The access problem arises in industries where services are produced using inputs from both monopolistic and competitive markets. In the port industry, a number of services need to be jointly provided to complete the logistics chain: pilotage, towage, stevedoring, storage, etc. In ports where a terminal constitutes a natural monopoly, an integrated operator has incentives to deter competition from providers of the competitive services. To avoid such situations from occurring, regulators have two options: forbidding integration or implementing access policies under which all service providers are allowed to use the terminal under reasonable conditions. The first option may create non-trivial transaction costs that result in higher prices for the consumers, for which the second option constitutes a more desirable policy.

Unfortunately, formulating access policies is not an easy task. If conditions are too high, a limited number of firms will enter the market, allowing existing ones to obtain economic rents. If conditions are too relaxed, an excess of entry may occur, thus reducing the operator’s incentives to adequately maintain and expand the terminal. This may produce dear consequences in developing countries, where there is urgency for updating and expanding port infrastructure.

The objective of this thesis is to propose a model to deal with the access problem in the port industry using as inputs the lessons learned in the telecommunications, electricity, natural gas and railways industries.
Access Regulation for Naturally Monopolistic Port Terminals: Lessons from Regulated Network Industries
Access Regulation for Naturally Monopolistic Port Terminals: Lessons from Regulated Network Industries

Toegangsregimes voor natuurlijk monopolistische haventerminals: Lessen uit gereguleerde netwerk industrieën

Thesis
to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus
Prof.dr. H.G. Schmidt
and in accordance with the decision of the Doctorate Board
The public defense shall be held on
Wednesday the 30th of June 2010 at 09:30 hours

by
Enzo Fabrizio Defilippi Angeldonis
born in Lima
Peru
«Nel mezzo del cammin di nostra vita,
mi ritrovai per una selva oscura,
che la diritta via era smarrita.»

- Dante Alighieri, La Commedia

«Leave this world a little better than you found it. »

- BP
Acknowledgements

It has been a long journey since late 2002, when I made the decision to pursue a PhD. The world has changed; Peru has changed; I have changed.

But no journey is made alone and there are several people I would like to thank at the end of mine. First, I am grateful to Professor Haralambides for his guidance, support, and above all, patience. I would not be writing this if he had lacked any of these. I am also thankful to Chantal Cheung, Michele Acciaro and the rest of the mid-2000s MEL team, who withstood endless stories about my home country. In Peru, I would like to thank Carlos E. Paredes and Richard Webb, who provided me with the means and freedom to finish this thesis.

In the personal domain, I would like to thank my family and friends on both sides of the Atlantic. I am especially thankful to my lifetime friend Eugenio Garcia, who practically adopted me when I arrived in The Netherlands. Also, to Alonso Segura, for his friendship and (even financial) support; and to Milagros Monteverde, for just being always there.

Lima, September 2009.
Contents

1. Introduction 1
   1.1. Research objectives 1
   1.2 Overview of the study 3
   1.3 Limitations of the Research 5

Part I: The access problem in the port industry 7

2. Monopolies and the need for regulation 9
   2.1 Monopoly theory 10
   2.2 Regulation theory 16
   2.3 Options to introduce competition in natural monopolies 35

Appendix 1: The Creation of a Pilotage Monopoly in the Netherlands 39
Appendix 2: Theory of Contestable Markets 41

3. Topics in access regulation 47
   3.1 Vertical integration versus vertical separation 47
   3.2 Access pricing 50
   3.3 The Essential Facilities Doctrine 53

Appendix 3: United States v. Terminal Railroad Association 61
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Privatization and regulation</td>
<td>63</td>
</tr>
<tr>
<td>4.1 Private sector involvement in infrastructure industries</td>
<td>63</td>
</tr>
<tr>
<td>4.2 Privatization assessments and the need for better regulation</td>
<td>65</td>
</tr>
<tr>
<td>4.3 Privatization assessments in the port industry</td>
<td>70</td>
</tr>
<tr>
<td>5. Port economics and the need for an access regime</td>
<td>73</td>
</tr>
<tr>
<td>5.1 Port economics</td>
<td>73</td>
</tr>
<tr>
<td>5.2 Port reform and trends</td>
<td>87</td>
</tr>
<tr>
<td>5.3 Access, competition and regulation in the port industry</td>
<td>104</td>
</tr>
<tr>
<td>Part II: Access policies in regulated network industries</td>
<td>117</td>
</tr>
<tr>
<td>6. Network economics</td>
<td>119</td>
</tr>
<tr>
<td>6.1 Architecture of networks</td>
<td>119</td>
</tr>
<tr>
<td>6.2 Economic characteristics of networks</td>
<td>122</td>
</tr>
<tr>
<td>6.3 Consequences of network effects</td>
<td>125</td>
</tr>
<tr>
<td>6.4 Network effects and public policy</td>
<td>126</td>
</tr>
<tr>
<td>7. Access regulation in the telecommunications industry</td>
<td>127</td>
</tr>
<tr>
<td>7.1 Economic characteristics of the telecommunications industry</td>
<td>127</td>
</tr>
<tr>
<td>7.2 Markets and services</td>
<td>130</td>
</tr>
<tr>
<td>7.3 Reform and regulation in the telecommunications industry</td>
<td>132</td>
</tr>
<tr>
<td>7.4 Access arrangements in the telecommunications industry</td>
<td>134</td>
</tr>
<tr>
<td>7.5 Reform, regulation and access regimes in selected countries</td>
<td>138</td>
</tr>
<tr>
<td>7. 6 Lessons for a port access regime</td>
<td>144</td>
</tr>
<tr>
<td>8. Access regulation in the electricity supply industry</td>
<td>147</td>
</tr>
<tr>
<td>8.1 Economic characteristics of the electricity supply industry</td>
<td>147</td>
</tr>
<tr>
<td>8.2 Markets and services</td>
<td>149</td>
</tr>
<tr>
<td>8.3 Reform and regulation in the electricity supply industry</td>
<td>153</td>
</tr>
</tbody>
</table>
8.4 Access arrangements in the electricity supply industry
8.5 Reform, regulation and access regimes in selected countries
8.6 Lessons for a port access regime

9. Access regulation in the natural gas industry
9.1 Economic characteristics of the natural gas industry
9.2 Markets and services
9.3 Reform and regulation of the natural gas industry
9.4 Access arrangements in the natural gas industry
9.5 Reform, regulation and access regimes in selected countries
9.6 Lessons for a port access regime

10. Access regulation in the rail industry
10.1 Economic characteristics of the rail industry
10.2 Markets and services
10.3 Reform and regulation in the rail industry
10.4 Access arrangements in the rail industry
10.5 Reform, regulation and access regimes in selected countries
10.6 Lessons for a port access regime

Part III: An access regime for the port industry

11. Lessons for formulating access policies in the port industry
11.1 Similarities among ports and network industries
11.2 Lessons learned from implementing access policies
11.3 Elements of an access regime for the port industry
Appendix 4: National access regimes in the studied countries

12. A model of access regime for the port industry
12.1 Basic elements
12.2 Operation of the model 241
12.3 Analysis 245

13. Conclusions 249

Bibliography 257

Summary 277

Samenvatting 281

Resumen 283

About the Author 287
List of Figures

Figure 2.1 Costs and Demand Curves in a Monopoly Situation 11
Figure 2.2: Social Loss Caused by a Monopolist 12
Figure 2.3: Sub-Additivity of Costs without Global Economies of Scale 15
Figure 5.1: Port’s Logistics Chain 76
Figure 5.2: Matarani Port Terminal (Peru) 79
Figure 5.3: Port of Algeciras (Spain) 80
Figure 5.4: Port of Zeebrugge (Belgium) 81
Figure 5.5: Port of Rotterdam (The Netherlands) 82
Figure 5.6: Main European ports 83
Figure 5.7: The Hinterland of Matarani port 85
Figure 5.8: Hub Ports in the East-West Route (World Bank, 2007) 86
Figure 5.9: World trade by type of transport service 92
Figure 5.10: Annual Throughput in the Port of Matarani, 1998-2003 103
Figure 5.11: Labor productivity at the port of Matarani, 1998-2003 104
Figure 6.1: A simple star network 120
Figure 6.2: A simple ring network 120
Figure 6.3: A Network with all points directly connected 121
Figure 6.4: Two Star Networks Connected by a Trunk Link 121
Figure 7.1: A typical telecommunications network 128
Figure 8.1: A typical electricity network 148
Figure 8.2: Great Britain’s Transmission Network 157
Figure 8.3: Organized Markets in the United States 161
Figure 8.4: Transmission networks in Australia 162
Figure 9.1: The natural gas supply chain 170
Figure 9.2: Natural gas consumption in the US, 2002 (Ferc, 2004) 174
Figure 9.3: The UK Natural Gas System 178
Figure 9.4: The US Natural Gas System (Ferc, 2004) 181
Figure 9.5: The Australian Natural Gas System (ACCC, 2000) 183
Figure 10.1: Rail activities 190
Figure 10.2: UK’s Rail System (Bueker, 2005) 202
Figure 10.3: US Rail System (USDA, 2005) 206
Figure 10.4: Australian Rail System 209
Figure 12.1: Rationale to Determine Regulatory Option 243
Figure 12.2: Possible interactions between incumbent and an entrant 246
List of Tables

Table 5.1: Port infrastructure and superstructure 76
Table 5.2: Main port markets 77
Table 5.3: Port management models 77
Table 5.4: Price elasticities for HLH range container ports 82
Table 5.5: Port projects with private participation by type (1990-2006) 94
Table 5.6: Private investment in port projects by type (1990-2006) 95
Table 5.7: Global player’s world throughput, 2002 and 2006 97
Table 5.8: Gross Productivity Yields in Colombian Ports 1998 (ton/hour) 100
Table 5.9: Performance Indicators for the Port of Cartagena 101
Table 5.10: Antitrust agency and regulator: typical role distribution 105
Table 5.11: Container terminals: competition options 109

Table 7.1: Telecommunications markets: typical competitive structure 128
Table 7.2: Cost characteristics of diverse transmission technologies 130
Table 7.3: Regulation of telecommunications in OECD countries 134
Table 7.4: Main characteristics of telecommunications access regimes 145

Table 8.1: Cost structure of main generation technologies 150
Table 8.2: Main characteristics of ESI access regimes 165
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 9.1</td>
<td>Main characteristics of natural gas access regimes</td>
<td>186</td>
</tr>
<tr>
<td>Table 10.1</td>
<td>Share of rail in total passenger transport, ECMT countries</td>
<td>193</td>
</tr>
<tr>
<td>Table 10.2</td>
<td>Share of rail in total freight transport, ECMT countries</td>
<td>193</td>
</tr>
<tr>
<td>Table 10.3</td>
<td>Reform processes in selected countries</td>
<td>195</td>
</tr>
<tr>
<td>Table 10.4</td>
<td>Structure and ownership of Australian railways</td>
<td>210</td>
</tr>
<tr>
<td>Table 10.5</td>
<td>Main characteristics railways industry’s access regimes</td>
<td>212</td>
</tr>
<tr>
<td>Table 11.1</td>
<td>Economic characteristics of ports and network industries</td>
<td>218</td>
</tr>
<tr>
<td>Table 11.2</td>
<td>Options for the design of an access regime</td>
<td>220</td>
</tr>
<tr>
<td>Table 11.3</td>
<td>Vertical structure of network industries</td>
<td>221</td>
</tr>
<tr>
<td>Table 11.4</td>
<td>Access pricing in the network industries</td>
<td>223</td>
</tr>
<tr>
<td>Table 11.5</td>
<td>Non-price access terms in the network industries</td>
<td>225</td>
</tr>
<tr>
<td>Table 11.6</td>
<td>Mechanism to expand the infrastructure in network industries</td>
<td>227</td>
</tr>
<tr>
<td>Table 11.7</td>
<td>Characteristics of an Access Regime for the Port Industry</td>
<td>230</td>
</tr>
<tr>
<td>Table 12.1</td>
<td>Types of Markets in Ports with Monopolistic Terminals</td>
<td>243</td>
</tr>
<tr>
<td>Table 12.2</td>
<td>Appropriate Regulatory Option for Matarani</td>
<td>245</td>
</tr>
<tr>
<td>Table 12.3</td>
<td>Interactions between Incumbent and Potential Entrant</td>
<td>246</td>
</tr>
</tbody>
</table>
### Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCC</td>
<td>Australian Consumer and Competition Commission</td>
</tr>
<tr>
<td>AER</td>
<td>Australian Energy Regulator</td>
</tr>
<tr>
<td>ARTC</td>
<td>Australian Rail Track Corporation</td>
</tr>
<tr>
<td>BG</td>
<td>British Gas</td>
</tr>
<tr>
<td>BR</td>
<td>British Rail</td>
</tr>
<tr>
<td>BT</td>
<td>British Telecom</td>
</tr>
<tr>
<td>CLEC</td>
<td>Competitive Local Exchange Company</td>
</tr>
<tr>
<td>ECMT</td>
<td>European Conference of Ministers of Transport</td>
</tr>
<tr>
<td>ECPR</td>
<td>Efficient Component Pricing Rule</td>
</tr>
<tr>
<td>EFED</td>
<td>Essential Facilities Doctrine</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FCC</td>
<td>Federal Communications Commission</td>
</tr>
<tr>
<td>FDC</td>
<td>Fully Distributed Costs</td>
</tr>
<tr>
<td>FERC</td>
<td>Federal Energy Regulatory Commission</td>
</tr>
<tr>
<td>ICCC</td>
<td>Independent Consumer and Competition Commission (Papua New Guinea)</td>
</tr>
<tr>
<td>ISO</td>
<td>Independent System Operator</td>
</tr>
<tr>
<td>IXC</td>
<td>Long Distance Carriers</td>
</tr>
<tr>
<td>kV</td>
<td>Kilovolts</td>
</tr>
<tr>
<td>kW</td>
<td>Kilowatt</td>
</tr>
<tr>
<td>LATA</td>
<td>Local Access Transport Areas</td>
</tr>
<tr>
<td>LDC</td>
<td>Local Distribution Company</td>
</tr>
<tr>
<td>LEC</td>
<td>Local Exchange Company</td>
</tr>
<tr>
<td>LLLU</td>
<td>Local Loop Unbundling</td>
</tr>
<tr>
<td>LNG</td>
<td>Liquid Natural Gas</td>
</tr>
<tr>
<td>LRIC</td>
<td>Long Run Incremental Costs</td>
</tr>
<tr>
<td>LRMC</td>
<td>Long Run Marginal Costs</td>
</tr>
<tr>
<td>MMC</td>
<td>Monopolies and Mergers Commission</td>
</tr>
<tr>
<td>MW</td>
<td>Megawatt</td>
</tr>
<tr>
<td>NCC</td>
<td>National Competition Council</td>
</tr>
<tr>
<td>NECA</td>
<td>National Electricity Code Administrator</td>
</tr>
<tr>
<td>NEM</td>
<td>National Electricity Market</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Nemmco</td>
<td>National Electricity Market Management Company</td>
</tr>
<tr>
<td>NETA</td>
<td>New Electricity Trading Arrangements</td>
</tr>
<tr>
<td>NGC</td>
<td>National Grid Company</td>
</tr>
<tr>
<td>NLC</td>
<td>Nederlandse Loodsen Corporatie</td>
</tr>
<tr>
<td>OECD</td>
<td>Organization for Economic Cooperation and Development</td>
</tr>
<tr>
<td>Ofcom</td>
<td>Office of Communications</td>
</tr>
<tr>
<td>Ofgas</td>
<td>Office of Gas Supply</td>
</tr>
<tr>
<td>Ofgem</td>
<td>Office of Gas and Electricity Markets</td>
</tr>
<tr>
<td>OFT</td>
<td>Office of Fair Trading</td>
</tr>
<tr>
<td>Oftel</td>
<td>Office of Telecommunications</td>
</tr>
<tr>
<td>PPIAF</td>
<td>Public-Private Infrastructure Advisory Facility</td>
</tr>
<tr>
<td>PPP</td>
<td>Public-Private Partnership</td>
</tr>
<tr>
<td>PSTN</td>
<td>Public Switched Telephone Network</td>
</tr>
<tr>
<td>PTO</td>
<td>Public Telecommunications Operator</td>
</tr>
<tr>
<td>PUC</td>
<td>Public Utility Commission</td>
</tr>
<tr>
<td>RPI</td>
<td>Retail Price Index</td>
</tr>
<tr>
<td>RTO</td>
<td>Regional Transmission Organization</td>
</tr>
<tr>
<td>STB</td>
<td>Surface Transportation Board</td>
</tr>
<tr>
<td>TPA</td>
<td>Third Party Access</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States of America</td>
</tr>
<tr>
<td>VTS</td>
<td>Vessel Traffic Services</td>
</tr>
<tr>
<td>wh</td>
<td>watts per hour</td>
</tr>
</tbody>
</table>
1. Introduction

1.1. Research objectives

During the last 50 years, trade in finished goods, components and services has been growing faster than the world’s GDP; a trend that is expected to continue in the future (Kumar and Hoffman, 2002). To be able to compete in a global scenario, an increased number of developing countries have been removing barriers to trade and resorting to the private sector to build and improve their port infrastructure. Indeed, according to PPIAF (2007), between 1990 and 2006 the private sector participated in almost 300 port projects in developing countries, worth in excess of US$ 33 billion. More than ninety percent of these projects consisted of concessions, greenfield projects and divestitures.

However, due to the economic characteristics of port operations (economies of scale, scope and density, and lumpy investments) some port terminals constitute natural monopolies for carriers and shippers located in their hinterlands (Flor and Defilippi, 2003). Transferring monopolistic port terminals to the private sector without an adequate regulation of prices and access would allow their operators to obtain rents at expense of users (Train, 1991).

Unlike railways and transmission or distribution networks in telecommunications, electricity and natural gas industries, which almost always exhibit natural monopoly characteristics, in the port industry natural monopolies arise mostly in developing countries and small islands where traffic is too limited to justify a second terminal. Not surprisingly, literature regarding regulation of natural monopolies in network industries is abundant, while in ports, is scarce. Apart from few publications from multilateral institutions such as the United Nations and the World Bank (referring to transport regulation in general) and a handful of research papers, port regulation remains an unexplored matter. Moreover, its practice has been limited to a handful of countries. The only countries known to have established port regulators are Australia, Colombia and Peru. The further development of the discipline would allow governments to involve the private sector in the operation and construction of monopolistic terminals characteristics without the fear of harming carriers and shippers.
Regulation implies implementing two sets of policies: (i) *price regulation* (setting the price of services provided by the monopolist); and (ii), *access regulation* (setting the rules to ensure the monopolist will not deter access to related markets) (Defilippi and Flor, 2008). The objective of this thesis is to analyze the characteristics of access policies implemented in the telecommunications, electricity supply, natural gas and railways industries, and to use the lessons learned from these experiences to propose a model suitable for the port industry. Its relevance resides in the importance of the subject for the formulation of development strategies in developing countries, and the dearth of previous studies in the field.

**The access problem in the port industry**

The problem of access arises in industries where inputs from monopolistic and competitive markets are complementarily needed to provide a service. In these circumstances, the firm controlling the monopolistic segment has incentives to deter competition in the competitive segments (markets) to recover profits foregone by regulation (Paredes, 1997). In the port industry, for example, a number of services need to be provided to complete the logistics chain: pilotage, towage, stevedoring, storage, etc. Without any of these, cargo cannot be delivered. In ports where a terminal constitute a natural monopoly, an integrated terminal operator\(^1\) has incentives to deter competition in the markets of services that are necessary to complete the logistics chain, since this would allow him to charge disproportionate prices and recover monopolistic rents. This strategy can be implemented by preferential treatment to itself or sister companies, or by restricting competitors access to the terminal.

OECD (2006) give several examples of terminal operators having sought to favor vertically-related carriers by impeding or complicating the operation of competing firms: in the UK, a terminal operator allowed a related ferry line to schedule its service to disrupt an entrant’s loading and unloading of passengers. In Denmark, the terminal denied access to a ferry line on the grounds that it would prevent existing companies from expanding operations. In Bulgaria, a catering firm was unable to enter the market because the terminal management refused to dispose the waste brought by the company or to allow its trucks entering the terminal. Similar behavior by operators of naturally monopolistic terminals has been reported in Peru (Alcazar and Lovaton, 2005) and Colombia (World Bank, 2007). Likewise, the expansion into developing countries of carrier-related terminal operation companies (such as Cosco Group, MSC, Evergreen) to operate common-user facilities has also raised concern that similar attitudes may arise. One example is the concern expressed after the Cameroon Government granted APM Terminals (a company related to Maersk) the concession to operate Douala’s common-user container terminal (Port Strategy, 2003 and 2004).

