may be appropriate; and (3) the practice of scaling generation by the ratio of peak demand to registered generation capacity has the effect of reducing the zonal differentials for generation charges, compared with those for demand charges, so distorting the locational signals.
- 17 NECA is the Australian National Electricity Code Administrator.
- 18 This point mentioned above under the NECA's review is related to load dynamics and price stability. This argument is strongly correlated with the Australian regulatory price control that is based on the revenue cap regulation. As the allowed revenues are determined for the duration of the regulatory period, load changes could cause recalculation of charges and redistribution of revenue collected from the customers. In this sense, the argument is relevant for the intertemporal charge dynamics and the regulatory price control modus. The regulators often used methods imposing side constraints on the price changes in order to limit the price fluctuations.
- 19 Day time versus nights and weekends.
- 20 Statnett recovers only about 15% of its total network-related cost from marginal loss payments.

- 21 This can be explained by the fact that marginal losses grow faster than load. Marginal losses will therefore be higher than average losses, resulting in a net income from marginal loss payments for the transmission system operator (TSO). For the same reason, marginal losses cannot directly be applied under a system of self-procurement, i.e., where network users have to supply/purchase a certain amount of energy to compensate network losses. In that case, loss coefficients must either be based on average losses or scaled down to avoid overcompensation and minimize system imbalances.
- 22 This may, for instance, be the case under seasonally changing flows patterns.
- 23 A physical bilateral transaction can be scheduled as if the injection submitted a zero offer and the load submitted an infinite bid.
- 24 Benchmarking is sometimes used by the regulators to disclose the efficient cost that should be allowed to the regulated service providers. The output of the benchmarking process is the relative degree of inefficiencies of a particular regulated network service provider in comparison with the best industry practice. Electricity networks use a wide range of inputs (capital, labour) to provide services to customers. While all network service providers use broadly the same inputs, some providers may use proportionately more of some inputs and less of others. The mix of inputs used depends, among other things, upon management practices and the operating environment. Similarly, the nature of services provided by networks varies according to the nature of customer demands. For example, some network service providers may need to maintain significant network capacity to transport power to a small number of customers while others may serve a large number of customers with limited and highly variable demand. The benchmarking technique must be able to accommodate these different supply and demand conditions. Further, the technique must recognise that it is possible to operate at equal efficiency while accepting different inputs and supplying different outputs.
- 25 In the Australian regulatory practice, this approach is called "optimised depreciated replacement cost." In this way, additional emphasis is put on the regulatory option to eliminate redundant network assets. See: IPART, Independent Pricing and Regulatory Tribunal of New South Wales (1999): Draft Statement of Principles for Regulation of Transmission Revenues, May.
- 26 The bottom-up calculations give the revenues that the company should be allowed to recover if its asset values have been fixed. If its prices have been fixed, then the calculations can be reversed (top-down) to derive the implied asset value, as discussed in the previous sections (using recoverable amount approach). This is effectively how the value of a company in a competitive market would normally be estimated by stock analysis deciding whether to buy or sell shares. In this way, the regulatory asset value is set equal to the predicted market value of the company (i.e., the present value of the expected stream of free cash flows).
- 27 Even if a privatisation value for the Russian network service providers would exist, this would not solve the circularity problem. The regulatory framework

- should pre-date the privatisation process and, on the other hand, the privatisation value is needed for the determination of the regulatory asset base. Moreover, experience shows that in many cases, the regulators and energy policy makers seek to establish a pragmatic approach that leads directly to the same quantitative answer instead of searching for economically "perfect" solutions with a low degree of implementation potentials.
- 28 See IPART, Independent Pricing and Regulatory Tribunal of New South Wales: (1999) Rolling Forward the Regulatory Assets Base in Electricity and Gas Industry, Discussion Paper, January.