---

\(^1\) A terminal operator that also provides shipping or logistics services (by itself or through related companies).
To avoid the situations from occurring, regulators have two options. They can either (i) forbid integration between terminal operators and carriers or, (ii) establish a framework under which all service providers are allowed to access and use the terminal under reasonable conditions. As suggested by Vickers (1995), the first option (vertical separation) may create non-trivial transaction costs that result in higher prices for the consumers, for which the second option (the implementation of access policies) constitutes a more desirable policy. However, formulating access policies is not an easy task. If access conditions are too high, a limited number of entrants will use the terminal, allowing providers to obtain economic rents. If conditions are too relaxed, an excess of entry may occur, thus reducing the terminal operator’s incentives to adequately maintain and expand the infrastructure (Laffont and Tirole 1994).

This research analyzes access policies implemented in three countries: UK, US and Australia. These countries were chosen because they have approached the access problem from different perspectives, and because they are considered best practice cases among regulation practitioners (ADB, 2000).

The UK was the first developed country that carried out a comprehensive privatization program, which required formulating access policies without the benefit of previous experiences. As we will see in the following chapters, this lack of experience, but also of sound economic assessments, led UK regulators to underestimate the monopolists’ ability to deter competition despite the implementation of open access policies. On the other hand, the US has a long tradition of encouraging competition and allowing the private supply of public services. In this country, the reform of network industries focused more in restructuring than in changing the nature of their ownership. The Australian strategy, on the other hand, consisted on implementing the National Access Regime whose provisions cover all relevant infrastructures (regardless of the nature of their ownership and the industry they belong).

1.2 Overview of the study

The thesis comprises three parts.

**Part I** presents the theoretical developments that are necessary to understand the nature of the access problem, and the choices to address it. It also analyzes the particular nature of the access problem in the port industry and the reasons why governments need to address it before concessioning naturally monopolistic terminals.

---

2 An optimal entry is the number of competitors that would enter the market if the terminal was not monopolistic (Laffont and Tirole 1994).
Chapter 2 presents the basic elements of regulation theory: monopoly pricing, regulation rationale and regulatory failures. It also introduces the five basic dilemmas a regulator faces when addressing the problems caused by monopolies. It also discusses the main options to introduce competition in network industries: competition for the market, competition over existing networks and competition among networks. These options are suitable for industries with different economic characteristics.

Chapter 3 discusses the most relevant topics related to the access problem: whether vertical separation is convenient or not, the theory behind access pricing and the Essential Facilities Doctrine (EFD). Chapter 4 analyzes the evidence on the effects of privatization processes around the world. It starts by discussing the reasons argued by governments to embark on privatization programs and the extension of these in the world. It also reviews diverse assessments on efficiency, tariffs and welfare, concluding that the majority of them support the idea that privatization had brought net positive effects to the implementing countries. Finally, the chapter reviews privatization assessments in the port industry. This section concludes that successful privatization processes are those complemented by sound regulation, thus supporting the idea that better regulatory policies are needed to deal with natural monopolies in the port industry.

Chapter 5 provides an overview of the basics of port economics; discusses the main drivers for port reform around the world and illustrates the processes undertaken in three developing countries. Its aim is to demonstrate that privatization of naturally monopolistic terminals could be counterproductive in countries unless access policies are implemented. The chapter also discusses the relationship between competition and access and the dilemmas regulators face when implementing such policies. It shows that designing access policies involves making decisions or setting rules regarding four main issues:

a. Vertical structure;

b. Pricing;

c. Non-price terms and conditions; and,

d. The mechanism to expand the infrastructure.

Part II analyzes the access policies implemented in regulated network industries in UK, US and Australia. Its goal is to summarize the lessons learned during their implementation and to use them to propose an access model for naturally monopolistic terminals. It starts with chapter 6, which describes the economic characteristics of network industries and their relationship with diverse market structures. This chapter is necessary to understand how the complementarity between the components of a network may allow a monopolist controlling an essential input to deter competition in related markets.
Chapters 7 to 10 analyze the access policies implemented in telecommunications, electricity, natural gas and rail industries. To facilitate comparisons, they follow the same structure. With the aim of contextualizing the analysis, the economic characteristics of the industry and its reform trends are first presented. This is followed by a discussion on the industry’s access arrangements and a study of the UK, US and Australian cases. Each chapter concludes with a section summarizing the lessons that can be drawn for formulating access policies in the port industry.

The aim of Part III is to propose an access model for the port industry using as inputs the lessons drawn from analyzing access policies implemented in network industries.

Chapter 11 aims at drawing these lessons. It starts by discussing the similarities between the economic characteristics of ports and networks industries. The next section summarizes the lessons learned from implementing access policies in these industries. The last section uses the lessons learned to determine the characteristics that an access regime for the port industry should have.

Chapter 12 proposes an access model to regulate naturally monopolistic port terminals using the lessons learned from network industries. The main objective of the proposed model is to allow competition to occur in markets that otherwise would need to be price-regulated. By limiting regulatory intervention only to situations when it is effectively needed, this policy is less costly and less likely to spawn market distortions.

The analysis of the possible interactions between incumbent and service providers shows that the implementation of the proposed access regime is likely to reduce entry barriers, increase contestability and introduce competition in markets for port services. Five out of eight possible outcomes are clearly competitive, and the above-the-market returns obtained in the remaining three are expected to attract competitors until these rents disappear.

Finally, chapter 13 presents the conclusions of the research.

1.3 Limitations of the Research

This thesis has several limitations that need to be pointed out.

The most important limitation is the one indicated in section 1.1: there is a limited number of academic articles in the subject of port regulation\(^3\), an even more reduced number of port terminals subject to economic regulation and only few port regulators

\(^3\) A search for the words “port” and “regulation” in IDEAS (http://ideas.repec.org/; Internet’s largest open database of economic papers) yielded only 30 results as of December 2008.
in the world. As indicated earlier, only the cases of private terminals located in
Australia, Colombia and Peru are known to be subject to port regulation, and only the
latter has been studied in academic journals.

A second limitation is that even though examples of regulatory policies implemented
in countries other than the UK, US and Australia are mentioned when possible, it was
not possible to analyze them in detail. The author considers, though, that the analyzed
country cases provide a perspective broad enough to illustrate the options regulators
have when implementing access policies.

A further limitation is caused by the inter-industry approach of the research, which
prevents a more detailed analysis of the motivations policy makers had when favoring
one regulatory option over another. Although this is not required to analyze the
consequences of the policies actually implemented, its inclusion would have
substantially enriched the research. Likewise, the descriptions of the implemented
regulatory, and the events that occurred in parallel or as a consequence of them, are
based on secondary sources. Although an adequate effort has been put into describing
facts rather than interpretations of them, the use of secondary sources may have cause
inaccuracies the author was not able to correct.

A final limitation of this thesis is that it is impossible to contrast the proposed access
regime with reality unless it is implemented and enough time passes to analyze its
consequences. Further research will be necessary when this occurs.
Part I: The access problem in the port industry
2. Monopolies and the need for regulation

Regulation is one of the forms of government intervention aimed at assuring the correct functioning of economic activities. Through laws, regulators condition agents to behave in a different manner than what they would do in the absence of institutional restrictions. There exist three main types of regulation: economic, social and process regulation. *Economic regulation* refers to restrictions in prices, quantity and access conditions for specific industries. *Social regulation* refers to rules that apply to several industries, like health, safety and environmental ones. *Process regulation* refers to government management of the operation of the public sector, like administrative requirements and costs incurred by producers and consumers (Guasch and Spiller, 1999).

Although certain regulations fall into more than one of these categories (like economic regulations caused by environmental concerns) and some others do not fit in any of the described ones, this classification serves to clarify what kind of regulation this research deals with: the economic regulation of transport infrastructure.

This chapter introduces the theories of monopolies and regulation. The first section presents the basic economics of both the general and natural monopoly cases, describing their economic characteristics and consequences on welfare. The second section introduces economic regulation and explains the theoretical arguments behind its existence. It also presents the main theoretical dilemmas a regulator has to face when implementing regulatory policies. The last section presents the menu of regulatory options to introduce competition in natural monopolies.
2.1 Monopoly theory

2.1.1 The general monopoly case

In the absence of regulation, the monopolist would choose a price that maximizes its profit function $\Pi(p)$, which is the difference between the cost of producing goods and the revenues obtained by selling them:

$$\max_p \Pi = p \cdot D(p) - C(D(p))$$

Where $D(p)$ and $C(p)$ are the demand and cost functions, respectively. Applying the first-order condition $\frac{d\Pi}{dp} = 0$, the following expression is obtained:

$$D(p^*) + p^* \cdot D'(p^*) = MC \cdot D'(p^*)$$

Where $MC$ is the marginal cost. Reordering and dividing into $p^*$:

$$\frac{p^* - MC}{p^*} = \frac{D(p^*)}{p^* \cdot D'(p^*)}$$

This expression is equivalent to: $\frac{p^* - MC}{p^*} = \frac{1}{\varepsilon}$, where $\varepsilon$ is the price elasticity of demand at equilibrium, defined as: $\varepsilon = \frac{D'(p^*)}{p^*} \cdot \frac{1}{D(p^*)}$

Consumer surplus is a measure (in currency units) of the utility generated by consuming a determined quantity at a determined price $p^*$. It is defined as the consumers’ willingness-to-pay, less what they actually pay:

$$E(p^*) = \int_0^{D(p^*)} D(p) \cdot dp - p^* \cdot D(p^*)$$

Therefore, the variation of the consumer surplus when prices vary is:

$$dE(p^*) = -D(p^*) \cdot dp$$

Figure 2.1 shows the cost curves of and demand that are typical of a monopoly situation. Notice that the demand curve is negatively-sloped, since the monopolist supply the entire market, for which the demand it faces matches that of the entire market.
It can be seen that the price that assures the achievement of allocative efficiency is where demand (D) and marginal costs (MC) are equaled. In this point, the value of the last produced unit matches the marginal cost of producing it. In figure 2.1, this corresponds to point c, where a quantity $y_c$ is produced and a price $p_c$ is reached. At this point, the sum of consumer surplus (abc) plus the net benefit of the monopolist (bcd) is the maximum. Since this is the point where the welfare of the society is maximized, it is called the first-best.

However, when the market structure is a monopoly, the interests of the monopolist do not match the interests of the rest of society. In this case, the point that maximizes benefits to the monopolist is not point c but point g, where the revenue produced by the last unit sold (its marginal revenue) equals the cost of producing it (its marginal cost). Therefore, a profit-maximizing monopolist will use its market power to limit the number of units in the market until marginal revenue (MR) matches marginal cost (MC). In the terms of figure 2.1, the quantity that achieves this equivalency is $y_m$, causing the price to rise up to $p_m$. This behavior, however, will cause negative consequences to consumers, who are obliged to pay a price of $p_m$ instead of $p_c$ due to the artificial scarcity caused by the monopolist by restricting production.

Additionally, the monopolist’s behavior also causes economic inefficiency. It can be seen in figure 2.2 that raising the price to $p_m$ reduces consumer surplus by the area emcb, but the monopolist only appropriates the area emfb. Therefore, the triangle mcg is lost, causing a social cost in terms of inefficiency called triangles of Dupuit, Harberger or deadweight loss (Pindyck and Rubinfeld, 1998).
There are other inefficiencies caused by monopolies. One of them is the so called “X-inefficiency” that arises when monopolies do not minimize their costs. These productive inefficiencies have been interpreted as a consequence of the lack of effort to reduce costs of firms facing limited competition. Nevertheless, these arguments against monopolies are less important when other aspects such as technological innovation or economies of scale and scope are considered. Tirole (1990) argues that monopolistic firms have more incentives than potential entrants to innovate when these innovations are non drastic, while entrants have more incentives to introduce innovations when they are more drastic. In this view, competitive structures do not necessarily deliver better results regarding technological innovations than monopolistic ones. On the contrary, Guasch and Spiller (1999) argue that the regulation required to prevent abuses from monopolies may cause adverse effects in the economy by delaying the introduction of technologically available services. They use the case of cellular telephony and voice messaging in the US to illustrate how regulation can slow the introduction of new products and discourage innovation, costing the consumers hundreds of millions of dollars.

It is worth noting that from a theoretical point of view, it may be beneficial for society to maintain a monopoly if economies of scale or scope exist, since, in this case, the average production cost of one firm operating in the market would be lower than the average cost of two or more firms (regardless the quantity produced). In these circumstances, the introduction of competition may not be desirable. This is the natural monopoly argument used to restrict the access of competitors in several industries where economies or scale or scope are significant.
2.1.2 The natural monopoly case

The concept of natural monopoly has been usually linked to the existence of economies of scale and scope. However, relatively recent research has shown that the appropriate definition of natural monopoly lies on the concept of \textit{sub-additivity of costs} (Tirole, 1990). These concepts are formally defined below.

**Economies of scale**

To assure the presence of economies of scale in the case of a multi-product firm, it is necessary to bear in mind the following definitions:

a. \textit{Average Multi-Product Cost (AMPC)}. In the case of a firm producing “n” goods, $AMPC = \frac{C(Y)}{a \cdot Y}$, where $Y$ is the production vector and $a > 0$ is a vector of relative weights.

b. \textit{Decreasing Average Multi-Product Costs}. AMPC is decreasing in “Y” if AMPC(tY) is a decreasing function of the scalar “t” when t=1:

$$\frac{\delta AMPC (tY)}{\delta t} \bigg|_{t=1} < 0$$

There exists evidence of the presence of economies of scale in the case of a Multi-Product firm if the AMPC is decreasing. It can be said that a firm’s cost function is sub-additive, when the presence of economies of scale is supposed, if the following conditions are met (Gallardo, 1999):

a. The cost function is convex through cutting rays departing from the origin;

b. The cost function is quasi-convex; and,

c. The cost function has decreasing AMPC

**Economies of scope**

There exist economies of scope in the production of a combination of goods, when the cost of producing such a combination by a single firm is lower than the cost of producing it by two or more firms that do not produce the same good (Tirole, 1990). In the case of a multi-product firm, economies of scope refer to the existence of synergies in the production of two or more goods derived, for example, from the shared use of a production factor. This would allow a single firm to produce goods.
cheaper than competitors who do not share production factors, even if economies of scale are not present.

To assure the presence of economies of scope in the case of a multi-product firm, it is necessary to bear in mind the following definitions:

a. **Incremental Cost of Producing “j”**. The incremental cost of producing “j” is the difference between producing all the goods and producing all the goods minus “j”:

\[ IC_j(Y) = C(Y) - C(Y_{-j}) \]

Where \( Y_{-j} \) is the same production vector “Y”, but placing “0” instead of “j”.

b. **Average Incremental Cost**. The average incremental cost of “j” is:

\[ AIC_j = \frac{IC_j(Y)}{Y_j} \]

It can be said that the firm’s cost function is sub-additive, when the presence of economies of scope is assumed, if for any good \( Y_j \) belonging to the production vector \( Y \), the average incremental cost is decreasing (Gallardo, 1999).

### Sub-additivity of costs

An industry is said to be a natural monopoly when technology imposes a cost function that makes it cheaper, for a relevant demand interval, to produce a good or service with only one firm in the market (Baumol, Panzar and Willig, 1982). Formally, a cost function “C” associated to a production vector “Y” is strictly sub-additive if the cost of producing \( Y \) by one firm is lower than the cost of producing the same vector by two or more firms using the same technology, for any sub-group “i” of \( Y \). Therefore, a cost function is sub-additive if it satisfies the following condition for the relevant demand interval:

\[ C(Y) < \sum_i C(Y_i) \]

Where:

\[ \sum_i Y_i = Y. \]

Then, **an industry can be defined as a natural monopoly if the cost function is strictly sub-additive for the relevant demand interval** (Tirole, 1990). As it can be seen, the concepts of sub-additivity of costs and natural monopoly are mutually implied.
It is worth noting that only in the case of single-product producer, economies of scale constitute a sufficient (although not necessary) condition for the existence of a natural monopoly. If the firm produces several goods, economies of scale are neither a necessary nor a sufficient condition for the existence of natural monopolies. Graphically, this can be seen in figure 2.3. Assume all firms that might provide the good or service in question have identical cost structures. In the figure, each firm’s average cost curve $AC(y)$ declines up to the production level $y_a$, increasing thereafter. The market demand intersects the average cost curve at the output level $y_b > y_a$. Given the shapes of the curves in figure 2.2, it can be seen that a single supplier could serve the entire market at lower unit cost than any industry configuration with two or more firms. In this sense, the industry is a natural monopoly even if economies of scale do not exist for all levels of output up to $y_b$ (Breautigam, 1989).

The natural monopoly constitutes one of the cases where market forces cannot obtain efficient allocation of resources, and this has been the principal argument to justify the regulation of infrastructure-based industries. In this case, regulation is used to avoid excessive prices, thus allowing a socially efficient use of the produced goods.

Natural monopolies are more likely to arise where the total cost has a large fixed-cost component, like in most public utilities such as in electricity distribution or in a fixed-line telephone network. As we will see later, large transport infrastructures such as ports, airports, railways and highways may also constitute natural monopolies depending on the size of the demand they face.
2.2 Regulation theory

According to Brown et al (2006), regulation refers to government-imposed controls on business activity, concerning the setting, monitoring, and enforcing of maximum tariffs and of minimum service standards.

2.2.1 The need for regulation

According to Viscusi, Vernon and Harrington (2000) there are both positive and normative views that offer answers to the question of why there is regulation. From the positive point of view, regulation is justified when the fundamental theorems of welfare economics fail. These theorems constitute the theoretical support of free markets as an efficient resource-allocation mechanism (Varian, 1999).

a. First Theorem of Welfare Economics. Under the following conditions, equilibrium in a set of competitive markets is Pareto-efficient: (i) there are no consumption externalities; and (ii), there are enough agents to ensure that each one behaves competitively.

b. Second Theorem of Welfare Economics. If all agents have convex preferences, every efficient allocation is a competitive equilibrium for some initial allocation of goods.

These theorems state that, under certain conditions, market forces lead to an efficient allocation of resources without the need of external intervention. The first theorem implies that a private market in which each agent seeks to maximize his or her own utility will result in a Pareto-efficient allocation. The second theorem implies that whatever allocation of resources is considered socially fair can be supported by the market mechanism From a normative point of view, regulation is justified when a market is a natural monopoly or it is plagued by externalities. In these circumstances, free markets do not lead to an efficient allocation of resources (Pindyck and Rubinfeld, 1998).

The existence of a natural monopoly obliges a society to face a conflict between productive and allocative efficiencies whose solutions require government intervention in the form of regulation. To achieve productive efficiency, it would be necessary to allow only one firm in the market, because it is the only case when the value of the inputs used to supply the market is minimized. However, this lack of competition would encourage the monopolist to set prices above marginal cost, therefore impeding the achievement of allocative efficiency that is produced when prices are set as close as possible to production costs.

Externalities exist when the actions of an agent cause positive or negative consequences on other agents’ welfare and they are not internalized by the causing
party (Fernandez-Baca, 2006). The most evident example is the use of natural resources, like water, in the production of goods. Without government intervention water would be inadequately priced, thus leading to an inefficient waste that would deplete its sources or prevent other producers from using it. Therefore, in the presence of externalities, competition does not result in optimal allocation of resources, for which regulation is required.

Regulation is usually applied to utilities such as electricity, natural gas and telecommunications, all of which exhibit characteristics of natural monopolies. As it will be shown in the following sections, under certain circumstances key transport infrastructures such as ports and airports may also constitute natural monopolies requiring regulation. Other reason argued for regulation, besides the standard distortionary and distributive effects of monopoly pricing, is the way these industries affect the competitiveness of the rest of the economy, since the production of most goods uses transport services, energy or telecommunications as inputs. If they are priced high, the country is unnecessarily vulnerable to off-shore competition (Tarzijan and Paredes, 2001).

According to Guasch and Spiller (1999), in the specific case of developing countries, where the institutional framework is weak and regulatory decisions may be biased due to interest groups or political considerations, the demand for regulation comes from the need to solve three types of problems:

a. *Government behavior that may distort incentives to invest.* These are the problems that may distort the investment incentives of companies owning large and expensive infrastructures, such as port terminals. To invest in these long-lived assets, operators require prices to be set at levels that allow them to recover their costs. However, given the weak institutional framework of developing countries, political considerations may tempt governments to set prices below its optimal level. As shown by Defilippi and Flor (2008), regulatory agencies enjoy a large degree of discretion in the selection of estimation methodologies when estimating regulated prices, which they could use to set artificially-low prices.

The solution to this problem requires credible regulation. It means the design of an institutional arrangement that deters the government from acting opportunistically. For example, by granting operational independency to regulatory agencies or reducing the causes for which their commissioners or directors can be sacked.

b. *Problems between governments and interest groups.* These problems refer to the problems that may encourage governments to distort regulatory policies to favor of special-interest groups.

According to the positive theory of regulation, this occurs when an interest group is able to convince the government to use its coercion power to improve
its welfare, even though harming other groups. This situation is more likely to occur when the harm produced, although large in aggregate, represent just a small loss of welfare for each individual consumer. Examples of this behavior can be seen in the transport industry. Appendix 1 illustrates the case of pilotage in the Netherlands, although similar arrangements are found in many other countries.

The government may benefit from this behavior if, as a consequence, receives political support from a well organized group. Well organized groups are more likely to gain from this behavior than less organized ones. Producers are typically well organized, while consumers are not (Viscusi, Vernon and Harrington, 2000).

c. Problems between firms and their customers, as a consequence of the firms’ market power. These are the problems that arise as a consequence of the asymmetric bargaining power between consumers and a single producer. If left unregulated, this asymmetry may allow the producer to raise its prices above marginal costs or provide poor quality levels. Regulation could solve this contractual problem by setting both price and quality levels.

In response to these problems, an optimal regulatory framework should set prices adequately, provide adequate investment incentives, create a framework for productive efficiency and minimize opportunities of interest groups to lobby for inefficient policies. However, in the real world the environment for the production of regulation differs from optimal, and it is strongly influenced by political considerations, interest groups and producers’ bargaining power (Laffont, 1995). In order to solve these problems, regulation may take many forms, depending on the characteristics of the specific sector and the country’s economic, legal and political context. Countries may choose to have industry-specific regulators (UK), one regulator covering several industries (US at state level) or grant regulatory powers to its antitrust agency (Australia).

2.2.2 Regulatory failures

As we have seen, regulation is needed to resolve problems that cannot be solved by market mechanisms alone. However, regulation also fails, and according to the theory of public choice⁴, the failures generated by government intervention may result in higher economic costs than those caused by market failures (Viscusi, Vernon and Harrington, 2000). The most common regulatory failures are the following:

---

⁴ This theory is based on the application of the criteria of individual choice (maximization of individual preferences subject to determined restrictions) to decisions made by public institutions (Lasheras, 1999)
a. Information asymmetries. This problem arises because regulatory agencies have less information about technologies, costs and the characteristics of the demand than the regulated firms.

In the absence of complete information about the firm’s technology and opportunity costs, governments face a trade-off between granting adequate incentives to the firm and reducing their potential rents. As a matter of fact, schemes that provide good incentives to a regulated firm to reduce costs are those that reward it with €1 for each €1 of cost reduction, such as fix-price contracts or RPI-X price-setting methodology (which will be analyzed in detail in section 2.2.4). However, if the government cannot observe if a cost reduction is a consequence of exogenous variables (such as technology changes), or endogenous ones (such as cost-reduction efforts enabled by the firm), this scheme generates substantial rents. Alternatively, if prices were set according to incurred costs, regulated firms would not have incentives to reduce them. According to Laffont and Tirole (1993), these errors may introduce more inefficiencies to the economy that than those the government tries to avoid.

Policies to overcome these informational disadvantages require sophisticated administrative capabilities and a substantial amount of regulatory discretion. Unfortunately, as mentioned before governments in developing countries tend to face institutional and financial restrictions that limit their administrative capabilities. Granting significant discretion to public officers in contexts of weak institutional development increases regulatory risks and may deter private investments.

b. Regulatory capture. These failures arise when regulatory decisions are biased towards the interests of a determined group, causing inefficiencies by improving the welfare of this group at the expense of others. This problem, which constitutes one of the reasons that explain the demand for regulation, was explained in the previous section.

To understand why governments can get away with policies that favor certain groups, Laffont and Tirole (2000) explain the relation among general public, government and regulated firm, in terms of two agency problems. The regulatory problem consists of finding the right incentives the government (principal) should give to the regulated firm (the agency) so that the latter will behave consistently with the policies set by the government. Likewise, the policies set by the government, in theory, constitute a response to the interests of general public, who in this second problem act as the principal, the government being the agency.

Nevertheless, the individuals that constitute the general public have no incentive to make the investment of having a good understanding of the technology and economics of utilities. In order to solve this problem, they
elect political representatives and support the creation of regulatory agencies who act as informational intermediaries. The role of these intermediaries is to reduce the information asymmetry between the industry and the rest of the society. But this expertise creates the second agency problem: it is precisely because the government possesses an expertise the general public lacks, that it can get away with policies that favor a determined group to the detriment of the rest of the society. This also explains why the influence of a determined group is not necessarily offset by interest groups with opposite interests.

c. Regulatory commitment. Given that governments, as public decision makers, have short-term horizons (usually, until the next election), static efficiency may collide with dynamic efficiency. Under certain circumstances (political reasons, general macroeconomic policies or other constraints), governments maybe reluctant to raise tariffs when needed, which might discourage appropriate maintenance and new investments. Governments may be tempted to implement this sort of policies because their effects might only be noticeable in the long-term.

An example of this occurred in Argentina after the large devaluation of the currency that occurred in the period 2000-2001. The government broke its commitment to set tariffs for utilities maintaining its value in us dollars by setting new ones in (devaluated) Argentinean Pesos, causing great losses to investors. As a consequence, exploration activities were drastically reduced: in 2002, no wells were drilled, and in 2003, only one. The country's natural-gas reserves dropped 21% between 2000 and 2003 (Honoré, 2004). However, the government’s approval ratings soared at the time the decision was made (The Economist, 2005b).

Moreover, the lack of regulatory commitment may exacerbate the negative consequences of a problem known as the ratchet-effect (Viscusi, Vernon and Harrington, 2000). As said before, schemes that provide good incentives to a regulated firm to reduce costs are those that reward it with €1 for each €1 of cost reduction. This problem will arise because, under these schemes, the firm will only receive this reward at the next tariff revision, when cost reductions will be passed on to consumers in the form of reduced tariffs. This reduction, however, will impose higher performance requirements from that period on. The ratchet-effect consists on the reluctance of regulated firms to pursue full-extend cost-reductions to avoid higher performance requirements in the future.

Given the economics of long-lived infrastructures, the lack of regulatory commitment may translate into governments trying to renegotiate concessions conditions during the period the regulated firm is making profits (after long periods of cost-recovery). This likelihood may increase a regulated firm’s reluctance to cost-cutting.
On the other hand, the regulated firm may also force the government to renegotiate and offer more favorable terms if it can make credible threats of bankruptcy, or when it can show that the initial terms of the regulatory contract are not profitable enough to allow it to carry on the investments the government regards as necessary. As we can see, the likelihood of ex post contract renegotiation (which is inversely proportional to the government’s degree of regulatory commitment) punishes the firm’s cost-reduction efforts, thus rewarding its inefficiency.

Regulation may also fail due to a mismatch between regulatory framework and a country’s institutional characteristics. For example, it is hard to find in developing countries the sophisticated professional and administrative capabilities that regulation requires. Therefore, an institutional arrangement that does not take into account these limitations may fail because regulated firms can manipulate the process or because regulators do not follow the right procedure (Estache and Ginés de Rus, 2000). To avoid this problem, it may be necessary to reduce the regulator’s discretionary powers by introducing rigidities in decision making. However, this solution may also cause inefficiencies by imposing higher administrative requirements and unnecessarily dilate regulatory decisions.

2.2.3 Regulator’s objective function

It has been stated before that regulation can take many forms, depending on the country’s economic, political and social characteristics. However, considering the special knowledge and skills that regulation requires, is common for governments to create regulatory agencies to oversee one or several industries (Banco Mundial, 1997).

In formal terms, the regulator’s objective function is usually defined as the sum of consumer and producer surplus, although not necessarily with similar weights (Braeutigam, 1989). Therefore, the problem for the regulator is to set a price that maximizes its objective function $W(p)$:

$$W(p) = E(p) + \alpha \cdot \Pi(p)$$

Where $E(p)$ and $\Pi(p)$ represent the consumer and producer surpluses, respectively; and the factor $\alpha$ reflects the social valuation of the interests of producers and consumers. $\frac{dE(p)}{dp} < 0$ and $\frac{d\Pi(p)}{dp} > 0$, which means that consumer surplus reduces as prices rises and the contrary occurs to producer surplus. Therefore, there is an inverse relation between $E(p)$ and $\Pi(p)$.

Transforming $W(p)$, we obtain:
\[ W(p) = \left[ D(p) \cdot dp - p \cdot D(p) + \alpha [p \cdot D(p) - C(D(p))] \right] \]

Where \( D(p) \) denotes the demand function and \( C \) denotes the total cost of production. \( W(p) \)'s first-order condition is \( \frac{dW(p)}{dp} = 0 \).

Since \( \frac{dE(p)}{dp} = -D(p) \):

\[
\frac{dW(p)}{dp} = -D(p) + \alpha [D(p) + p \cdot D'(p) - MC \cdot D'(p)] = 0
\]

\[ 1 - \alpha \cdot D(p) = \alpha \cdot D'(p) [p - MC] \]

Since \( \varepsilon = \frac{D'(p^*)}{D(p^*)} \cdot \frac{1}{p} \), the price that fulfills the first-order condition and maximizes the regulator's objective function is:

\[ \frac{p - MC}{p} = (\alpha - 1) \cdot \frac{1}{\alpha \cdot \varepsilon} \]

According to this expression, when \( \alpha = 1 \), the optimal price is the one that matches the marginal cost. For \( 0 < \alpha < 1 \), the optimal price is located below the marginal cost, whilst for \( \alpha > 1 \), the optimal price approaches the monopoly price.

According to Armstrong, Cowan and Vickers (1994), there are two main reasons why \( \alpha \) may take values lower than 1. First, it may be socially preferable to transfer income from shareholders to consumers, since the average income of the shareholders tend to be higher than those of the consumers. And second, the regulator may try to balance the shareholder’s capacity of influencing regulatory decisions by adopting \( \alpha \) lower than 1.

In countries such as US and UK, where governments tend to refrain from intervening in market, regulators are expected to adopt \( \alpha \) closer to one. In more socially-oriented economies like the ones in continental Europe, a smaller \( \alpha \) would be expected (Lasheras, 1999).

### 2.2.4 Regulation methodologies

There are two pricing methodologies that are most used by regulators: rate-of return and price-caps (Kahn, 2004).
Rate-of-return

This methodology has been widely used in monopoly pricing, and is common in Japan, Canada and the United States (Intven, Oliver and Sepúlveda, 2000). It consists of the regulator setting prices so that the regulated firm earns only a fair return on its capital investments. Prices must satisfy the following condition:

\[ \sum_{i=1}^{N} p_i y_i = C(Y) + gI \]

Where \( p_i \) and \( y_i \) are the price and quantities sold of the good “i”, “C” represents the costs and expenses incurred to provide “N” goods, “g” is the return on investment and “I” represents the investments which will earn such return.

This regulation has three components:

a. The rate base, constituted by the investments that will be allowed to earn a rate of return (usually plant minus depreciation and working capital).

b. The rate level, which refers to the relation of overall revenues to costs.

c. The rate structure, which refers to the individual prices charged for different services to different consumers.

The most important activities of this form of regulation is deciding what kind of investments will be included in the rate base, and selecting the appropriate rate of return. Since both variables are estimated using methodologies that are selected through subjective criteria, rate-of-return regulation makes regulatory decisions more easily challengeable before the courts, and may convert the judiciary into the ultimate regulator. This characteristic may introduce more inefficiency to the regulatory process.

The main strength of this price-setting policy is that it insulates the firm against unforeseen events by allowing it to recover unexpected costs. However, it suffers from three essential problems (Guasch and Spiller, 1999):

a. It provides little incentive for productive efficiency, since it allows firms to pass all production costs onto consumers in the form of higher prices.

b. Since the firm earns a return over its capital investments, this method encourages the firm to invest excessively.

c. The regulator needs a high degree of discretion to implement it. This facilitates regulatory capture by regulated firms, thus inducing rent-seeking behavior from monopolists.
Price-caps

Price-cap regulation is increasingly replacing rate-of-return as preferred pricing methodology. It is the most common form of regulation in Europe for privatized gas, telephone, electricity and water companies. It is also the price-setting policy used by regulators in the US, Singapore, and several countries in Latin America (Intven, Oliver and Sepúlveda, 2000). It has been the methodology to regulate prices for monopolistic port terminals in Colombia, Australia and Peru.

Under price-cap regulation, the firm is allowed to raise its prices between review periods, at the rate of the Retail Price Index (RPI or inflation rate), minus some amount (the X factor) chosen to reflect expected increases in productivity (this is the reason why this scheme is also called RPI-X). The main benefit of this methodology is that allows producers and consumers to share the risks and rewards. It also provides incentives for cost-cutting, since the firm may keep any cost savings until the end of the review period. In this way, a profit-maximizing concessionaire will maintain its incentive to operate efficiently, and this effort will be passed on to the consumers in the form of price reductions at the next review period (Bernstein and Sappington, 1998).

As we can see, this system avoids the drawbacks of rate-of-return regulation by setting incentives to reduce costs and achieving productive efficiency. However, the use of this price-setting policy has some shortcoming as well (Guasch and Spiller, 1999):

a. **Regulatory capture.** The profitability of a regulated firm depends on the price set by the regulator. If governments want to concession more infrastructures or attract further capital to the industry, the regulator may be tempted to keep prices high.

b. **Regulatory risk.** The use of the RPI-X methodology may distort investment incentives if the public policy towards private provision of public services changes with the government’s political agenda.

c. **Cost calculations.** The use of this methodology does not reduce the information asymmetries between regulator and regulated firms described before, and may even exacerbate them because of the use of forward-looking costs to set the X factor.

d. **Cross-subsidization among regions.** The regulator may receive public pressure to set the same prices across different regions even when costs differ in different geographical markets, thus forcing the monopolists to cross-subsidize certain services based on the location of the infrastructure.

e. **Entry deterrence.** The use of price-caps that allow rate rebalancing among services may be used by monopolists to cross-subsidize services sold in
competitive markets with those sold in monopolistic ones, thus deterring entry in potentially competitive complementary markets.

f. Access pricing. As we will see in the next chapter, access prices should be set to allow the monopolist to recover all costs incurred, but, at the same time, they should not deter competition in complementary markets (pilotage, towage, warehousing). Regulators receive pressures to use access-pricing methodologies that favor either the monopolist or its competitors.

g. Regulatory lags. Cost reductions that can be gained from privatization of utilities tend to diminish over time. This increases the possibility of the government overestimating cost reductions. If the review period is longer that the lag between expectations and reality, it can lock the regulator and the regulated firms into an inappropriate framework that is difficult to adjust. On the other hand, too frequent reviews may increase regulatory risks and deter investments.

Comparing rate-of-return and price-caps

This section compares both regulation methodologies under several criteria. It can be seen that despite the increasing popularity of price-caps, they are not better for all scenarios.

Productive efficiency

As we have seen, rate-or-return regulation biases investments decisions towards a more-intensive-than-optimal capital use. Price-cap regulation corrects this bias because it allows the regulated firm to appropriate the surplus generated by cost cuts until the end of the review period. By doing that, price-caps incentivize incumbents incentives for cost reduction. Considering these reasons, we can conclude that, in principle, price-caps offer better incentives to allow productive efficiency than rate-or-return regulation.

Allocative efficiency

Rate-or-return regulation constitutes a better scheme to correct allocative inefficiencies, since it relates the firm’s revenues with the costs actually incurred. On the contrary, since price are fixed while costs decrease between review periods, price-cap regulation dissociates revenues and costs, thus allowing allocative inefficiencies.

Speed and likelihood of convergence towards the optimum

With price-cap regulation, an inadequate estimation of the X factor may prevent prices from converging towards the optimum. Indeed, if this factor is set below the optimum, regulation will generate monopolistic rents and consequent inefficiencies.

---

5 This section is based in Gallardo (1999)
that will slow down convergence towards the optimum. On the contrary, setting this factor above optimum will discourage future investments, therefore deterring innovation. For these reasons, it is preferable to use rate-of-return regulation when non-economic factors play a substantial role in price-setting.

**Monopolistic rents**

Price-cap regulation explicitly allows the firm to obtain monopolistic rents, and the appropriation of such rents is the incentive used to promote productive efficiency. However, rate-of-return may also generate rents due to information asymmetries discussed earlier. However, since information problems affect both schemes but are only explicitly allowed in the former case, it seems logical to conclude, *a priori*, that rate-of-return regulation delivers better results in avoiding monopolistic rents.

**Information requirements**

Both schemes require large quantities of information, but under rate-of-return regulation this requirement is higher for the regulator, who needs to know not only the cost of the capital investments incurred by the regulated firm, but also its current value. This requirement also exacerbates the problem of information asymmetries discussed before. For these reasons, rate-of-return regulation requires more information than the price-cap scheme.

**Learning costs**

As we have seen, the effectiveness of regulations depends on the information available to the regulator. However, information is usually valid only for a certain period of time, after which becomes obsolete.

Gallardo (1999) argues that price-cap regulation seems to be more appropriate in industries characterized by rapid innovation, since the accumulation of information by the regulator becomes less relevant. On the contrary, if an industry is characterized by slow technological change, the cost of learning is lower, which reduces the usual information asymmetry between regulator and regulated firm. In this context, rate-of-return regulation seems to be more appropriate.

### 2.2.5 Regulatory dilemmas

A regulator has to face five theoretical dilemmas when trying to maximize its objective function (Gallardo, 1999).
Dilemma 1: Productive or allocative efficiency

When the costs of an industry are sub-additive, it is desirable to have only one firm supplying the entire market. The entry of more firms to the market would increase the industry's average costs, which would lead to productive inefficiencies. This is the argument in favor of restricting the entry of firms to a market, which is especially relevant in industries with substantial fixed-costs such as telecommunications, energy or transport infrastructure. This restriction would allow a coordinated investment program, avoiding duplication of investments and destructive competition.

However, the existence of only one firm in the market would concede it substantial market power, increasing the probability of occurrence of the undesired behavior described in section 2.2.1. As we have seen, a monopolist has incentives to charge a mark-up above its costs, increasing prices and causing allocative inefficiencies. Therefore, the solution to the problem of productive inefficiency creates a problem of allocative inefficiency.

There are four methods a government can organize economic activities within an industry to solve this problem, all of them with advantages and disadvantages.

a. State-owned companies. A commonly used solution has been the creation of public enterprises or state-owned companies, which are believed not to have incentives to use their market power.

b. Concession to a regulated firm. Another way to organize the production of goods in the context of natural monopolies has been through the concession of the industry to a private firm whose activities are supervised and regulated by a regulatory agency. This case usually combines access restrictions to solve the problem of productive inefficiencies, with price regulation to solve the problem of allocative inefficiencies.

c. Ex ante competition. A third alternative to solve this dilemma is the introduction of ex ante competition through auctions, with the winner being the bidder who offers the lowest prices (Demsetz, 1968). In this case, the problem of allocative efficiency is minimized, whilst the granting of monopoly rights solves the problem of productive inefficiencies.

d. Liberalization of the industry. The fourth option to solve this problem is the liberalization of the industry (Baumol, Panzar and Willig, 1982). Under certain circumstances (price sustainability, for example) this can reduce the degree of inefficiency. As a result, potential competition would impede extraction of economic rent, considering that overcharging may allure competitors, thus solving the problem of allocative inefficiencies. If the threat of competition disciplines monopolists, the likelihood of firms entering the market (thus causing productive inefficiencies) would also be minimized, and
incumbent’s prices would not diverge much from marginal social opportunity costs.

However, all of these options have important disadvantages to consider. In first place, the creation of public enterprises was supposed to solve the problem because of their lack of incentives to abuse their market power. However, this view overlooks the agency problem between society and the management of such enterprises, something that, in countries with weak institutional arrangements such as developing ones, may lead to gruesome inefficiencies. Indeed, the view that state-owned companies will serve the public interest better than private ones implicitly supposes that the interests of their management are aligned with those of the consumers, which is not necessarily true. According to Tirole (2006), there are various ways in which management may not act in the owner’s best interests (in this context, the public’s best interests): insufficient effort, unnecessary investments and self-dealing, among others. These agency problems (Laffont and Martimort, 2001) seem to be exacerbated in developing countries, where the lack of an adequate system of checks and balances facilitates public officers making decisions that favor their political parties or themselves.

According to Kessides (2004), a main cause of deteriorating infrastructure performance in developing countries has been underinvestment, which was largely due to the failure of governments to prescribe cost-reflective tariffs. Under state ownership, prices fell to levels that could not cover the investment needed to meet the demand, which led to a significant infrastructure deficit and substantial welfare losses. These inefficiencies constrained domestic growth, reduced international competitiveness, and discouraged foreign investment. The World Bank (2003) estimates that technical inefficiencies related to public management in developing countries’ roads, railways, power, and water, caused losses estimated at US$55 billion a year in the early 1990s. This figure is equivalent to 1% of their GDP, a quarter of their annual investment in infrastructure, and twice the annual development finance for infrastructure in the developing world. In the port sector, many of the large public monopolies have been quite effective in preventing competition even where there are allowances for private facilities, such as in Philippines or India (ADB, 2000).