- 29 Often the companies' accounting uses accelerated depreciation methods. The mostly used such method is the method of declining balance. According to this method, a fixed percentage of the written value of the asset is charged as depreciation each year. The effect of this is that decreasing amounts are charged each year in contrast with the straight-line method that produces an equal charge each year. The fixed percentage to be used is the percentage that should be deducted from the written down value each year so that, over the life of the asset, the total installed cost is reduced to the net scrap value.
- 30 Straight-line depreciation is not immune to critique and FEC and FNC should be aware of this. One of the major issues regarding the determination of regulatory depreciation is the need to achieve a time profile of revenues that is economically efficient. It could be argued that straight line depreciation fails to capture the important features of economic depreciation that are evident from the sale value of assets or the pricing of products over the life-cycle of productive assets.
- 31 Essentially, a transmission pricing methodology based on a multilateral concept must respond to the following issues: identify the system that must be paid, the costs that must be covered and the allocation of payments among the different agents. Such charges for transmission infrastructure have to cover all the capital and operating costs of the transmission network and are consequently not marginal costs for the system as a whole (sunk cost). Although depending on the chargeable basis in the transmission pricing regime, they could be perceived as marginal by the transmission service users.
- 32 See: Ramsey, F. (1927): A Contribution to the Theory of Taxation, in: Economic Journal, Vol. 37.
- 33 Additional results on Ramsey pricing may be found in: Baumol, W. J. / Bradford D. (1970): Optimal Departure from Marginal Cost Pricing, in: American Economic Review, Vol. 60, No.3.
- 34 In Australia, the National Electricity Code determines a 50/50 split between locational (cost reflective network pricing) and postage stamped charges. This was initially determined as a pragmatic approach for estimating LRMC across the shared network. Later, it was discovered that this approximation may hold for South Australia, Victoria and New South Wales but is more likely an underestimate for Queensland where projected network investments were high.
- 35 Simple aggregation of connected customer load for the purposes of cost allocation, however, would fail to reflect the actual use of transmission network and the coincidence of time occurrence of peak demand at each voltage level (assuming that cost should be allocated in proportion to the contribution to the peak demand at each network voltage level). The time of occurrence of peak demand at a particular transmission network voltage level could be different from the time of occurrence of peak demand at an upper transmission network voltage level. Consequently, the contribution of this particular network voltage level to the peak demand at the upper network voltage level will be less than its individual peak demand. Therefore, there is a need for coincident metering
- Cost allocation according to energy consumption might be more easily implemented as the energy measurements exist and the energy consumption on the different network voltage levels could be easily aggregated. It could be argued that demand is the relevant criteria for network planning and the customers with low utilisation of their connected load will pay lower amounts for the use of transmission network in comparison to those that they should have paid in case of demand allocation criteria. It is correct; however, that the economic benefits of proper use of "causality principles" will depend on the availability and reliability of the demand data. It is doubtful whether demand criteria would lead to better allocation when the input information is not available or distorted.
- 36 Robustness, i.e., little volatility with respect to input data or the absence of arbitrary decisions such as the choice of a slack bus that influence the results in other methods, is a very desirable characteristic.
- 37 Such have been applied in South America (and also in New Zealand). The New Zealand approach determines grid "usage" by tracing flows, as does the "area of influence" regime employed in a number of South American countries, where it is also common to charge generators rather than consumers for transmission assets. In Chile, the area of influence is defined as the set of assets upon which flows are deemed to be directly and necessarily affected by the connection of a generator to the grid. Generators pay for the lines to which they contribute a positive flow in proportion to their usage of the capacity of those lines. They are then free to use those lines, but a separate toll, similar to a wheeling fee, is charged for trading outside the area of influence.
- 38 In the case that a pure demand charge is applied, incentives for the customers would appear to be to avoid paying charges for transmission service by limiting his usage to only off-peak hours.
- 39 In the practice, the so-called "threshold approach" is used. The costs are allocated to the time periods where demand exceeds a pre-determined threshold level. The threshold is set as a percentage of the maximum system demand.