Secondly, downside of using regulation to deal with monopolies is the occurrence of the regulatory failures described in the previous section; and the use of auctions may cause substantial problems when unforeseen events occur due to the rigidity of the approach (this will be discussed in detail when analyzing the third dilemma). Finally, the liberalization of an industry may not deliver the results expected by Baumol, Panzar and Willig, (1982) if the presence of sunk costs (typical in the case of infrastructure-based industries) or asymmetric information is taken into account. In these cases, liberalization may lead to “cream-skimming”, i.e., competitors entering only to the profitable segments of the market and leaving the incumbent with no option to finance the deficit in the unprofitable ones. This view also supposes that firms can enter or leave a market easily, which is not likely when sunk investments need to be made.
Since there is not an *a priori* better solution to this dilemma, this will depend on the advantages and disadvantages of these options relative to the characteristics of each country or industry. In ports, for example, it is common to find port authorities of state-owned companies providing services, whilst in telecommunications; the most common way to deal with this problem is to concession monopolies to regulated firms.

**Dilemma 2: Optimal, subsidy-free, sustainable or equitable pricing**

If the government chooses to organize economic activities by concessioning the natural monopoly to a private firm, the problem of selecting the most appropriate pricing philosophy arises. As we will see, the prices that maximize the regulator’s objective function are not necessarily subsidy-free, sustainable or fair, and the criteria of optimal pricing may collide with other efficiency or distributional concerns, such as the minimization of inefficient entry or the elimination of cross-subsidies.

**Optimal prices**

As we have seen, the regulator’s objective is to set prices that maximize the function $W(p)$. We have also seen that the price that maximizes this function is the one that equates to marginal cost (the first-best). However, with large infrastructures such as ports, where fixed-costs represent a large share of total costs, setting prices at marginal cost would not allow firms to recover their fixed-costs. For these infrastructures, the regulator has to set prices that cover the long-run average or marginal cost of the monopolist.

There are two main options for doing this. The first consists on setting prices at marginal cost and cover the consequent revenue deficit with government transfers. This option, that would allow the achievement of the first-best, is called the “Hotelling approach” and will lead to Pareto-efficient results if the transference is financed through a lump-sum tax paid by all the population, regardless of their level of consumption of the service (Hotelling, 1929). However, this characteristic makes this tax-and-subsidy scheme practically impossible to be implemented by industry regulators.

The second alternative is to set prices equal to marginal costs plus a surcharge that allows the covering of fixed costs. Since this option would introduce economic inefficiencies, it would only allow achieving a “second-best” solution; i.e., welfare maximization subject to the restriction that the firm covers all its costs. The main approach for second-best pricing is known as the “Ramsey rule” or “invert-elasticity”, and consists of adding a surcharge to marginal cost that would be higher for consumers with low demand elasticity and lower for those with high elasticity.
The prices obtained through the Ramsey rule are considered optimal ones. The conditions that allow obtaining the invert-elasticity rule according to Ramsey criteria are the following:

a. There are no cross-elasticities among different market segments.

b. The rent-effect caused on the individual behavior of consumers does not substantially alter the analysis of elasticities on demands that do not contemplate such effects.

When these conditions are met, the Ramsey criterion lead to a relatively simple rule of price setting for different types of consumers that can be defined by the following statement: when there exist different products or market segments with different demand elasticities, the optimal way to distribute fixed-costs is by a margin over the marginal costs, which is inversely proportional to the demand elasticity in each segment. This is an optimal criterion, since the loss of welfare produced by increasing the price above marginal cost is minimal when such mark up is charged to consumers with more rigid demands.

To obtain the expression that sets optimal prices as a mark up over marginal costs, it is necessary to define \( p = (p_1, p_2, p_3, \ldots, p_n) \) (where \( p_i \) represents the price of a service in segment “i”) and incorporate the concept of fixed costs (FC). The function \( \Pi(p) \) then becomes:

\[
\Pi(p) = \sum_{i=1}^{n} p_i \cdot D_i(p_i) - FC - \sum_{i=1}^{n} C(D_i(p_i))
\]

Where \( C \) is the firm’s cost function.

The problem for the regulator is to set prices that maximize the welfare function \( W(p) \) subject to the restriction \( \Pi(p)=0 \) (in this way the monopolist is allowed to recover all its costs) (Lasheras, 1999). To maximize \( W(p) \) subject to the restriction \( \Pi(p)=0 \), we need the following Lagrange function:

\[
L : W(p) - \lambda [W(p)] = 0
\]

Where \( \lambda \) represents the multiplier of the restriction. Differentiating this function with respect to the price \( p_i \), we obtain the following first-order condition:

\[
\frac{\delta W(p)}{\delta p_i} - \lambda \frac{\delta \Pi(p)}{\delta p_i} = 0
\]

\( ^6 \) This solution was initially proposed by Ramsey (1927) for tax purposes and was applied by Boiteaux (1956) and Baumol and Bradford (1970) to the problem of optimal tariffs for natural monopolies facing marginal costs below average costs.
If for the sake of simplicity, the same weight is given to producer and consumer surpluses in the $W(p)$ function, this expression equals to:

$$- D_i(p_i) + (1 - \lambda) \cdot \left[ D_i(p_i) + p_i \cdot \left( \frac{\delta D_i(p_i)}{\delta p_i} - MC \cdot \frac{\delta D_i(p_i)}{\delta p_i} \right) \right] = 0$$

Regrouping terms and substituting the demand elasticity of the good “$i$” with respect to its price, we get:

$$\frac{p_i - MC}{p_i} = -\frac{\lambda}{1 - \lambda} \cdot \frac{1}{\varepsilon_i} = \frac{\delta}{\varepsilon_i}$$

The expression “$\delta$” is called the “Ramsey number” and represents the global price level of the regulated firm, whilst each market segment’s elasticities configure the relative structure of the price vector.

The use of the Ramsey rule has been criticized, among other reasons, because it requires the regulator to have large amounts of information that is very difficult to estimate, such cross-elasticities among different market segments (Miller, 2007).

Equitable prices

As seen in the previous section, the use of optimal (Ramsey) prices would imply that consumers with inelastic demand pay prices above marginal cost, while consumers with elastic demands pay prices below marginal costs. In the case of utilities, low-income consumers are usually those whose demand for public services is more rigid, and who, under the Ramsey rule, would be charged with higher prices. In this context, the implementation of the rule for public services may not be socially viable, especially in societies with unequal income distribution, such as the developing countries (Prieger, 1996). In ports, where terminals are increasingly privatized, the use of Ramsey prices would lead to carriers paying a disproportionately large share of the port’s costs, since they face a rigid demand.

It can be seen that whereas the Ramsey rule leads to economic efficiency, it does not guarantee equitable prices.
**Subsidy-free prices**

A natural monopoly may have incentives to set prices that imply the use of cross-subsidization. For example, if a monopolist produces two goods, but one of them is sold in competitive markets, the firm has incentives to eliminate competition by cross-subsidizing this good with the revenues obtained in the regulated market. If joint costs exist, this could be difficult to prove, since determining the fraction of the cost that must be imputed to each good is a complex process and often impossible. This situation is more likely to happen when the Ramsey rule is used, since goods sold in competitive markets are more elastic than those sold in monopolistic ones. In the port sector, this situation may arise when port authorities that provide infrastructure to third parties are also involved in cargo-handling activities. They are able to cross-subsidize total handling costs by not taking into account port dues.

Naturally, the fact that optimal prices are not necessarily subsidy-free creates a problem for the regulator, who might have to opt between prices that maximize consumer surplus and subsidy-free ones, especially in contexts where an industry is being liberalized or competition is being used to increase economic efficiency. The solution to this dilemma becomes even more complex when the monopoly provides the same service in different regions, for example, telephony or electricity in rural areas. In this case, the regulator will have to opt between setting prices that reflect the different costs of providing the service in diverse areas or setting the same price in all areas, thus allowing cross-subsidization.

As we can see, optimal (Ramsey) prices are not necessarily subsidy-free.

**Sustainable Prices**

When several products are provided by the monopolist, the incursion of competitors is possible when some goods, segments or lines are profitable. For example, the operation of a port terminal may not be profitable, but the provision of pilotage, stevedoring or towage services may be lucrative and attract other firms who want to provide these services without having to operate the port. In this case, the operation of the unprofitable service is only sustainable if the monopolist can obtain surpluses in related services.

According to Baumol, Panzar and Willig (1982), for the prices of a monopoly to satisfy the condition of sustainability, there should not be a competitor that covers a segment of the demand, charges lower prices, and, at the same time, generates profits. However, Ramsey prices do not necessarily satisfy these conditions. In the described case, the competitors in the provision of pilotage, stevedoring or towage services may very well charge lower prices and make profits (Train, 1991).

As we can see, optimal (Ramsey) prices are not necessarily sustainable.
Dilemma 3: *Ex ante* competition or flexible regulation

As stated before, an alternative to price regulation consists on implementing auctions to grant the concession of the natural monopoly to the firm offering the lowest prices. The central idea was initially proposed by Demsetz (1968), who claims that it is possible to achieve optimal results without regulation by generating *ex ante* competition (the “Demsetz Approach”). This process is called competition “for the market” because it differs from the usual one that occurs between rivals within a market (competition “in the market”).

In the absence of collusion, equal access to essential inputs and symmetric information among the bidders, the auction would make prices approach the average cost of the most efficient firm, thus minimizing simultaneously productive and allocative inefficiencies.

Besides this argument, the use of auctions might be desirable for two other reasons:

a. Auctions would solve the problem of the government’s lack of information about costs and demand.

b. Auctions may reduce the cost of regulation by not requiring the existence of a regulatory agency.

However, Demsetz’s proposal works well only when it is possible to specify, in the concession contract, how prices will be adjusted to reflect changes in market conditions. This implicitly supposes that it is possible to anticipate the occurrence of future events, or that it is possible to modify existing contracts without generating important costs for society. Following this argument, Williamson (1976) criticizes Demsetz’s approach because it supposes the existence of perfectly designed contracts that specify the actions to follow before any contingency. He argues that contracts are in essence incomplete for the following reasons:

a. It is difficult to anticipate future contingencies;

b. Even when contingencies may be anticipated, it may still be difficult to negotiate measures to take in each case;

c. It is difficult to write contracts in a language that allows third parties to interpret them and settle all kinds of disputes.

Two mechanisms have been suggested to incorporate modifications to concession contracts: long-term contracts that allow planned investments, and repeated auctions of short-term contracts that allow the incorporation of new information.

Williamson also criticizes both types of contracts. The use of long-term contracts may lead to efficient results only if the gains or losses generated by unanticipated events
are shared by the parties, i.e., if it is possible to introduce modifications that reduce allocative inefficiencies. However, inefficiencies may arise because the monopolist has incentives to manage the information strategically, or because the government may prefer to allow the contract to expire before acknowledging any error in its design. On the other hand, the success of the use of consecutive short-term contracts depends on the equality of conditions among the concessionaire and the potential entrants. The main problem is that incumbents always have advantages over potential entrants, such as a deep knowledge of the demand, better information about operation and technology costs and human resources with specific experience.

The introduction of government supervision on the firm’s accounting or operations would reduce the problems or both types of contracts, but this would not differ from standard regulation. For this reason, Williamson does not consider the use of auctions as a substitute to regulation, but as an alternative method of organization.

As we can see, the concession of an unregulated industry through the use of auctions may generate problems that reduce its apparent advantages over the use of standard regulation, making it impossible to decide \textit{a priori} which mechanism is superior.

Dilemma 4: Economic efficiency or informational rents

The theory of optimal prices assumes the regulator has the same information as the regulated firm. However, this is not true. As a matter of fact, firms possess better information than the government about important endogenous and exogenous variables, such as technology, costs and demand. In a context like this, where the regulator faces informational restrictions, it will have to opt between achieving productive and allocative efficiencies, and avoiding the monopolist obtaining extraordinary rents at the expense of the consumers. To illustrate this, Baron and Myerson (1982) developed a model in which there is hidden information about costs. In this model, the regulator’s aim is to maximize productive efficiency by setting prices at marginal cost, for which he is willing to grant a transfer to encourage the monopolist to disclose its marginal cost. However, for the incentive to work, the monopolist’s profit from the transfer must be larger than the profit produced by declaring any other cost; for which this transfer has to be as large as the consumer surplus. In terms of productive efficiency, this strategy would lead to a first-best solution, but in terms of distribution, the firm would enjoy excessive profits paid by the consumers with their surplus.

If this option is not viable, the regulator can improve allocative efficiency by granting a smaller transference, although it would have to be at the expense of reducing productive efficiency. The price, then, will be set above marginal costs. Nevertheless, if the regulator has an objective function (as the one described earlier in this section) with $0 < \alpha < 1$, where $\mathcal{E}$ of consumer surplus is worth more than $\mathcal{E}$ of profits, it will be satisfied with the monopolist obtaining economic rents if by doing so it increases consumer welfare.
As we can see, in a context of information asymmetries, a regulator will have to decide whether to achieve productive efficiency by granting an economic rent or to improve surplus distribution by setting prices above marginal costs.

Dilemma 5: Rate-of-return or price-cap regulation

When implementing a price-setting policy, the regulator will have to opt between the two main schemes: rate-of-return and price-cap regulation. As we have seen in the previous section, the rate-of-return methodology translates cost increases directly to prices, thus reducing the risk for the monopolist but eliminating its incentives to operate more efficiently. Price-cap regulation provides better incentives for achieving efficiency by making the firm the residual claimant of costs reductions.

Both are opposed in terms of the relation costs-revenues and none can be considered superior *a priori*.

2.3 Options to introduce competition in natural monopolies

According to Klein (1996), there are three main options to introduce competition in natural monopolies: competition for the market, competition over existing networks and competition among networks.

2.3.1 Competition for the market

It consists on granting the right to operate a monopoly to the operator who offers to charge the lowest price. This approach was explained while discussing the dilemma between ex ante competition or flexible regulation, in section 2.2.4.

This option works well with monopolistic activities that *do not* involve a large amount of sunk costs, such as towage services in Australian ports (Ergas, Fels and Soon, 2004). In these cases, the concession can be auctioned periodically and new information can be easily incorporated in the contracts. But since most infrastructures have the characteristic of being long-lived, and it is impossible to write complete contracts that foresee every contingency, problems are found in cases when long terms are required. In these situations, concessions are granted using auctions but incorporating a schedule for price-revisions and even the methodologies to be used for that activity.
2.3.2 Competition over existing infrastructures

Typically, large infrastructures are used to produce services for several markets, not all of them naturally monopolistic. Therefore, it is possible to minimize the risks of regulatory failure by introducing competition in non-monopolistic markets (potentially competitive ones). Klein and Gray (1997) argue that there are three options to introduce competition using existing infrastructures: open access, time-tabling and pooling arrangements. The suitability of each for a particular industry depends on the nature of the services and the infrastructure over which these services are to be provided.

Open access

The main characteristic of network industries is that they require complementary inputs, produced in both monopolistic and competitive markets, to produce a single service (see chapter 6). In the telecommunications industry, long-distance and mobile services are usually considered competitive, while fixed telephony is not. In electricity and natural gas, generation is considered a competitive segment whilst transmission and distribution naturally monopolistic ones. In these circumstances, competition in non-monopolistic markets is not possible unless competitors have access to the monopolistic segments of the network. For example, long-distance carriers cannot compete unless they have access to the fixed-telephony network, since a call cannot be completed unless both networks are interconnected (Intver, Oliver and Sepulveda, 2000).

In transport, air, sea, rail and road freight are usually considered competitive activities, while some port terminals, airports, highways and railways may be natural monopolies. It requires services from both kind of markets to complete the transport chain. For example, neither airlines nor those who provide services to them (ground handlers, caterers) will be able to compete unless they have access to the airport (Betancor and Rendeiro, 1999). Although a port terminal is not a network, the same rationale applies. Providers of pilotage, towage or storage cannot enter the market unless they have access to terminals. The problem is that in small ports, usually located in remote or poor areas, building a second terminal may be uneconomical. This fact converts the terminal operator into a monopolist. If the terminal operator also provides other services that are needed to complete the transport chain (shipping, pilotage, towage, storage, etc.), regulation of the monopolistic activity gives the monopolist incentives to try to foreclose competition in the related activities it participates (Paredes, 1997).

---

7 The terminal operator can be a private company (such as the operators of Matarani port terminal, in Peru; or Cartagena’s container terminal in Colombia), a state-owned company (such as Göteborgs Hamn AB, the operator of several terminals at Gothenburg’s port) or a port authority (such as Cyprus Port Authority, operator of several terminals at five Cypriot ports).
Under open access, the government, usually through a regulator, fixes the terms and conditions under which firms participating in competitive markets acquire access to naturally monopolistic facilities (Valletti and Estache, 1998). The aim of this intervention is to enhance the contestability of these markets (see appendix 2 for the characteristics of contestable markets). When implementation is possible, open access does not require vertical separation of activities. This means that the owners of naturally monopolistic facilities are not restricted to participate in related competitive market, therefore it is possible to preserve economies of scope and coordination. According to OECD (2001), the implementation of an open access regime is most effective when the capacity and the quality of the non-competitive segment are easy to observe. In this case, the regulator has only to ensure that all requested capacity is made available under non-discriminatory terms.

The open access option will be analyzed in more detail in the next chapter.

**Timetabling**

A theoretical option to introduce competition is timetabling, which allows several firms to use the same infrastructure according to a predefine schedule. In theory, this option would allow rail tracks or time slots at airports to be used by several competitors, thus generating competition and fostering the efficient use of expensive infrastructures. However, as indicated by Klein (1996), the implementation of this option faces practical complexities that deter its use. Indeed, in industries such as electricity or gas the homogeneity of the product makes the identity of the producer irrelevant. But in shipping, airlines of railways, where passengers or freight need to reach a particular destination; the requirements for route optimization are more complex than matching electricity or gas inflows with outflows. For example, in order to define rights to use rail tracks and allocate them to multiple parties, secondary trading would yield the optimal set of paths that maximizes welfare given valuations by producers and consumers. However, since the value of each right to use a segment of track depends on what happens with all the other segments, it is not clear if an optimal timetable can be generated through decentralized bargaining. Apart from the allocation of pairs of slots in some airports, there are no experiences with this system.

**Pooling**

This option applies only to electricity networks, where the physical characteristics of electric power and some distinctive features of electricity markets make it difficult to implement open access policies. Demand for electricity is highly seasonal (it even changes during the same day), but since power cannot be stored efficiently; supply must match demand at every time. In such a situation, competition can be introduced in the generation segment by establishing a central system that selects the generators who will dispatch electricity to the network according to their production costs, matching supply and demand at every time (Hunt, 2002). In Peru or Chile, for example, a central system dispatches orders on the basis of the audited costs of power
plants. Those that use cheaper technologies (such as hydropower) are dispatched first. Those which use more expensive technologies only dispatch at peak time. Under this system, transmission and distribution activities remain natural monopolies that require regulation.

2.3.3 Competition among infrastructures

The third option for introducing competition in network industries consists of promoting competition among different types of infrastructure, for example, by fostering competition among short-sea shipping, railways and highways, or between fixed and mobile telephony networks. Klein (1996) argued that two problems may arise with this approach. The first one is that in the presence of a natural monopoly, the establishment of an alternative network supposes welfare losses associated with the duplication of investments (Vickers, 1995). The author, however, claims that the pressures generated by competition to work hard, to learn and to innovate outweighs many costs of duplication. This view implicitly argues that the dynamic benefits of competition are larger than the static allocational benefits of preventing duplication. One example of this is the competition that occurs among global supply chains, of which ports are crucial nodes. The competition among these networks forces ports to be more competitive, thus limiting their potential to abuse of their market power. The second problem is the imperfect substitutability between products. In fact, competition between transport modes is only possible for a limited number of products and origin-destination pairs just as wireless telephony is an effective substitute for fixed telephony for some types of clients and geographical areas only. However, it can be argued that this substitutability, although imperfect, limits the market power of monopolists by making the demand curve more elastic.
Appendix 1: The Creation of a Pilotage Monopoly in the Netherlands\textsuperscript{8}

“In 1988, The Netherlands Pilotage Service became an independent organization, with pilots acting as private entrepreneurs. The objectives of the government in the privatization of pilot services were to reduce the governing executive burden and to improve efficiency and adequacy of pilot services.

A public entity, the Nederlandse Loodsen Corporatie (The Netherlands Pilot Corporation, NLC) was created to manage the register of licensed pilots and be responsible for education and training of licensed pilots. All licensed pilots constitute the NLC. (…) The licensed pilots are all shareholders of the Loodswezen Nederland BV (Pilotage Service of the Netherlands Ltd.) which is responsible for the exploitation of the independent private enterprise. All supporting staff is employed by this company. The company collects the pilotage fees and makes payments to the pilots in accordance with the financial statute. (…)

Privatization in The Netherlands did not bring an end to the debate about pilot services. The Government Audit Office directed harsh criticism at the privatization process and asserted that the efficiency improvements did not benefit the shipping lines or the government, but solely the pilots. Notwithstanding counter arguments, the Government Audit Office’s criticism, The Netherlands’ privatization of pilot services is not considered a successful one.

To a certain extent, the government’s objectives have been attained. The increase in the amount of pilot activity and the reduced number of licensed pilots have led to higher efficiency. However, pilotage became a virtual monopoly and efficiency improvements have led primarily to a very substantial rise in pilots’ incomes.