- 40 Part of the costs falling under this category could be identified explicitly, e.g., the costs of the communication lines, network planning, general overhead, etc. On the other hand, part of the costs is included in the network infrastructure in form of additional investments necessary to meet contingency standards. In the

- Australian pricing practice, the costs that could be individually identified are separated and form an explicit charge called Transmission Common Service Charge. In the UK pricing practice (as already described), the reverse approach is used; i.e., the transport related costs are explicitly calculated by means of modelling (ICRP approach).
- 41 From an economic perspective, the actual allocation of sunk costs is largely an arbitrary decision, provided that allocation is done in such a way as to avoid any impact on future decision-making. In particular, this implies that any charge levied to recover such costs should be "fixed" in the sense that it should not depend on any present, or future, actions of the parties involved. If these charges vary in any way with current or future "usage," then grid users may change their usage, or build new assets — such as embedded generation — to avoid the charge, even though this makes no difference to the costs, which must be recovered. These users would incur additional costs in order to gain a net advantage for themselves, but leave others to cover the cost of the assets already committed, thus causing an overall loss to the sector as a whole.
- 42 Network asset costs are related to peak capacity, and if costs must be allocated in some way, an allocation on the basis of contribution to peak demand may be appropriate. The exact form of a peak charge can vary: (1) the peak can be defined as the single highest period of demand or supply over a number of years, within each year, or the average demand/supply over a small number of period within the year. An energy charge is then simply the peak charge averaged over all periods; and (2) a peak can be defined as the contribution to system peak demand or regional peak demand, the customer's own peak or the peak for each supply point servicing a customer.
- 43 If peak usage does not change significantly over time, then a peak usage charging regime corresponds closely to the ideal of a fully fixed transmission charge.
- 44 Fixed charges — appropriately applied — are less distortional than variable charges, are simple to apply, could be differentiated by size of customer, and would reduce the taxation impact on consumption.
- 45 The assessment period is defined as the settlement period in which peak load occurred and a given number of settlement periods either side of the settlement period in which the peak load occurred. This number could vary; e.g., in the UK, the system peak demand is supplemented by two additional system peak demands (load triad) with at least ten days time difference to avoid peak demand sequences. In New South Wales, Independent Pricing and Regulatory Tribunal (IPART) proposed in March 1996 that demand charge (25% of the network costs are covered by the demand charge) will apply to 10 winter daily demand days and 10 scaled summer daily demands. The summer demands are scaled up by the ratio of winter/summer demand for the system. The rationale for this price structure is that the need for augmentation of TransGrid network may be attributed roughly equally to summer and winter limitation.
- 46 The distribution ratio shown in the example should be considered as an example and not as a prescriptive coefficient.
- 47 Transmission network assets include only core transmission network assets and exclude transmission connection assets.
- 48 Generally, the transmission connection assets on the generation side (entry transmission connection assets) will include: transmission switchgear plant (circuit breakers and isolators) and associated plant; station establishment and building cost; land. The transmission connection assets on demand side (exit transmission connection assets) will include transmission switchgear plant (circuit breakers and isolators) and associated plant; transformers that supply the subtransmission voltage level and associated switchgear; station and establishment and building cost; land. Taking this into account, the transmission connection charges should cover the connection cost required to provide physical access to the transmission network. This, in turn, consists of the following services: (1) the design, project management and provision of connection assets (including purchase and installation); (2) repair, maintenance of connection assets to ensure that they remain fit for use throughout their replacement period, including the replacement of failed assets during the replacement period; and (3) the removal of the connection assets either at the request of the user or at such time as the connection becomes unsafe.
- 49 This could occur if costs resulting from certain connection had not been recovered and become stranded after the user's decision to leave the system. However, the same is also true for deep connection charging, unless all investment costs for the deep connection are immediately covered by an up-front payment.
- 50 See, e.g.: Kahn, A. E. (1971): *The Economics of Regulation: Principles and Institutions*, Vol. I, II, John Wiley and Sons, Inc., New York; Weizsacker, C. C. v. (1982): *Staatliche Regulierung — positive und normative Theorie*, in: *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, Heft, Vol. 3.
- 51 Incomplete information on the regulator's side is a major problem in all of the above-described drawbacks of "rate or return" regulation. Most apparent is the "cost immunisation" problem. The regulator is not in a position to ultimately judge whether certain inputs would have been required under efficient operation (e.g., whether the staffing is at an appropriate level or whether it is excessive) or whether incurred prices of inputs (e.g., staff wages) were at appropriate levels where no exact comparator exists. The utility is practically allowed to pass all costs actually incurred through to customers.
- 52 See: Joskow, P. / Schmalensee, R. (1986): *Incentive Regulation for Electric Utilities*, in: *Yale Journal of Regulation*, Vol. 4, No.1; Laffont, J. J. / Tirole, J. (1993): *A Theory of Incentives in Regulation and Procurement*, Cambridge, The MIT Press.
- 53 See: Shleifer, A. (1985): *A Theory of Yardstick Competition*, *Rand Journal of Economics*, Vol. 16, pp. 319–327.

- 54 Implementation of price cap regulation, under the name "RPI-X regulation," has been advocated by Littlechild (1983) as an alternative to rate of return regulation. It has been adopted in a large variety of circumstances, most notably in the regulation of telecommunications in Great Britain and the United States. See: Littlechild, S. C. (1983): Regulation of British Telecom's Profitability Report to the Secretary of State, London, UK, Department of Industry.
- 55 The relationship between profit and volume will depend partly on the extent to which the utility can re-balance prices within the basket. For example, a utility may be able to increase its rate of return by increasing the price of a service for which demand has grown and decreasing the price of a service with falling demand.
- 56 The networks are characterised by scale economies and therefore a tariff system based on "pure" marginal costs is not able to raise sufficient revenue for cost recovery. In order to ensure cost recovery, tariffs should be increased above marginal costs. Although theoretically the cost recovery should be ensured through an additional charge that does not depend on the usage (kWh, kW), the practical solutions usually consider a energy and/or capacity charge and incorporates the sunk cost in the tariff design through usage-dependant tariff components.
- 57 See IPART, Independent Pricing and Regulatory Tribunal of New South Wales (1999): Regulation of Network Service Providers. Price Control Issues and Options, Discussion Paper 34, March.
- 58 This gives the maximum revenue that the network business is allowed to earn in year t to cover the provision of its regulated services.
- 59 See IPART, Independent Pricing and Regulatory Tribunal of New South Wales (1999): Regulation of Network Service Providers. Price Control Issues and Options, Discussion Paper 34, March. DTe (1999): Price Caps Regulation in the Electricity Sector, Information and Consultation Document.
- 60 Under a strict formulation of the control, any changes in the underlying cost drivers are not reflected in the allowed level of revenue. There is therefore a danger that the regulated service provider will earn sustained excess profits, or face sustained losses, if cost drivers turn out to be different from forecasts. This risk has led to cost-pass through formulations of some pure revenue controls, in order to take account of unpredictable changes in costs.
- 61 Since there is no necessary link between marginal revenues and marginal costs under revenue yield formula, there is a danger of systematic trends in cost and revenue drivers, such as a shift in demand conditions, leading to persistently higher profits or losses. Persistently higher profits, not anticipated by the regulator at the time the price control was set, may prompt calls for intervention to reopen a price review. Persistent losses borne by the company will put at risk its willingness and ability to continue to provide the network service. The lack of a proportionate link between revenues and costs under this form of control therefore exposes the regulated service provider to a degree of risk, and may increase unnecessarily the cost of capital. See: Office of Regulator General, Victoria (1998): Consultation Paper No.4, December, and Office of Regulator General, Victoria (2001), Electricity Distribution Price Review, Form of Price Control.