The cost structure of the pilotage organization is not transparent. The fees are non-negotiable, contrary to the fees for other marine services and pilot fees in other ports. The magnitude and rigidity of pilot fees create strong pressures to reduce other cost elements in the highly competitive maritime transport sector.

Overall, the present situation has proven unsatisfactory to port users”.

\textsuperscript{8} World Bank (2001) Module 3, P. 72
Appendix 2: Theory of Contestable Markets

A2.1 Contestable Markets

The “Theory of Contestable Markets” was initially proposed by William Baumol, John Panzar and Robert Willig (Baumol, Panzar and Willig, 1982). In their book, the authors proposed the concept of “contestable markets”. A contestable market is one in which entry is absolutely free, and exit is absolutely costless.

As explained by Train (1991), the terms “free” and “costless” have a particular meaning. “Free entry” does not mean that a new firm need not incur any cost to enter an industry, but that a new firm does not have to incur any cost that is not also incurred by a firm that is already producing in the industry. This concept therefore requires the entrant having access to the same technology and inputs as the incumbent. On the other hand, “costless exit” means that any firm that decides to leave an industry is able to recover all the costs it incurred when entering –minus the respective depreciation. This implies that no sunk costs are incurred by entrants.

According to this theory, when markets are contestable, the threat of potential competition can be more important than actual competition in generating competitive discipline to producers. Under perfect contestability, this threat would even force monopolists to behave as if they were in perfectly competitive markets. Indeed, if a monopolist is charging a price above competitive levels, a new firm can enter the market, charge a slightly lower price and capture the whole market. If the incumbent retaliates by lowering its own price, the entrant could simply leave the industry, recovering all costs incurred. Analogously, if the incumbent is producing inefficiently, a new firm could enter the market producing without waste, charge a slightly lower price and capture the whole market, earning a positive profit.

As it can be seen, a crucial feature of a contestable market is its vulnerability to hit-and-run entry. Nevertheless, this hit-and-run mechanism is only possible if there is an asymmetric time lag between the incumbent and the entrant. Indeed, the entrant must be able to enter the market, enjoy profits at prices above competitive levels, and leave before the incumbent can retaliate. As it will be explained later, this critical assumption has been a source of criticisms to the theory.

The authors recognize that perfectly contestable markets are as scarce as perfectly competitive ones in the real world. On the basis of this Baumol claims that “perfect
contestability is not useful to describe reality, but it should be seen as a benchmark in industrial organization that is more flexible and more widely applicable than perfect competition”⁹. In fact, one of the advantages of this theory is that the concept of the perfectly contestable market is a generalization of the concept of the perfectly competitive market, and it is possible to apply its analysis to any industry structure, including monopoly and oligopoly.

In the case of oligopoly, under perfect contestability, oligopolistic structure do not need to depend on the assumptions made about the incumbents’ behavior, but both, structure and behavior, are determined uniquely by the pressures of potential competition. In the case of a monopoly, however, the behavior of a monopolist under perfect contestability does not guarantee that the second best optimum is achieved, despite the restriction of financial viability and the presence of economies of scale.

It is worth noting that in contestable markets, firms need not be small or numerous or independent, not even produce homogeneous products. Therefore, the existence of a perfectly contestable market is a necessary condition for the existence of a competitive market, but not a sufficient one.

Baumol also argues that while the theory extends in some directions the domain in which market forces operate, it restricts it in others. For example, where the market structure is such as to yield a satisfactory allocation of resources within the period, it will also perform well between periods. But where there are economies of scale in the production of durable capital, a contestable monopoly that performs relatively well in a single period will not do so as time passes. In such a case, the least costly producer is in the long run vulnerable to replacement by rivals, creating a potential waste of social resources.

Until this theory was presented, industry structure was thought to be determined exogenously. Previous economic theory hardly explained the reasons why one industry was organized as monopoly or another as an oligopoly. In contrast, under perfect contestability, the structure of an industry is determined explicitly, endogenously, and simultaneously with the pricing, output and other decisions of the firms operating in it.

**A2.2 Properties of Contestable Markets**

According to the author, the two most important properties of contestable markets are their welfare attributes and the way in which they determine industry structure.

---

⁹ Baumol (1982), p.3
A2.2.1 Welfare Attributes

The welfare properties of contestable markets are drawn from their characteristics of free entry, costless exit and vulnerability to hit-and-run incursions from rival firms. The first welfare property is that, given that under perfect contestability even monopolists have to behave as if they were in competitive markets, contestable markets never offer more than a normal rate of profit. Any positive profit would encourage any firm to enter the market, undercut the price of the incumbent and earn a profit.

The second welfare property of contestable markets is the absence of any sort of inefficiency in production. The argument is similar as above. Any waste, like any abnormal profit, constitutes an incentive to entry. This is valid for inefficiencies derived from misallocation of inputs, inefficient management, x-inefficiencies, or inefficient organization of the industry.

The third welfare characteristic of contestable markets is that in the long-run, no product can be sold at a price lower than its marginal cost. If a firm lowers its price below marginal cost to drive competitors out of the market, it will not be able to recover its losses once they are left, since any price above marginal cost will entice hit-and-run entry. This characteristic has an important implication for regulatory and antitrust policy, since no predatory pricing can be used for unfair competition.

Baumol (1982) uses a different rationale to explain why in the long-run, no product can be sold at a price lower than its marginal cost. He explains that if some firm charges a price lower than its marginal cost and still makes a profit, then it is possible for an entrant to offer a smaller quantity at a slightly lower price and still make a profit. If the price $p$ is lower than marginal cost $MC$, then the sale of $y - \zeta$ units at price $p$ must yield a profit $(\pi + \Delta \Pi)$ which is greater than the profit $\Pi$, that can be earned by selling only $y$ units of the output at that price.

A2.2.2 Determination of an Industry’s Structure

To understand the way industry structure is determined under perfect contestability, it is important to recall the fact that contestable markets are incompatible with any sort of inefficiency, in particular, with inefficiencies in the organization of an industry. To illustrate this point, we can suppose that a certain output of an industry can be supplied by two or 1,000 firms, as done by Baumol (1982). But, if under the first arrangement the output can be produced at a lower cost, the industry cannot be in equilibrium if there are 1,000 producers. Therefore, we can expect a wave of mergers, acquisitions and/or bankruptcies in this industry until equilibrium is reached. The way market forces will lead this industry towards equilibrium will be described in the following paragraphs.

As explained when natural monopoly was defined, if economies of scale and/or scope hold throughout the relevant range, production by a single firm will be most
economical. Similarly, in the single product case and under constant returns to scale, if the minimum cost is obtained by producing 10,000 units and there is demand for 50,000 units, the most economical way to produce this output is by five firms. Therefore, the industry will be organized as an oligopoly. In the multi-product case, the analysis is slightly more complicated, but the logic is the same. When the industry’s output vector is small compared to the output vector a single firm can produce at relatively low costs, an efficient industry will be characterized by few firms.

From this analysis, we can see that the most efficient number of firms will vary with the location of the industry’s output vector. An industry may be a natural monopoly with one output vector, and a competitive market with another.

Note that the optimal structure of an industry depends on its output vector, but this output vector in turn depends on the prices charged by its firms. However, since pricing depends on industry structure, pricing behavior and industry structure must be determined simultaneously and endogenously.

The theory says nothing about how the structure of an industry is determined when the market is not perfectly contestable. But the author argues that while the industry structures that emerge in reality are not always those which minimize costs, they constitute reasonable approximations to efficient structures.

A2.3 Criticisms

As pointed by Waldman and Jensen (2001), the theory of contestable markets represented a dramatic departure from conventional economic views, for which its introduction drew much interest and scrutiny. However, as time passed, it became clear that the new theory constituted more an evolution than the revolution in economic thinking that its followers initially expected.

As Tye (1990) argued, critics have focused more on the theory itself than on its application. In particular, criticisms have concentrated on the assumptions required to allow hit-and-run entry.

The first main criticism to the theory was presented by Schwartz and Reynolds (1983). The authors argued that perfect contestability requires two conditions that are “implausible”: (i) in response to high prices, an entrant have to be able to enter the market instantaneously, with no time lag; and (ii), the entrant can exit the market with no loss of fixed costs before the incumbent can adjust his price, i.e., there must exist a time lag between the entrant entering the market and the incumbent’s response. Moreover, the authors claim that the theory is not robust, since any relaxation of these assumptions will make results differ from those obtained under perfect contestability.
Shepherd (1984) argues that while the theory focuses on several aspects of entry, their results hold only for the limiting and extreme case of perfect contestability. He agrees with Schwartz and Reynolds about the lack of robustness of the theory, by stating that under any departure from perfect contestability, the analysis becomes speculative. But the author mainly criticizes the consistency of the assumptions and the generalization of its recommendations. Indeed, he claims that the assumption of both total and trivial entry—the cases when the entrant takes all or some part of the incumbent’s market, respectively—are inconsistent. And even in the eventuality that it holds, it would lack of generality because of its extreme character.

Shepherd also claims that the assumption that external market conditions dominate internal ones is “eccentric”. He clearly states his belief that industrial organization is primarily about internal conditions, and criticizes the fact that the theory ignores actual competitors. Moreover, he claims that research and experience confirm the expected dominance of internal over external conditions.

Baumol and Willig (1986) argue that Shepherd incorrectly associates their position with an all-pervasive laissez-faire position on the role of regulation and antitrust. They argue that contestability theory provides guidance in determining when intervention is socially warranted, and that it has the value of providing a more applicable benchmark for policy makers when intervention is required. And about the criticism of the inapplicability of the theory in real-life situations, the authors state the following:

“The economy of reality is composed of sectors which vary widely in the degree to which they approximate the attributes of contestability. Thus, the conclusion that perfectly contestable markets require no intervention claims little more than the possibility (which remains to be proven, case by case) that some markets in reality may automatically perform in a very acceptable manner despite the small number of firms that inhabit them.”10

10 Baumol and Willig (1986), p. 10
3. Topics in access regulation

This chapter introduces the main topics in access regulation. The first section describes the debate on vertical integration in these industries and whether this should be permitted or prohibited. As we will see, this decision would have substantial consequences on the incumbent’s incentives to allow access and thus, alter the focus of regulation. The second section discusses one of the main issues of access regulation, namely access pricing. It describes the main methodologies used for this task, as well as their implementation problems. The last section introduces the concept of “essential facilities” and describes the doctrine that is commonly used in antitrust to prevent the abuse of market power by the owners of this kind of infrastructure.

3.1 Vertical integration versus vertical separation

As said before, naturally monopolistic and competitive segments co-exist in infrastructure-based industries. Typically, access to the former is necessary to render services in the latter. In these circumstances, when a firm that controls a naturally monopolistic segment is allowed to participate in the competitive ones, it has incentives to try to monopolize the competitive markets with the aim to regain the profits forgone by regulation. Therefore, some form of regulatory intervention is warranted. There are two main regulatory approaches to address this problem:

a. To forbid owners of the non-competitive segments to participate in potentially competitive ones, i.e., to enforce the vertical separation of the industry. Vertical separation implies, for example, that neither terminal operators nor their related firms can supply other services that are necessary to complete the logistics chain.

b. To regulate the terms and conditions under which participants in competitive markets acquire access to non-competitive segments of the industry.

OECD (2001) advocates vertical separation as a preferable approach. This document argues that vertical integration increases incentives on the incumbent to restrict
competition in competitive activities. Vertical separation, in contrast, would lessen the regulatory burden, thus enhancing the quality of regulation and the level of competition. They claim, for example, that as long as prices are set above costs, the incumbent has incentives to sell as much of its product at those prices. Therefore, rather than refuse access, the incumbent “has an incentive to welcome access, as each new entrant in the competitive market will enhance competition, innovation and product differentiation (...) enhancing demand for the non-competitive service”\textsuperscript{11}. Furthermore, the authors argue that allowing integration makes regulatory tasks more complex, since regulation must overcome the incumbent’s incentive to deny access. Given the information asymmetry between regulator and regulated firm, the risks of regulatory failure are increased. In contrast, by removing this incentive, vertical separation would allow lighter regulation which, in turn, may permit the incumbent to use its information in a more efficient way. For example, the regulated firm could have more discretion to use complex access pricing schemes, such as multi-part or peak-load pricing.

The incentive system that vertical separation introduces may also facilitate investments in new capacity that integrated firms would not have incentives to undertake, like new capacity that would be mainly used by rival firms. These incentives become crucial when, as in most cases, regulators lack legal powers to force incumbents to invest. Another reason the authors cite to advocate for vertical separation of related activities are the longer time that negotiations between incumbents and entrants may take, given the incentives the former have to delay entry, raise prices and lower quality. In addition, they claim that separation improves information by eliminating the use of transfer prices and, at the same time, reduces the possibility of cross-subsidization between regulated and non-regulated activities. Paredes (1997) analyzed the issue in the context of the debate on the convenience of vertical integration initiated in Chile in the mid-nineties. The author firstly presents Coase’s theory about the origin and nature of the firm (Coase, 1939). According to this theory, a firm is an institution that avoids the costs of using the market. For Coase, the activities carried out by a firm in an integrated manner can also be subcontracted to the market. However, the use of the market is not free. It generates transaction costs that under some circumstances can prevent using it. Therefore, the integration of those processes that substitute the costly use of markets constitutes the essence of the firm.

According to the author, economic literature identifies three factors that cause non-trivial transaction costs that avoid using the market:

a. Uncertainty about the future that causes the incompleteness of written contracts;

b. Uncertainty about the compliance word or implicit contracts; and,

\textsuperscript{11} OECD (2001), pp. 21.
c. Specificity of investments to third parties.

For example, a producer may contract transport services to carry his products to the market or integrate this activity within the firm. According to this theory, the producer will be more inclined to integrate transport within the firm if he or she is uncertain whether the supplier of transport services will comply with the contract or that will be able to find another one in the market. In this case, the likelihood of finding another supplier is reduced when a substantial part of his investments are specific to the original supplier, because of packing specifications, IT systems, etc. Therefore, according to the circumstances, the most efficient solution may be vertical integration. Forcing the separation of activities in these cases raises costs and damages consumers. However, if uncertainty can be somehow reduced or the cost of the service being provided by a specialized party is so low that compensates for any risk involved, the producer will naturally contract out transport services. The cost of using the market would be trivial and it would not be necessary to force separation. In fact, the advancements made in logistics and the rapid surge of third-party logistics providers confirms that this is what has occurred in the manufacturing industry.

Paredes also analyzes the argument that vertical integration facilitates the extension of monopoly power. If a monopoly is left unregulated, it is possible for the monopolist to maximize profits in the monopolized market. Thus, integration does not produce extra benefits to the monopolist. But when the monopoly is regulated (and the monopolist’s profits capped), he has incentives to recover foregone profits by integrating downstream and excluding rivals. It is worth noting that an integrated monopolist could only extract monopolistic rents if he is able to exclude rivals in the competitive segment. The author argues however that this is not the general case.

Vickers (1995) presents a theoretical model that is consistent with Paredes’ assertion. In this model, both imperfect information in the monopolistic segment and imperfect competition (there are more firms than needed) in the downstream market are supposed. The information asymmetry makes the regulation of the upstream (monopolistic) segment to be imperfect, hence allowing the incumbent to charge a price mark-up. Imperfect competition in the downstream market causes duplication of fixed costs. According to this model, when the incumbent is allowed into the downstream market, prices are higher and output is lower than in the case of vertical separation; but the number of firms in the imperfectly competitive downstream market, and hence, the number of times that fixed costs are incurred, is lower. Therefore, concludes the author, the overall effect on welfare is ambiguous. It depends on whether the reduction in the duplications of fixed costs offsets the greater price mark-up.

It can be seen that, as in many other topics in regulation theory, there is not an approach that can be considered, a priori, the best. As we will see in Part II, vertical separation is the norm in liberalized electricity markets. In telecommunications, the owners of fixed-line networks are commonly allowed to render mobile services. In
ports, the experience is mixed. In Australia, for example, regulated terminal operators can provide complementary services or be related to companies providing them (pilotage, towage, shipping, etc.). In Chile, there are limits to vertical integration. Up to 40% of a port concessionaire’s shares may be owned by “relevant players”, i.e., shipping companies or cargo owners accounting for more than 25% of the traffic at the concessioned terminal or more than 15% of the traffic at ports in the same region (Foxley and Mordones, 2000).

3.2 Access pricing

As seen before, fair access rules to naturally monopolistic facilities are necessary to promote competition within network industries, regardless of vertical structure. Laffont and Tirole (2000) claim that access charges reflect multiple objectives. They must “induce an efficient use of networks, encourage incumbents to invest while minimizing costs, generate an efficient amount of entry into infrastructure and services, and do all of this at a reasonable regulatory cost”12. Furthermore, according to Valletti and Estache (1998), the failure to design these policies is one of the main reasons why potential gains from restructuring utilities are not realized or shared between the users and the owners of these facilities.

3.2.1 Access pricing methodologies

In principle, pricing access poses the same difficulties for regulators as pricing final goods produced in monopolistic situations. If access is priced according to marginal costs, the fixed costs incurred in its provision have to be recovered through government transfers. If a surcharge is added to recover fixed costs, the regulator faces the dilemma between optimal (Ramsey) prices, and fair, subsidy-free and sustainable ones (see section 2.2.5). In this case, access to monopolistic facilities is seen just as any of the products sold by the incumbent. However, in the context of access pricing, Ramsey pricing will require price discrimination when the same product is sold to firms facing different price elasticities.

Even in a context of vertical separation, the problem of access pricing is complex due to the large consequences of not doing it adequately. For example, if access is priced in a way such as to allow the upstream monopolist to obtain a mark-up, the final effect in retail prices would be further magnified, leading to the double marginalization problem. This problem occurs because access charges constitute a cost for downstream operators, upon which they will charge their own mark-ups. In this way, part of the downstream mark-up will thus be charged upon the upstream mark-up, increasing retail prices even further (Motta, 2004).

---

On the contrary, if access is priced too low, excessive and inefficient entry may occur (arguably, subject to physical limitations). In this case, the waste of resources due to the duplication of fixed costs discussed by Vickers (1995) and described in the previous section would be produced. Even more, low access charges may discourage incumbents from maintaining and upgrading the infrastructure. The problem becomes even more intricate under vertical integration. As discussed in the previous section, the monopolist has in this case incentives to recover the foregone benefits by foreclosing entry in related markets (where it has to compete with other operators). High access charges erect barriers to entry and maintain the incumbents’ market power in the potentially competitive markets. Moreover, given the information asymmetry between the regulator and the regulated firm, this may lead to a practice called “price-squeeze”: the incumbent charges higher-than-optimum access charges and uses the extra revenues to undercut competitors’ final prices. In this way, competitors are “squeezed” between expensive costs and reduced retailing prices.

A first approach to price access is by using the “Efficient Component Pricing Rule” or ECPR, proposed by Baumol and Sidak (1994). This approach reflects the following rationale: if final prices are already fixed, the price of access has no effect on allocative efficiency. Therefore, if final products are homogeneous and the market is contestable, the access charge should equal the difference between the final price and the marginal cost of producing the good. Alternatively, the rule says that the access charge should be equal to the direct cost of providing access plus the opportunity cost of providing access.

It is important to explain the rationale of the ECPR in more detail. According to Laffont and Tirole (2000), the premise is that in absence of government transfers, the monopolist must recover its fixed costs by offering some goods or services above marginal cost. Therefore, setting access charges any differently than what would be done under Ramsey pricing will introduce distortions by requiring the incumbent to recover fixed costs in some other way. ECPR pricing reflects this view. This rule is neutral for the incumbent’s revenue, i.e., every good or service sold contributes with the recovery of fixed costs, even if it is not provided by the incumbent. Furthermore, the use of this rule secures a level playing field: entry is profitable only for those firms that are more efficient than the incumbent in the downstream activity. In this respect, the rule sends the right signal to new entrants.

This methodology has been repeatedly criticized (Tye, 1994; Economides and White, 1998). First, since the ECPR assumes that final prices are fixed, its critics claim it provides no mechanism for forcing retail prices to competitive levels. Secondly, if the incumbent is earning supra-normal profits, the use of the rule guarantees these rents. Valletti and Estache (1998), however, argue that this observation is inappropriate because ECPR assumes that final prices are optimally set. A third criticism is that if an entrant is effectively more efficient than the incumbent, the latter will disappear from the downstream segment, for which the regulation of the final price is irrelevant.
Other criticisms are that under the ECPR approach the entrants’ full costs on the competitive market are likely to exceed the incumbent’s incremental price; that the use of the rule will facilitate “price-squeezing” and that it would force the entrants to contribute to the incumbent’s cost inefficiency.

3.2.2 Implementation problems

The first problem of implementing any access pricing rule is to calculate and allocate costs adequately. The costs implicitly supposed in economic theories differ from the costs kept in accounting books, mainly because economic and accounting costs serve different purposes. A further complication is that accounting costs are usually kept according to historic costs, while economic costs usually suppose replacement costs.