- 62 Consider, for example, a situation when the price control ties revenue closely to the volume of energy consumed but underlying costs are not greatly affected by changes in volumes; i.e., the variable cost of distribution is low compared to the fixed cost. If there is an unexpected increase (decrease) in demand, profitability will also increase (decrease) under the operation of the price control, compared with the level initially anticipated. Persistent excess profits are not consistent with the objectives of monopoly regulation in promoting efficiency and protecting consumers. Persistent losses are inconsistent with safeguarding the investment of regulated businesses, will discourage new investment and result in the potential unwillingness to continue providing the regulated services. A form of price control which permits revenues and costs to move out of line with each other therefore exposes the regulated firm to a degree of risk, and is a suboptimal outcome from an economic efficiency perspective.
- 63 In the 1994 UK price review, the electricity regulator, OFFER, altered the form of the control so that, whereas previously revenue increased proportionately with the number of units sold, it would now increase at half the rate. The Director General of Electricity Supply concluded that: "On the evidence available to me, it does not seem that distribution business costs can be said to move entirely with units sold. At the same time, I am conscious of the importance of retaining a general incentive for companies to seek out and meet the needs of their customers. To balance these considerations, I propose to halve from 100 to 50 per cent the weight of units in the 'revenue driver' of the price control and to relate the remaining 50 per cent to customer numbers. This should avoid any artificial disincentive in the distribution price controls to the companies' pursuit of energy efficiency, while at the same time, retaining an appropriate marketing incentive."
- 64 At distribution level, there might be a risk that the distribution network businesses may make relatively larger reductions in the prices paid by large consumers than in the prices paid by smaller consumers. This could raise the companies' profits if smaller consumers are less responsive to prices than larger consumers, so that the company gains more sales to large customers than it sacrifices to small consumers. See also the position of the Dutch Regulator and the Regulatory Office of Victoria (Australia) to this issue: DTe (2000): Guidelines for Price Caps Regulation in the Dutch Electricity Sector in the period from 2000 to 2003, and Office of Regulator General, Victoria (2000): Consultation Paper No.4, 2001, Electricity Distribution Price Review, Form of Price Control.
- 65 Another rather conventional and widely used method to place some discipline on the process is a requirement to publish the tariff setting methodology and expose that to critical scrutiny of network users and researchers.

- ⁶⁶ Transmission service includes transmission network service (paid via transmission use of network charges) and transmission connection service (paid via transmission connection charges). Transmission network service is a service associated with the ability of the FNC to transport energy via the transmission network in a reliable and efficient way (including network planning and network construction) and to ensure secure and efficient operation and regular maintenance of the transmission network assets as required in the Technological Rules and the relevant Russian technical standards.
- ⁶⁷ The FNC's annual capital cost consists of annual depreciation plus a reasonable annual return on its invested capital.
- ⁶⁸ Transmission connection service is associated with the ability of the FNC to provide reliable and efficient connection to the transmission network and regular maintenance of the transmission connection assets as required in the Technological Rules and the relevant Russian technical standards.
- ⁶⁹ Contestable markets for some of the network services may exist which means that the price will be established in the competitive environment of the market. The classical example is the debate on the contestability of connection and exclusion of connection charges from the regulatory control. In UK, this debate led to the deregulation of some of the connection services (e.g., construction of connection, but not design of connection). Another example for contestability is the metering service. Meter installation and meter maintenance could also be potentially competitive. The author is not aware whether similar arrangements exist in Russia currently. Further examples for non-regulated services are the so-called miscellaneous services. Miscellaneous services are services that are different from the transmission service. Examples for such miscellaneous services are: engineering consultancy services; or other related services, such as telecommunications; or generation of electricity in own generation facilities or in capital subsidiaries.

*This publication has been produced by KEMA Consulting
with the assistance of the European Union. The contents
of this publication are the sole responsibility of KEMA Consulting
and can in no way be taken to reflect the views of the European Union.*