The costs incurred in providing access are of two kinds: incremental costs, defined as those directly related to the increase in production caused by the demand for access; and common costs, those incurred in the supply of a group of services that cannot be directly attributed to any one service, typically as a consequence of economies of scope (Fernandez-Baca, 2006). The problem is that allocation rules are arbitrary and, under monopolistic situations, common costs are so important that, to avoid anti-competitive situations, have to be allocated in the right proportions to the various services provided with the infrastructure.

The most straightforward way to allocate costs is under the Fully Distributed Costs (FDC) methodology. Under FDC, the costs are allocated mechanically among the different products following several criteria: output shares, revenue shares, direct costs involved, contribution margins, etc. This allocation method has the advantage to be well understood and easy to implement. However, even though it allows recovery of investments, it does not account the demand nor encourages cost minimization. In the presence of competition, FDC may induce an inefficient amount of entry.

Valletti and Estache (1998) sustain that the current dominant paradigm is the use of cost-based methodologies such as the Long-Run Incremental Cost (LRIC). Their rationale is to set access prices on the basis of an efficient cost benchmark rather than on the incumbent’s actual costs. The LRIC is calculated based on the cost of the currently most efficient technology derived from an engineering model, and on a forecast of the most likely use of the infrastructure. The main drawback of this methodology is that it requires too much information and projections that in practice give the regulator a large amount of discretion. Indeed, this methodology requires knowing the lowest current cost of the incumbents’ equipment, how intensively it will be used and also the rate of technological progress of the industry. The method is popular, however, because it promotes competition although, according to the cited authors, prevents incumbents from making money in the bottleneck, which may bar perverse consequences in the long-run.
A more usage-based approach to calculate access charges is through a Global Price-Cap (Laffont and Tirole, 2000). This regulation methodology consists of setting a price-cap in the form of a weighted average of the goods or services produced by the monopoly. Given that the monopolist possesses better information about the characteristics of the demand than the regulator, it is allowed to choose any combination of prices that better matches the needs of his customers, as long as the weighted average does not exceed the global price-cap. This approach has the advantage that when a cap is set properly, the regulated firm is induced to choose the optimal Ramsey prices, without the need for the regulator to know demand functions. Since, for the incumbent, providing access is not much different than providing any other good or service, the rationale is to treat access not differently and include access charges in the computation of the global price-cap.

The main criticism to this approach is that it allows the monopolist to engage in anti-competitive practices. In fact, this methodology allows the incumbent to set high prices in the monopolistic segments, where no entry is encouraged, and set artificially low prices in competitive segments, thus discouraging entry. For this reason, this approach might need to be used with price floors and caps (Boyer, 1997).

3.3 The Essential Facilities Doctrine

The third important topic necessary to understand how access policies are implemented is the Essential Facilities Doctrine (EFD). This is a principle used to promote competition in markets where an (tangible or intangible) asset necessary to compete is controlled by a firm with the ability and incentive to foreclose competitors. According to this doctrine, under certain circumstances the incumbent should be mandated to grant access to competitors to this asset (the “essential facility”) under “reasonable” conditions.

3.3.1 The EFD under Antitrust Law

The EFD was first articulated by the US Supreme Court in 1912, in the case United States v. Terminal Railroad Association. In this case (see Appendix 3 for details), the company Terminal Railroad Association had monopolized all ways to cross the Mississippi river, thus controlling all the freight that crossed the city of St. Louis. Instead of mandating the divestiture of this bottleneck, the US Supreme Court restored competition by granting access to competitors to the monopolized facilities.

Since then, the US Supreme Court had reached similar decisions in a series of cases. According to Pitofsky, Patterson and Hooks (2002), the most cited are: Associated
Press v. United States\textsuperscript{14}, Lorain Journal v. United States\textsuperscript{15}, and Otter Tail Power Co. v. United States\textsuperscript{16}. In these and other decisions the Court made clear that the EFD portrays a unilateral refusal to deal, potentially liable as a monopolization attempt. According to these and other authors (Bergman, 2001; Werden, 1987), the leading US essential facilities case is \textit{MCI Communications Corp. v. AT&T}\textsuperscript{17}. In that case, the Seventh Circuit Court established that there were four elements necessary to establish liability under the EFD:

a. Control of the essential facility by a monopolist;

b. A competitor's practical or reasonable inability to duplicate the essential facility;

c. The denial of the use of the facility to a competitor; and,

d. The feasibility of providing the facility.

Since the establishment of this precedent, Pitofsky, Patterson and Hooks (2002) claim that virtually every court in the US has used these criteria to analyze cases related to essential facilities claims. However, as a result of the practical difficulty to demonstrate that the facility controlled by the incumbent is essential to competition; American courts rarely impose liability under the EFD.

The situation is somewhat different under European competition law. Van den Bergh and Camesasca (2001) argue that essential facilities cases constitute a special kind of “refusals to deal” practices under European law. According to Bergman (2001), any exclusivity agreement that limits trade to a pair or group of firms is considered to have the “object or effect” to restrict competition, and constitutes a violation of the article 81 of the EC treaty; unless it is explicitly allowed. Furthermore, any unilateral selective refusal to deal may also infringe the same article if the purpose or effect is to establish a vertical agreement, i.e., the setting of prices and conditions under which retailers are allowed to sell their products. But if there is no exclusivity element present, European law does not interfere with a firm’s right to decide with whom to deal, unless it is involved in abusive practices forbidden in the article 82 of the EC treaty.

Bergman (2001) argues that the EFD is also related to the prohibition against excessive pricing stated in the mentioned article 82 of the EC treaty. However, in this case, the doctrine is not seen as a criterion to define when excessive pricing constitutes an abusive practice, but a tool to be used against a certain type of predatory action.

\textsuperscript{14} 326 U.S. 1 (1945)
\textsuperscript{15} 342 U.S. 143, 146-149,156 (1951)
\textsuperscript{16} 410 U.S. 366, 377-79 (1973)
\textsuperscript{17} 708 F.2d at 1132-33
Van den Bergh and Camesasca (2001) argue that, in contrast with US practice, the EFD in Europe has been overused. They argue that frequently the underlying analysis, starting with market definition, has been conducted in a “rather patchy manner, and this then resulted in an overtly summary appraisal of the essential character of the facility at hand.”\textsuperscript{18} They use as an example the Sea Containers v. Stena Sealink case\textsuperscript{19}, in which the later evolution of the market proved the European Commission wrong in defining a harbor facility as essential.

The use of the EFD in Europe substantially proliferated until the European Court of Justice ruled the case Oscar Bronner GmbH & Co. KG v. Mediaprint Zeitungs-und Zeitschriftenverlag GmbH & Co. KG et al\textsuperscript{20}. In that case, the Court held that refusing access does not constitute an abuse of a dominant position under the article 82 of the EC treaty. To arrive at such decision, the Court set a series of conditions for a facility to be deemed as essential (lately called the “Bronner test”), that since then have substantially limited the scope of the EFD under European law. These conditions are the following:

a. The facility is controlled by a monopolist;

b. The facility is considered essential because it is indispensable in order to compete in the market with the controller of the facility;

c. Access is denied or granted on unreasonable terms;

d. No legitimate business reason is given for objectively justifying the denied access (as to the feasibility of providing the facility); and,

e. A competitor is unable (practically or reasonably) to duplicate the essential facility.

Consequently with this restrictive interpretation, the exceptional circumstances required to apply the EFD only exist if a monopolist’s refusal to deal eliminates all competition in a downstream market; and if this input is indispensable for competitors to carry out their business.

It is worth noting that although the standard formulation of the doctrine supposes two vertically-related markets, this is not necessarily a requisite to apply it under American antitrust law. In fact, in the case Aspen Highlands Corp. v. Aspen Skiing Co.\textsuperscript{21} the Tenth Circuit Court of Appeals applied the EFD when a ski resort decided to stop its participation in a “multi-area ski ticket” that allowed customers to use several resorts at a discounted price. The Court established that this multi-area ticket was an

\textsuperscript{18} Van der Bergh and Camesasca (2001), p. 275
\textsuperscript{19} OJ L 15/8 (1993)
\textsuperscript{20} C-7/97
\textsuperscript{21} 738 F.2d 1509, 1520-21 (10th Cir. 1984)
essential facility to which the firm was denying access with the intent of driving competitors out of business and monopolize the market. The Court also stated explicitly that vertical integration is not essential to finding a violation of antitrust laws (Pitofsky, Patterson and Hooks, 2002).

This practice contrasts with European competition laws. Indeed, as argued by Bergman (2001), “the doctrine is only applicable in situations where a firm that is refused access to the facility is incapable of being active in another market, due to this refusal”22.

3.3.2 The EFD under regulation: the Australian National Access Regime

In the early nineties, the Australian Federal Government commissioned a study aimed to define a national competition policy. The report of this inquiry, known as the “Hilmer Report” recommended the establishment of a new legal regime under which firms could access specific essential facilities on reasonable terms. The new regime was implemented in 1995 and applies only to infrastructures considered “of national significance” (Maddock and Marshall, 1997).

The Australian access regime applies only to assets that had been declared “essential facilities” by the National Competition Council (NCC). The criteria used by the NCC to evaluate applications by potential entrants are the following (Productivity Commission, 2001):

a. That access to the service would promote competition in at least one market (whether or not in Australia) other than the market for the services;

b. That it would be uneconomical for anyone to develop another facility to provide the service;

c. That the facility is of “national significance”, regarding its size, importance to trade and commerce or the importance of the facility to the national economy;

d. That access can be provided without risks to human health and safety;

e. That access to the service is not already subject to an access regime; and,

f. That access to the service would not be contrary to the public interest.

If the facility satisfies these criteria, the NCC makes its recommendation to the respective Federal, State or Territory Minister, who decides whether or not to

---

22 Bergman (2001), p. 418
“declare” the facility based on NCC’s arguments. When an asset is declared to be an essential facility, access can be granted in two ways:

a. A contract, where the applicant and the incumbent agree on access price and conditions following a negotiation. Once an agreement is reached, the contract has to be registered by the Australian Consumer and Competition Commission (ACCC), who can deny registration if its terms harm the public interest. If the parties are unable or unwilling to agree, either party may apply to ACCC for dispute arbitration.

b. Mandate by the ACCC when a deal is not reached.

When registered, such contract has effect as if it were a decision of the ACCC. If a facility has been declared and an agreement has already been negotiated or arbitrated, further access seekers may apply to be granted access under the same conditions. Likewise, existing regimes can be certified by the ACCC if they are effective in granting access to third parties at reasonable terms. The telecommunications, electricity and natural gas regimes have been certified by the ACCC. Furthermore, the owner of any non-declared facility may submit to the ACCC a voluntary undertaking which outlines the conditions under which access would be granted to any third party. In these cases, the ACCC may accept or reject the undertaking based on criteria described before (Maddock and Marshall, 1997).

The criteria upon which the ACCC must base its determinations (both for registering a contract or enacting a mandate) are the following (Productivity Commission, 2001):

a. The legitimate business interests of the provider's investment in the facility;

b. The public interest, including the public interest in having competition in markets;

c. The interests of all persons having rights to use the service;

d. The direct costs of providing access to the service;

e. The value to the provider of extensions whose cost is borne by someone else;

f. The operational and technical requirements necessary for the safe and reliable operation of the facility; and,

g. The economically efficient operation of the facility.

Two problems have been identified (Pengilley, 1998). The first is the length of the process. Indeed, the legislation itself does not specify how long would it take to an applicant to go through the whole access system, with the possibility to appeal in many instances to the Australian Competition Tribunal. The second and most
important problem is that the legislation does not state what access pricing scheme will be used to compensate a facility owner. The encouragement or discouragement of private investments will then depend on the consistency upon which the ACCC exercises its discretionary powers to decide what pricing scheme will be used.

Some features of the Australian access regime are worth to be highlighted. First, the regime is general (as opposed to industry-specific) and only applies to facilities considered as of national importance, regardless if they are owned by federal, state or private parties. The second important aspect is that the regime is an administrative one. Court recourses are only allowed when the administrative process is exhausted. The third and most important characteristic is that rather than directly regulating the price of access, Australia has chosen to allow parties to negotiate privately. The government only intervenes when an agreement cannot be reached.

3.3.3 The EFD as criterion for access granting

The application of the EFD can be seen from an economic point of view as a trade-off between static and dynamic economic efficiency. On one hand, it is desirable to limit the incumbent’s property rights to avoid the loss of welfare caused by the reduction of competition in at least one market. On the other hand, this limitation on property rights also reduces the incumbent’s incentives to invest, innovate and improve the coverage of infrastructure. This effect should not be underestimated, especially in the case of scarce, expensive-to-build infrastructure, such as that needed in transport, energy or telecommunications industries. Given the seriousness of the undesired effects caused by the limitation of property rights, the misuse of the EFD carries large risks. For this reason, it is important to establish clear criteria to select the cases where it is correct to apply the doctrine (Laffont and Tirole, 1993).

The desirability of the application of the EFD seems to be clear in cases when the facility is a natural monopoly and its use constitutes an essential input to produce goods in a related market. As discussed in previous sections, the duplication of natural monopolies would waste valuable social resources and increase the production costs of the goods produced with these facilities. Furthermore, its essentiality would increase the loss of welfare by facilitating the extension of monopoly power to potentially competitive markets—especially in the case when the monopoly is regulated.

The case is much less clear however when the input is not produced within a natural monopoly or its use is not essential to produce further goods (Maddock and Marshall, 1997). Indeed, the feasibility of duplication makes the market more contestable, thus reducing the potential of monopoly pricing. The application of the EFD in this context may also be counter-productive, given that the elimination of duplication incentives is likely to perpetuate the monopoly. In other words, the potential short-term gains of sharing may not compensate the long-term effects of reducing the owner’s incentives.
to invest and innovate. Analogously, the arguments for limiting property rights are weaker when the output produced by a natural monopoly can be substituted, for example, in cases when transport is possible both by railways and highways. In this case, even in the presence of sub-additivity of costs, the possibility of a monopolist excising market power is limited by competition.

As important as knowing in which cases the application of the EFD is desirable, is to correctly assess price and conditions under which access should be granted. If the price is too high or conditions are too restrictive, less-than-optimum access will occur, producing economic rents and thus loss of welfare. On the other hand, if the access is priced below costs or its conditions make the incumbent bear some of the costs that should be paid by the entrants, excessive entry will occur, thus reducing or eliminating the incumbent’s incentives to invest in new infrastructure.

In this respect, an issue raised by several authors is the inability of courts to correctly assess in which cases the EFD should be applied and to supervise its correct application in cases that access requires continuous supervision (Maddock and Marshal, 1997; Van den Bergh and Camesasca, 2001). According to Pengilley (1998), courts in general do not have staff with the required business expertise to evaluate commercial decisions, not even in the US, where courts have probably the greatest experience in EFD. In effect, US courts have not usually mandated access unless it is under “nondiscriminatory” conditions (in cases where access has been previously granted) or when the issue has been delegated to a specialized regulatory agency.

As we can see, the EFD serves as criterion to evaluate when access should be mandated, but not to set prices and conditions of such access. Considering the poor performance of courts in this matter, and that the specialized knowledge required for this task is similar to that necessary to price goods produced in monopolistic situations; it seems logical to entrust this task to an economic regulator. Indeed, in an increasing number of countries, access to natural monopolies is regulated by specialized agencies.

The system devised in Australia has two main innovations. The first one, considering the magnitude of the potential negative effects of misusing the EFD, the legislation aims to reduce regulatory failures by limiting the application of the access regime only to infrastructure regarded as of national importance. Moreover, in its 2001 review of the National Access Regime, the Productivity Commission recommends to mandate access only if this measure will increase competition in a substantial fashion (Productivity Commission, 2001). This choice implicitly supposes that undesired anti-competitive behavior can be dealt using antitrust law.

The second innovation of the Australian regime is the preeminence given to direct negotiation between the incumbent and the entrant. This limits the participation of the regulator to cases in which the parties do not reach an agreement. According to the Productivity Commission (2001), the regime is not intended to replace commercial negotiations between facility owners and access seekers, but to enhance the incentives
for negotiation and provide a means of access on reasonable terms and conditions if
negotiations fail. This feature, aimed to reduce the typical information asymmetries
between regulators and regulated firms, is consistent with a theoretical development
known as the “Coase Theorem”. This theorem indicates that if property rights are well
defined and there are no transaction costs (or these are extremely low), negotiation
between the parties will lead to a better resource allocation than if there was
governmental intervention (Coase, 1960). In this context, the intervention of the
regulator is justifiable only if its own costs are lower than the transaction costs of a
negotiated agreement between the parties (Flor and Defilippi, 2003).

However, some authors are skeptical about the social efficiency of negotiated
agreements. Maddock and Marshall (1997), for example, argue that only two
outcomes are possible: access pricing according to the ECPR rule (under which the
incumbent maintains its monopoly rents) and collusion between both parties to share
the rents. These authors, however, do not analyze cases where other access pricing
methods are used and ignore that the existence of economic profits will attract further
entrants until this rent disappears. Furthermore, the authors implicitly suppose that the
information asymmetry between the parties and the ACCC is such that this agency will
fail to recognize potential threats to public interest.
Appendix 3: United States v. Terminal Railroad Association

“In 1889, the notorious financier Jay Gould organized a coalition to acquire railroad facilities in and around St. Louis, Missouri. The antitrust suit that resulted from Gould’s escapade (…) involved three different means of crossing the Mississippi when the government finally sued the combination. After Gould had obtained control of each crossing, his acquisitive urge was still well short of its goal. At this dominant regional and transcontinental railroad junction, twenty-four independent lines terminated—half on the bluffs forming the St. Louis side of the Mississippi, and half on the plains stretching away from East St. Louis, Illinois, on the opposite bank. Gould’s group, which included only fourteen of the twenty-four lines, acquired all of the railroad facilities of both cities: terminals and yards, and tunnels and tracks leading from the high bluffs on the Missouri side of the Mississippi down to the river crossing below.

In short, the acquisition gave Gould complete control of the facilities necessary to load or unload freight traffic or passengers anywhere within the area of St. Louis or East St. Louis, let alone carry anything or anyone across the Mississippi. Given that the assets under Gould’s control were absolutely indispensable to the railroads of the region, and considering the importance of the railroad to both passenger and freight transportation in that era, it is difficult to imagine an amalgamation today that could achieve a similar chokehold. Perhaps one might imagine the unification under common control of the highways, bridges, railroad facilities, airports, and city streets of St. Louis and East St. Louis.

The specific results of the combination’s power were predictable: The combination was able to impose premium rates on traffic moving within and through the St. Louis area, constrained with respect to the latter by the presence of a railroad bridge at Memphis, Tennessee, roughly 285 miles to the south. These rates were imposed in the form of supplemental charges called “arbitraries.” The term suggests the likely attitude of the parties most obviously aggrieved by the situation—namely, the railroads relying on those facilities that were not included within Gould’s ownership group.

The federal government brought suit in 1905, seeking, under sections 1 and 2 of the Sherman Act, to dissolve the Association and restore independent competition among the various entities united by Gould. But the Supreme Court, in 1912, found merit in the defendant’s argument that the consolidation of terminal facilities within this enormous transportation complex would permit more efficient coordination of railroad operations. Accordingly, the Court held that dissolution would not be required unless the parties could not agree on a remedy short of divestiture. This remedy was to require the Association to admit any railroad to ownership on the same terms and conditions as the railroads already allied with Gould. Moreover, railroads that wished to use the Association’s facilities without becoming owners would have to be charged usage fees that would “place every such [railroad] company upon as nearly an equal plane . . . as that occupied by the [member] companies.” The Court gave no further guidance on the principles by which such rates could be calculated.

Thus, the competition that had existed before Gould’s consolidation of the numerous independent terminal companies and other facilities operators could have been restored by a decree of divestiture. Rather than rekindle the competition extinguished by Gould, however, the Court permitted the entry of a decree that required regulation of (1) the terms and conditions of ownership in the monopoly established by the consolidation and (2) the relationship between the rates and terms of usage applied to owners and those applied to non-owner users of the monopoly facility.”
4. Privatization and regulation

The following section describes the main characteristics of the reforms carried out by a large number of governments during the last three decades, and the new role of the private sector in providing public services. The second section presents various assessments of the consequences privatization brought to the countries that implemented it. We will see that despite the obvious failures of some processes, privatization generated important gains to most countries that implemented it. We will also see evidence that sound regulation is required to promote investment and avoid private agents obtaining economic rents from operating infrastructures. The last section presents the findings of several studies that have analyzed the effects of privatization in the port industry. These studies do not present conclusive evidence that private management deliver higher levels of efficiency, although much of the research leans towards this direction.

4.1 Private sector involvement in infrastructure industries

Transport, telecommunications and energy infrastructure play a vital role in economic growth, since they provide the basis on which economic activities take place. It has been argued that the inadequacy of physical infrastructure constitutes the key characteristic defining developing and transition economies (Dutz, Ordover and Willig, 2000).

Some characteristics of infrastructure-based industries (the presence of naturally monopolistic segments and externalities) make the provision of their services difficult for societies to organize. In effect, the fear that the owners of facilities may abuse their market power by providing a sub-optimal quantity of services was such that in many cases the ownership and operation of infrastructure within these industries was exclusively limited to governments.
Unfortunately, the provision of infrastructure-based services by the government was not exempt of problems. The public nature of the utilities’ management often made pricing and investment decisions to be guided by political and not economic reasons, causing gruesome inefficiencies and hindering economic development, especially for poor countries. The seriousness of the problem became evident in the last decades of the 20th century, when the globalization of world markets made inefficiency a burden too expensive to bear.

In the early eighties, the UK government embarked on a series of reforms aimed at changing the role of the public sector in the economy. Although these reforms were not exempt of criticisms (the most acute ones were those stating that the reforms were guided by pure ideological reasons), a number of developing countries started to imitate them very soon. These reforms share three common elements (Kessides, 2004):

a. **Liberalization of the economy.** The reforms typically liberalized the country’s economy, cutting tariffs to imports, removing barriers and eliminating exclusive rights to enter markets.

b. **Changing the roles of the public and private sectors.** The reform’s rationale assigned the private sector the role of creating wealth and generating employment whilst the public sector regulates and promotes competition. As a consequence, many public enterprises were privatized and governments embarked in the promotion of public-private partnerships. Private investment was promoted in areas once reserved for the government, from infrastructure to social services.

c. **Implementation of pro-competitive policies.** New antitrust agencies were created and old ones were empowered to guarantee competition. Considering the special characteristics of transport, telecom and energy, however, the promotion of competition in these industries required that some previously-integrated activities were unbundled, setting new rules aimed at creating markets where previously was none, and the creation of regulating agencies in charge of overseeing the remaining naturally monopolistic segments in the industries.

The reasons behind the reforms, however, were different for developed and developing countries. In the former ones, besides ideology, the shift toward a bigger role for private sector was motivated by failures of previous reforms aimed at introducing business-like practices (i.e. commercialization) in public companies but fell short of ownership change; plus the short-term fiscal attraction of selling state-owned assets (Kikeri and Nellis, 2004). In developing ones, the main reasons argued were the following (Alexander and Estache, 1999):

a. Governments were increasingly convinced that their own resources are insufficient to meet investment demands.
b. Even when a government may have sufficient funds for investing in infrastructure, other demands—mainly for social programs—were leading governments to seek new sources of funding.

c. A belief that private operators will produce greater efficiency than can be achieved by the public sector.

Kikeri and Kolo (2005) show that between 1990 and 2003, 120 developing countries carried out 7,860 transactions involving the divestiture of state assets or some form of public-private partnership; generating US$ 410 billion in proceeds. Nearly US$ 200 billion correspond to infrastructure projects. During the same period, additional US$ 350 billion worth of greenfield projects were developed through public-private partnerships. Around half of these transactions occurred in Latin America. Europe and Central Asia accounted for 26%.

Transactions in infrastructure sectors have been, however, concentrated in telecommunications and power, amounting to 50% and 36% of the total occurred between 1990 and 2003. During the same period, 65 developing countries embarked in projects involving some form of private participation in telecommunications, and 72 in electricity or natural gas.

Compared to the large number of privatizations, there have been few nationalizations of privatized companies. These occurred when heavily indebted companies failed or in times of crisis or due to political change. The most prominent examples of re-nationalization of privatized companies are Rail Track in the UK (due to bankruptcy), Air New Zealand, water and sanitation concessions in Bolivia as well as postal and water concessions in Argentina.

4.2 Privatization assessments and the need for better regulation

In the early eighties, when privatizations started in the UK, there was neither great theoretical justification nor hard evidence that the performance problems of state-owned companies could be altered by changes in ownership. But later, as privatizations processes started in other countries, the number of assessments of these policies grew.

Kikeri and Nellis (2004) surveyed more than 100 papers assessing privatizations. They focus on performance and compare productivity, profitability, output changes, investments and capacity utilization before and after the sale. Most of these studies conclude that privatization improves performance and profitability for new owners, as well as efficiency, output and investment. They also indicate that in terms of welfare, privatizations increased the resources available in the economy and produced net
gains for stakeholders (consumers, workers, investors, competitors and the government). However, only in rare occasions all of them benefited at the same time. In most cases, the status of winner or loser was a consequence of the transaction’s structure and the degree of institutional development of the economy. As we will see later, regulation plays an important role in the distribution of these gains.

Megginson and Netter (2001) reviewed cross-cutting studies that evaluate the impact of privatization over firm performance. These studies (Megginson, Nash and van Randerborgh, 1994; Boubakri and Cosset, 1998; D’Souza, Nash and Megginson, 2000) show that profitability increased from 8.6% in average before privatization to 12.6. They also show that efficiency rose from 96.9% in the year of privatization to 123.3% afterwards and that output per worker increased in between 79% and 86% of firms. The main factors driving these improvements were better incentives, more adequate management structure and improved corporate governance.

Moreover, a study of more that 200 enterprises privatized in Mexico by La Porta and Lopez de Silanes (1997) concludes that these companies went from highly unprofitable to profitable, closing the performance gap with the control group of similar private sector firms. Output rose 53.4%, sales per employee doubled and profitability increased 24%. The authors infer that improvements were due to productivity gains resulting from better incentives and management associated with private ownership.

Several studies also conclude that privatization improved efficiency and profitability of enterprises in Brazil. Pinheiro (1996) analyzed the performance of 50 Brazilian companies before and after privatization and concluded that the process significantly improved performance, especially when it involved a change of control rather than a sale of a minority stake. A paper by Macedo (2000) showed that a large steel mill, which had been unprofitable before privatization, started to obtain profits afterwards, boosting investments dramatically.

Privatization has also improved performance in infrastructure sectors. According to a survey by Harris (2003), well designed schemes for private participation (those that align economic incentives with public policy goals) produce substantial increases in overall welfare. For example, private participation led to overall domestic welfare benefits in Guinea and Argentina, amounting to US$1.4 billion in the case of Buenos Aires.

Moreover, the evidence produced by Harris suggests that in many developing countries the private sector does as well (or even better), than the public sector at expanding the service. In Gabon, for example, where the same private operator runs water and electricity services, targets for increasing coverage for both services have been met or exceeded. The privatization of electricity led to near universal coverage in Lima, Peru; and in Chile, the largest coverage increase has occurred in low income areas.
Results have been particularly impressive in telecommunications, especially where competitive regimes have been established. In fact, evidence suggests that expansion has been far slower where competition has not occurred; but even in that case, private monopolies have expanded more quickly than public ones. In Uganda, for example, the entry of private mobile companies led to a major increase in the number of connections, which largely surpassed those of the fixed-line incumbent. According to the author, the private sector’s technical and managerial competence, combined with more sustainable pricing policies and better financial discipline provide more resources for investing in the infrastructure’s expansion.

A study by Galal et al. (1994) estimates the consequences of privatizing 12 companies in Chile, Malaysia, Mexico and the UK. The study compares the performance of these companies before and after privatization, analyzing consumer welfare, efficiency and investment. The authors conclude that divestiture substantially improved economic welfare in 11 cases, which equal, on average, 26% of the firm’s pre-divestiture sales. According to the authors, this result was due to increased competition, adequate regulation, more rational pricing and investment policies and productivity improvements.

Mckenzie and Mookherjee (2002) evaluated the distributive impact of privatization in four Latin America countries (Argentina, Mexico, Bolivia and Nicaragua). They analyze the effects of this policy on customers and workers, based on household and employment surveys. They find that there is no clear pattern concerning price increases, with prices actually going down in about half of the cases. And even when prices went up, their effects were outweighed by increases in access to basic services that occurred in the lower half of the income distribution. They also found that the quality of service improved significantly, in some cases because it was mandated by the government as part of the conditions for sale of the public company. For example, the privatization of electricity in Bolivia was accompanied by rules that established a system for measuring quality, setting out dates by which privatized companies had to comply with the indicators and determined penalties in the event of failure.

Clarke and Wallsten (2002) analyzed the performance of infrastructure firms in delivering services to poor and rural households. Using data from around the world, their results show the failure of state-owned monopolies to provide universal service everywhere, with the only exception of Eastern Europe. The evidence also suggests that privatizations have not harmed poor and rural consumers, and in many cases have improved their access to infrastructure services.

Andres, Foster and Guasch (2006) analyzed the impact of privatization on the performance of 116 electricity utilities in ten Latin American countries. They constructed a panel data of a wide range of indicators such as output, employment, productivity, efficiency, quality, coverage and prices. They evaluate three stages: before privatization, transition period (two years following divestiture) and after the process. Their results suggest that changes in ownership generate important improvements in key indicators, especially during the transition period. The most significant ones are in labor productivity (around 19% in the transition period and
5.5% afterwards), efficiency (5.5% during transition), and service quality (reductions on the duration and frequency of interruptions of 9.8% and 10.6%, respectively, during transition).

There are also cases of privatizations not achieving their policy goals, generating undesired effects or having been carried out in environments unlikely to improve performance. Boubakri and Cossett (1998), for example, studied 16 privatized African firms and found a significant increase in capital spending but insignificant changes in output, profitability or efficiency. Analyzing the privatization of the Central Electricity Generating Board in England and Wales, Newbery and Pollitt (1997) conclude that the process effectually produced substantial cost reductions, but the bulk of the financial gains were captured by the new shareholders, not consumers or government (the privatization of UK ports was severely criticized for similar reasons). Likewise, Brown (2002) discussed the failures of privatizing Dabhol and Hub river power plants in India and Pakistan, respectively, and that of Rio Light electricity distribution company in Brazil. The author points out that the lack of a proper regulatory and competitive framework is much to blame for these failures, concluding that this is a prerequisite for transferring infrastructure companies to the private sector.

Kikeri and Nellis (2004) claim that results tend to be the best when privatization is combined with proper competition policies and regulatory frameworks. Other assessments also conclude that good regulation is required to generate adequate levels of competition, increase investment levels and improve the distribution of the efficiency and welfare gains.

A study made by Chisari, Estache and Romero (1999) separates the effects of privatizations in Argentinean utilities with those of regulation itself. They find that privatization led to operational gains equivalent to 0.9% of GDP or 41% of the average expenditure on utility services. More importantly, effective regulation added gains worth 0.35% of GDP or 16% of the average expenditure on utility services. According to the authors, this improvement occurs because regulation acts as a mechanism to transfer rents from shareholders to consumers.

In a recent study by Gassner, Popov and Pushak (2007), the authors analyzed a panel of 302 utilities with private sector participation and 928 utilities run by different government levels in 71 developing countries. Their aim was to evaluate the impact of private management on firm performance in electricity and water distribution. They find evidence that private sector participation is associated with output increases in electricity, connection increases in water, and improvements in service quality and bill collection ratios in both sectors. However, they do not find conclusive evidence for a change in prices nor an increase in investment levels, which point to a lack of maintenance and expansion investment even when private management has generated gains in operational efficiency. These results call for better regulation policies.
To determine if and how regulation matters for policies promoting private investment in infrastructure, Andres, Guasch and Straub (2007) tested the impact of regulation on 1,000 concessions granted in Latin America from late 1980s to early 2000s. They tested the effect of regulation on aligning costs and tariffs, deterring opportunistic contract renegotiation and on productivity, quality of service, coverage and prices.

Assessing the impact of regulation on aligning costs with tariffs, they find that the better the quality of regulation the closer the alignment between financial results and costs of capital (which implies that monopolists’ profitability is similar to that of firms in competitive markets). Assessing the impact of regulation on deterring opportunistic contract renegotiation, they find that the setting of a regulatory framework prior to privatizations reduces the likelihood of contract renegotiation. And, assessing the impact of regulation on productivity, quality of service, coverage and prices; they find that, effectively, regulatory structure, framework and quality are important determinants of how regulated companies perform.

It is worth noting the dearth of literature analyzing privatization or regulation of transport infrastructure. The paper by Estache and Serebrisky (2004) is one of the few exceptions. The authors analyze the effects of liberalization and regulation of transport infrastructure claiming that for governments in developing countries, two of their most important objectives when privatizing ports, airports, railways and toll roads have been the reduction of their fiscal burden and the promotion of new investments in the sector. In developed countries, however, the decision to deregulate and restructure may have been more of an ideological choice and less of an urge to resolve pressing problems.

The authors claim that in terms of efficiency, the liberalization and regulation of transport infrastructure has brought significant gains, although their specific source has not been adequately documented. From a fiscal viewpoint, the short-term effects have generally been positive, but it is still too early to draw long-term conclusions. From the viewpoint of freight transport, the effects of the reforms are not clear. Since competition with the trucking industry has been quite strong, it is unlikely that the average users may have been penalized. Nevertheless, the situation of captive shippers (those who have no other choice but to use the privatized infrastructure) might have been worsened. As for quality, the evidence is mixed. Visible quality dimensions, such as punctuality or safety, seem to have improved significantly (at least in developing countries); but less visible ones, such as regard for the environment, may have improved but not as much as expected. Lastly, the authors argue that there is enough evidence to point out the need for an improvement in regulation in many countries, especially developing ones. One of the most important dimensions that require improvement is the access to key transport facilities.

Kessides (2004) supports this view. After analyzing reforms in infrastructure industries carried out in developing and developed countries, the author points out two important issues:
a. Designing and implementing stable and effective regulation remains one of the most critical tasks for policymakers in developing and transition economies. In these countries the most pressing issue is designing regulatory mechanisms for privatized companies in infrastructure industries.

b. To obtain the benefits of having reformed infrastructure industries, policies need to harmonize regulatory oversight of monopolistic activities with fostering competition. For the latter, it is critical to design policies that effectively allow firms to access key facilities (access policies). Otherwise, the relation between bottleneck components and competitive segments can create such severe distortions that results in terms of productivity and efficiency can be even worse than before the reforms.

In the specific case of the reforms carried out in the port industry, the author indicates the following regarding the need for better regulation:

“The primary objective of port policy is to support national development. Although some emphasis has recently been placed on port services that add value to products, the development objective is usually best served by securing cheap and fast movement of traffic through ports. To that end, the landlord port model introduces competition either in the market for the provision of port services (between or within terminals) or for the exclusive right to provide services where the market is too small to support multiple providers. This approach may require structural controls to secure or maintain an appropriately competitive framework or, where structural measures are insufficient, controls to prevent monopolistic exploitation or distortion.”

4.3 Privatization assessments in the port industry

In the port sector, evidence about the private sector achieving higher levels of efficiency than the government is not conclusive, although much of the research seems to point in that direction.

Cullinane and Song (2002) studied the theoretical support and practical validity of the claim that the transfer of ownership from public to private hands will lead to an improvement in economic efficiency, and financial and operational performance. The authors claim that in the case of UK ports, it is extremely difficult to conclude that ownership constitutes a significant factor in port performance and efficiency. Instead, factors such as geographical location and deregulation seem to have a greater influence. They conclude that privatization is only a partial cure for port’s economic problems and that, if implemented in isolation, it cannot deliver the desired results.

Cullinane, Song and Gray (2002) employed two different models to assess the efficiency of a sample of Asian container terminals. The authors find evidence that

---

privatization is related to the efficiency improvements observed in the studied terminals. Cullinane and Song (2003) also find evidence that efficiency levels in Korean container terminals increase with the degree of private property.

Two studies suggest that privatization brought overall efficiency gains to the Mexican port system. Estache, Gonzales and Trujillo (2002) used panel data from 11 independent Port Administrations. Their findings concluded that the decentralization and privatization of Mexican ports produced positive effects to practically every Port Administration; and that the efficiency gains reached between 6% and 8% per annum between 1996 and 1999 for the whole industry. Moreover, the findings of Estache, Tovar de la Fe and Trujillo (2004) conclude that the Mexican reform introduced incentives to port operators to increase capacity and adopt new technologies, which constituted the main contributors of efficiency gains.

Trujillo and Serebrisky (2004) assessed the privatization and deregulation of Argentinean ports. According to the authors, the process transformed Puerto Nuevo de Buenos Aires from the most expensive port in Latin America into the cheapest one (stevedoring costs dropped by 8% per container, 11% for grain and 22% for general cargo). Cargo volumes grew significantly (containerized traffic grew by 78% between 1996 and 1999), and tonnage handled per worker increased from 900 per year in 1991 to 5,417 in 2001; allowing a reduction in the average duration of ship’s stay by about around 79% between 1995 and 2002. The privatization process had the additional advantage of eliminating all cross subsidies and reducing fiscal burden.

Tongzon and Heng (2005) analyzed a sample of selected container terminals around the world. They find that increasing private sector participation in the port industry contributes to increase efficiency, but full privatization does not. Based of their findings, the authors claim that the best policy is to limit private participation within the “landowner and operator” functions, with the public sector performing regulatory functions.

Other studies, however, do not find higher levels of efficiency in private port operations. Liu (1995) compared the efficiency of publicly and privately operated ports using data from 28 ports in the UK. His findings failed to show that ownership has a significant effect on efficiency. Analyzing 36 European container terminals and four Asian ones, Noteboom, Coeck and Van Der Broeck (2000) also failed to find a clear relationship between ownership and terminal efficiency.

Baird (2003) claims that port privatization offers port users and the economy as a whole many advantages, but only when carried out properly and for the right reasons. Moreover, Baird (2000) argues that the sale of port land and the transfer of operational and regulatory functions may be counter-productive. In fact, even though private sector participation may increase efficiency levels, an almost total dependence on private operators to provide infrastructure may result in delayed investments. Therefore, public sector support, particularly in regard to provision of certain elements of port infrastructure, is still likely to be necessary.
5. Port economics and the need for an access regime

The aim of this chapter is to argue why, in countries which possess naturally monopolistic terminals, privatization (in the form of concessions, leases, divestitures or greenfield projects), could be counterproductive unless access policies are implemented. It starts by describing port characteristics, services and markets and the complexities behind their organization and forms of competition. It continues discussing the drivers of the reforms carried out during the last quarter century and presents the experiences of three developing countries during their implementation. The chapter concludes with a discussion of the relationship between competition and access in naturally monopolistic terminals and the dilemmas regulators face when formulating such policies.

5.1 Port economics

5.1.1 Economic characteristics of ports and terminals

Ports possess a series of economic characteristics that make them difficult to regulate. Their first distinctive characteristic is that they are multi-service entities; i.e., different services for different markets are produced within a single port: infrastructure services, cargo handling services, pilotage, towage, mooring, etc. (Kessides, 2004). In the case of cargo handling, each type of cargo constitutes a different market: containers, grains, oil, coal, mineral ore, etc.

Moreover, ports not only produce private goods, such as the services mentioned before, but also public goods. Public goods are those which possess the characteristics of non-rivalry (their consumption by one agent does not preclude their consumption by a second one) and non-exclusion (it is impossible to exclude a consumer from their consumption), and produce positive externalities when they are used (Baird, 2004). However, since private companies cannot profit from them, they have no incentives to
produce them. Port assets considered public goods are breakwaters and navigational aids, among others (Haralam bides, 1997). Other public goods produced by ports are trade enhancement, second order increases in production volumes and collateral increases in trade-related services (World Bank, 2007). However, ports may also produce negative externalities, such as water pollution and the loss of natural marine habitats caused by vessel traffic, congestion caused by the vehicles transporting cargo to and from the port, and others such as landscape degradation and the loss of shores.

An important characteristic of the port industry is that services from more than one market are required to complete the transport chain (to allow cargo and passengers to be embarked or disembarked). For example, for cargo to be transferred between ship and land, ships carrying it need pilotage, towage and mooring, and cargo itself needs handling services. If one of these services is not provided, the others become useless. As we will see, this is an important feature for ports with monopolistic markets, since the monopolization of one market would allow the monopolist the control of the whole logistics chain (Flor and Defilippi, 2003).

A fourth distinctive characteristic of the industry is that the degree of competition largely depends on factors that are generally beyond the control of the port authority or the terminal’s management, such as the port’s location relative to maritime routes, the size and development patterns of the regional economy (which determines demand) and the quality and coverage of the overall transport network (which determines the degree of inter-port competition). The capacity to develop new ports or terminals is also limited because suitable shores are usually hard to find.

A fifth important characteristic of ports is the presence of economies of scale and scope. In fact, a study by Heaver (1975) suggests that economies of scale in port operations amount to 0.8; i.e., that a 10% increase in traffic would only raise total costs by about 8% (average costs would diminish). Economies of scope emerge from the fact that ports are multi-product entities, which implies that the same inputs can be used to provide different services. For example, once a channel has been dredged, it can be used by both tankers and container ships at lower costs than in situations where two separate channels are needed.

Tovar, Jara-Diaz and Trujillo (2004) reviewed several studies measuring these economies in ports and reported that not only infrastructure but also cargo handling services present increasing returns to scale. Other studies also suggest the presence of economies of scope between cargo types for infrastructure and cargo handling services. In a previous study, the same authors found both economies of scale and scope in Las Palmas port, in Spain (Tovar, Jara-Diaz and Trujillo, 2003).

A sixth economic characteristic of ports is asset indivisibility. Indeed, port terminals cannot be expanded in a continuous way but in discrete amounts of relative large minimum sizes (De Rus, Campos and Nombela, 2003). In fact, in the case of container terminals, the investment needed to develop a new berth can be as high as
The fact that the minimal size of terminal investments is relatively large, constitutes an entry barrier, since it is harder for newcomers to raise large amounts of capital. Moreover, the difficulty to find suitable shores to develop new ports (and the slow pace at which new terminals are developed) constitute barriers for existing companies to enter specific markets. Likewise, the combination of economies of scale and scope also constitutes an entry barrier, since smaller and less diversified entrants would face higher costs than the incumbent.

5.1.2 Port markets and services

Ports are conformed by water and land areas upon which infrastructure and superstructure is developed. The infrastructure of a port is conformed by the facilities that allow ships to enter and leave the port (channels, locks, navigation aids), load/unload cargo (berths, wharfs, yards), and for cargo to be transferred to and from the port (connections to roads, rails and waterways). The superstructure of a port consists of equipment (cranes, conveyor belts, elevators and other quay and yard-handling equipment) and fixed assets built on the infrastructure (fuel tanks, warehouses, buildings) (Estache and De Rus, 2000). Both infrastructure and superstructure are used to provide port services.

Table 5.1 presents a list of assets that constitute the infrastructure and superstructure of a port.

---

25 There are several definitions of what is regarded as port superstructure. The one used here is used by Estache and De Rus (2000)
Table 5.1: Port infrastructure and superstructure

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Superstructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channels</td>
<td>Cranes and other container handling equipment</td>
</tr>
<tr>
<td>Locks</td>
<td>Conveyors belts</td>
</tr>
<tr>
<td>Navigation aids</td>
<td>Elevators</td>
</tr>
<tr>
<td>Berths</td>
<td>Pipelines</td>
</tr>
<tr>
<td>Wharfs</td>
<td>Fuel tanks</td>
</tr>
<tr>
<td>Yards</td>
<td>Warehouses</td>
</tr>
<tr>
<td>Road, rail and water connections</td>
<td>Office buildings</td>
</tr>
</tbody>
</table>


It is worth mentioning that the basic port infrastructure (channels, navigational aids, breakwaters) is typically provided by a public body such as a port authority or maritime administration. But under a Public-Private Partnership (PPP), the terminal’s infrastructure can be provided by either the port authority or the terminal operator, depending on the PPP agreement. Under a Lease, the port authority provides the infrastructure and the terminal operator the superstructure. Under a Concession or a Greenfield Project, the operator also builds most of the terminal’s infrastructure (World Bank, 2007). This topic will be discussed in detail in section 5.2.1.

Figure 5.1 shows port’s logistics chain and the services required to complete it.

There are many markets within a port. The most important service carried out within a port is stevedoring/cargo handling, which consists on the transfer of goods between the berth and the ship26. Each type of cargo is considered a different market, since many of the assets required to handle them are different. The main cargo handling markets are: containers, coal, mineral ore, grain (wheat, maize and soy, among

---

26 Since in previous time cargo handling and stevedoring were performed by different crews, they were differentiated. The movement of goods from the berth to the side of the ship (and vice-versa) was considered cargo handling, whilst the movement from the side of the ship to storage inside the vessel was considered stevedoring. Nowadays, both activities are provided by single firms and are considered synonyms (Trujillo and Nombela, 1999).
others), oil, liquefied gases, break-bulk and semi-commoditized products such as rolling cargo (ro/ro), forestry, steel products, fruit and fruit juices, etc. (Fearnley’s, 2004).

It is important to note that the amount of cargo handled in large ports, such as Rotterdam, Long Beach or Buenos Aires, justify investing in specialized terminals (Stopford, 1997). But in small local ports such as those which usually constitute natural monopolies (for example, Ilo, Paita or Matarani in Peru and Buenaventura in Colombia), different types of cargo are handled within single multi-product terminals.

A second group of port services is related to the processing of these cargoes through the port: storage, cargo handling and related services (stuffing and striping, bagging and packaging, etc). A third group of port services is related to serve ship’s needs during their port call. Table 5.2 shows the main port markets.

<table>
<thead>
<tr>
<th>Table 5.2: Main port markets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Services to the Cargo</strong></td>
</tr>
<tr>
<td>Shipping</td>
</tr>
<tr>
<td>Stevedoring/cargo handling</td>
</tr>
<tr>
<td>Storage</td>
</tr>
<tr>
<td>Transshipment</td>
</tr>
<tr>
<td>Stuffing and striping</td>
</tr>
<tr>
<td>Bagging and packaging</td>
</tr>
<tr>
<td>Consignment and delivery</td>
</tr>
</tbody>
</table>

Port services are provided under various organizational models in which the public and the private sector play different roles. According to Baird (1999), there are three dimensions of port management that can be used to characterize the extension of privatization within a single port: regulation, land ownership and management and operations. In a pure public port, all of these tasks are carried out by the public sector. Under an organizational model that the author calls PRIVATE/I, port operations are carried out by private companies but the public sector still owns and manages port land and performs regulatory functions. Under PRIVATE/II, the private sector not only undertakes port operations but also own and manages port land. Under PRIVATE/III, the private sector performs the same functions as in the previous case but also performs regulatory functions. These options can be seen in table 5.3.

<table>
<thead>
<tr>
<th>Table 5.3: Port management models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model</strong></td>
</tr>
<tr>
<td>PUBLIC</td>
</tr>
<tr>
<td>PRIVATE/I</td>
</tr>
<tr>
<td>PRIVATE/II</td>
</tr>
<tr>
<td>PRIVATE/III</td>
</tr>
</tbody>
</table>

Source: Baird (1999)

According to the author, in 1999 88% of the top 100 container ports of the world operated according to the PRIVATE/I option.
It is worth noting that these models are stylized and port services are rarely organized in such straightforward way. In some PUBLIC or PRIVATE/I ports, such as those in Cyprus, for example, some port services have been liberalized but the port authority still supplies some services directly. Likewise, the fact that some port services are provided by the public sector does not imply that they are provided by a single entity. For example, in the case of Gothenburg port, in Sweden, VTS services are provided by the maritime authority, while locks are provided by the municipality and hinterland connections are provided by the public road or rail administrations.

It is also worth noting that even though there is a clear trend towards reducing the degree of government intervention in the industry, the results of the European Commission’s Port Policy Consultation 2006-2007 indicate that the debate on the role of the public sector still continues (European Commission, 2007). Most stakeholders, with the exception of some in the UK, support the idea of public financing for sea access to ports, but opinions are divided on financial support for the construction of quay walls. In the case of services related to port safety (pilotage and towage), the debate continues on whether these should be provided as public services or viewed as regular commercial activities whose characteristics are regulated by the government but whose tariffs are determined by market forces.

5.1.3 Port competition

As said before, ports are motionless, expensive-to-build infrastructures, where many services are provided and different markets develop and interrelate in complex ways. Three concepts need to be grasped when analyzing port competition: that ports may be very different from each other, that port competition occurs in overlapping hinterlands and that port competition occurs in different ways.

Ports are very different from each other

Since many of the characteristics of a port depend on geographical and economic features, there are not two ports alike. Differences are not only physical but, most importantly, on the markets they participate.

Common-user ports tend to compete in most cargo markets, while private ones tend to specialize in few commodities (ADB, 2000). Large ports tend to have enough traffic to justify the construction of terminals specialized in a single cargo, while small ones may consist of just one terminal with few berths handling different types of cargo.

According to Stopford (1997), ports can be classified according to their size and importance for the local economy. There are four main types of ports:
a. Small local ports. These are ports consisting of single terminals handling a mixture of containers, dry-bulk and break-bulk cargo. These terminals usually possess only basic facilities such as general-purpose berths and storage areas. Their typical users are small vessels (feeders and short sea shipping) and local shippers. In small islands or developing countries where long distances are combined with relatively small regional economies, ports like these may become natural monopolies.

An example of this type of port is Matarani, in southern Peru. The port is conformed by just one multi-purpose terminal with three berths (see figure 5.2). In 2006, it handled less than 2 million tonnes, mostly break bulk (soy, fishmeal), dry bulk (coal, maize, fertilizers) and sulfuric acid. Containerized cargo accounted for less than 15,000 TEU. The port is owned by the Peruvian Government but it was concessioned to an important business conglomerate in 1999 (Alcazar and Lovaton, 2005).

![Figure 5.2: Matarani Port Terminal (Peru)](image)

It is important to notice that, although small, this port is of crucial importance for the Peruvian economy. It is the second largest common user port and practically constitutes the only door for the foreign trade generated in southern part of the country (an area that generates one third of the Peru’s GDP and comprises 30% of its population). Moreover, long distances and poor transport infrastructure impede shippers located in this area to use alternative facilities, for which they constitute a captive demand. The closest common user port (Ilo), handles one tenth of Matarani’s cargo and is even less connected to
international maritime routes (it receives, on average, less than one ship a week) (Flor and Defilippi, 2003). This situation grants monopolistic power to Matarani’s private operator, for which tariffs, access terms and service quality have to be regulated by the government.

b. Large local ports. These ports are mainly local but their demand is large enough to allow the construction of specialized terminals. However, some cargo is typically handled using general-purpose berths. One example of this type of port is Algeciras, located close to the Strait of Gibraltar, in southern Spain. The port has specialized terminals for containers but manages passengers, rolling cargo and dry bulk through general purpose berths (see figure 5.3). Typically of this type of ports, Algeciras also has facilities to accommodate the local industrial fishing fleet (www.puertoalgeciras.org).

Algeciras handles more than 3 million TEU and 25 million tonnes of non-containerized cargo and more each year. It faces competition from other ports in all the cargoes it handles. Although most of its non-containerized cargo traffic is handled for the local industry (coal for the local power station and steel factory and petrol products for the local refinery and petrochemical complex), the port also serves as an important distribution hub for containers destined to European market. In this case, the port not only faces competition from ports located in the surrounding areas but from all others in the European South: Valencia, Barcelona (Spain), Marseille (France), Genoa, La Spezia, Gioia Tauro (Italy), etc.

![Figure 5.3: Port of Algeciras (Spain)](source: www.puertoalgeciras.org)
c. **Large regional ports.** These are ports that handle significant level of long-haul traffic and require large investments in specialized terminals and facilities for specific goods. Some of these ports are Zeebrugge (Belgium), Santos (Brazil) and Durban (South Africa.)

Figure 5.4 shows the layout of the port of Zeebrugge. The port has several specialized terminals and the facilities to serve the kind of large ships used in long-haul bulk transport. It has an outer port and an inner port. The outer port has two specialized terminals for containers and one for liquid bulk, dry bulk and rolling cargo, respectively. The inner port is an area of logistics and distribution activities, with several terminals and warehouses that handle specialized cargo: fruit and fruit juices, tires, cars and spare parts, among others. The port has rail and road connections and is also served by barge. In 2006, the port handled almost 40 million TM, including 1.6 million TEU (www.zeebruggeport.be).

![Figure 5.4: Port of Zeebrugge (Belgium)](image)

Figure 5.4: Port of Zeebrugge (Belgium)\(^{28}\)

d. **Regional distribution centers:** These are the world’s largest ports, such as Rotterdam, Antwerp, Hong Kong and Singapore. These ports also act as hubs for long-haul ships. They have several specialized terminals and facilities to distribute cargo using several other transport modes (railways, road, and inland navigation).

The case of the port of Rotterdam is illustrative. It is Europe’s main port. It has 2,000 ha of infrastructure extended in an area of more than 10,000 ha, including 74 km of quay length (see figure 5.5). Cargo can be transported to/from the port by truck, train, inland shipping, short sea shipping, air or pipelines. The port contains more than 60 terminals: 11 for containers, in excess of 20 for bulk cargo, three for juice, two for fruit, and 17 multi-purpose ones. Its annual throughput amounts to more than 400 million tonnes, including 10 million TEU. The port is also the heart of a large petrochemical complex, comprising five oil refineries, 44 chemical and petrochemical

\(^{28}\) Source: www.zeebruggeport.be
companies, 19 storage and distribution terminals for oil and chemical products, four refineries for edible oils and fats and more than 1,500 km of pipelines (www.portofrotterdam.com).

Port competition occurs in overlapping hinterlands

As explained before, each kind of cargo represents a different market, since many of the assets required to handle them are different. Therefore, competition occurs in each market. Ports can face intense competition for certain types of cargo and a much more limited one for another. According to Haralambides, et. al, (2001) price elasticities are in general much lower for liquid and dry bulk cargoes than for containers, general cargo and Ro/Ro.

The port of Rotterdam illustrates this situation. It faces intense competition for the containerized cargo generated in the industrial areas of northern and central Europe. For northern European traffic, competition comes from the ports located in the “Hamburg-Le Havre range”, a group of ports located in the coast line between northern France and Germany. In the container market, competition among these ports is intense, as one can see from their elasticity figures shown in table 5.4.

<table>
<thead>
<tr>
<th>Port</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamburg</td>
<td>3.1</td>
</tr>
<tr>
<td>Bremen Ports</td>
<td>4.4</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>1.5</td>
</tr>
<tr>
<td>Antwerp</td>
<td>4.1</td>
</tr>
<tr>
<td>Le Havre</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Source: Haralambides, Verbeke and Musso (2001)

---

29 Source: www.portofrotterdam.com
30 Ro/Ro is a shortening of the term "Roll on/Roll off." A method of ocean cargo service using a vessel with ramps that allow wheeled vehicles to be loaded and discharged without cranes (World Bank, 2001).
For the traffic generated in central Europe, additional competition comes from ports located in the Mediterranean. Principally those in Spain, France and Italy (see figure 5.6). In the case of dry bulk, its low value-added per unit of weight does not allow its transport over long distances, for which competition mainly comes from ports located in the immediate surroundings (Amsterdam and Antwerp). Liquid bulk is mostly imported to be consumed at Rotterdam’s petrochemical complex, for which it faces very limited competition.

![Figure 5.6: Main European ports](source: www.portofrotterdam.com)

Port competition occurs when the influence areas or hinterlands of two or more ports overlap. The extension of a port’s hinterland depends on many factors, although the most important is the cost of transporting cargo between the port and its point of consumption or production. In turn, transport costs are largely influenced by geographical features, political factors and the quality and extent of the transport infrastructure (De Rus, Campos and Nombela, 2003).

Two factors allow the container market to be very competitive: that containers are standardized, i.e., operators can handle them without further specialization; and that they do not necessarily contain cargo that will be used or consumed in the port’s vicinity. The hinterland of Pusan port, for example, includes not only South Korea but also parts of northern China and even Japan, areas typically served by feeder services connecting to Pusan (Yap and Lam, 2006).

---

31 Source: www.portofrotterdam.com
Factors such as the relation between value-added and weight and time-sensitivity also play an important role in determining the extension of a port’s hinterland. Manufactured and semi-manufactured goods, for example, have higher value-added levels per weight unit than raw materials, which can be transported over longer distances. Thus, the location of new manufacturing areas may extend a port’s influence area and allow inter-port competition where previously was none. However, this competition is limited by another common feature of manufactured goods, which is time-sensitivity. Indeed, intense competition in high-end industries such as apparels and electronics is reducing the life-cycle of these products, therefore increasing the need to reduce the time between manufacturing and delivery. This is the reason why many goods manufactured in northern Asia and destined to the US east coast are unloaded in west coast ports. Shippers prefer to pay extra handling costs and have them delivered earlier by train than waiting three to eight days and have them delivered (more cheaply) directly to east coast ports using the Panama Canal. According to ACP (2006), 61% of the traffic between northern Asia and the US east coast route uses west coast ports.

The fact that competition is limited to overlapping hinterlands may lead to situations where ports face both a captive and a competitive demand for the same cargo at the same time. The abovementioned port of Matarani illustrates this situation. As mentioned before, shippers located in southern Peru constitute a captive demand for the port’s only terminal. However, its influence also reaches the production areas of neighboring Bolivia, as shown in figure 5.7. But unlike southern Peruvian shippers, Bolivian ones also have the choice of using the Chilean ports of Arica and Iquique. Therefore, even though Matarani’s terminal operator has a monopoly over southern Peruvian cargo, it has to compete for the Bolivian one (Gamarra, 2006). Similar situations may arise in ports around the world that handle both local (captive) and transshipment (competitive) cargo.
It is important to note that the boundaries of what constitutes a captive hinterland are being reduced due to the increased efficiency of transport operations. Both train and truck operations have become more efficient and greatly reduced inland transport costs (Estache and Serebrisky, 2004). Improved custom procedures have also increased rivalry among ports located in neighboring countries.

The existence of extensive hinterlands and monopolies over captive customers is more likely to occur in sparsely-populated areas with relatively limited economic activity and long coastlines. Although this situation is more likely to arise in the developing countries, it also occurs in developed ones such as Australia. Indeed, a number of Australian ports face limited competition and are thus under monitoring or regulation of government authorities. The ACCC, under its Container Stevedoring Monitoring Program, monitors prices, costs and profits of container terminal operators at the ports of Adelaide, Brisbane, Burnie, Fremantle, Melbourne and Sydney (ACCC, 2008). The natural monopoly problem is especially severe in the case of small islands such as Réunion Island (a French possession in the Indian Ocean), which generates cargo to support only one container terminal (World Bank, 2007).

On the contrary, rich countries as the US and the European ones enjoy intense port competition. In the US, there is intense competition within ports located in the Gulf and in the east and west coasts and among them (Fleming, 1989). The relatively small extension of continental Europe, the existence of many navigable waterways and first-class transport infrastructure allow competition to occur between ports located in the North Sea and the Mediterranean (Foschi, 2003). Competition for transshipment cargo
is also very intense in ports located in southeast Asia, such as Hong Kong, Singapore, and Pusan, among others (Slack and Wang, 2002).

Port competition occurs in different ways

Given the large number of markets related to port activities, it is not surprising that competition takes place at different levels. Ports compete among them, terminals compete with other terminals within the same port, and service providers compete to supply port services to vessels and cargo owners within terminals. Moreover, when alternatives exist, competition can also occur with other transport modes. Port competition can be defined as: inter-port, intra-port, intra-terminal and inter-modal (World Bank, 2007).

a. **Inter-port competition.** Occurs when two or more ports compete for the same trades, like Canadian and US ports in both coasts; between Singapore and Hong Kong and between other ports in the region; and between ports that compete inter-regionally, such as Japan and Australia (Van Niekerk, 2005).

The containerization of general cargo has produced an important form of port competition that is worth commenting. Unlike most other cargoes, containers are transported by ships that make several calls under a fixed itinerary. This allows shippers to reach not directly-connected ports through successive transshipments. This practice has created a new market, since ports compete to attract transshipment traffic (Foschi and Cazzaniga, 2002). The most successful ones are those located where east-west routes intersect with north-south ones, such as Hong Kong and Singapore in Southeast Asia, Algeciras, Gioia Tauro and Marsaxlokk in the Mediterranean, and Balboa in Panama. Ports with large shares of transshipment traffic are called “hubs”. Figure 5.8 shows the location of the main hubs along the east-west route.

![Figure 5.8: Hub Ports in the East-West Route (World Bank, 2007)](image-url)
b. *Intra-port competition.* It occurs when two or more different terminal operators within the same port compete for the same markets. As one can see, users of large ports enjoy the benefits of this type of competition in addition to the one generated by inter-port competition. However, for users located in areas where the latter is weak, intra-port competition produces the additional benefit of avoiding rent extraction by the closest terminal.

According to De Langen and Pallis (2006), intra-port competition also promotes innovation and specialization and is more relevant when users do not have a strong bargaining power. Examples of medium-sized ports where terminal operators compete head-to-head are Buenos Aires (Argentina), Laem Chabang (Thailand) and Bremerhaven (Germany).

c. *Intra-terminal competition.* It occurs when companies compete to provide the same services within the same terminal. Both cargo and vessel services can be competitively provided within a terminal. In many ports, pilotage, towage and cargo storage are provided by several competing companies. In Puerto Limon (Costa Rica), for example, stevedoring is provided by several competing companies (Kent and Fox, 2004). Moreover, one of the aims of the proposed EU’s Directive for market access to port services was that mooring could be provided by ship owner’s own land-based staff (Comisión Europea, 2001).

d. *Inter-modal competition.* This is a form of competition that arises when cargo can be transported using alternative transport modes, such as railroads and trucks; or a combination of them with maritime transport. This situation allows competition to occur between ports situated at long distances.

The US provides a good example of inter-modal competition. As mentioned before, goods manufactured in northern Asia and destined to areas close to the US east coast can be unloaded in nearby ports (New York, Charleston or Savannah) or unloaded in west coast ports (Los Angeles, Long Beach, Oakland, Seattle/Tacoma and Portland) and distributed via rail to their final destination. Another choice is to use gulf ports (Houston or Mobile) and transported from there by train or truck (ACP, 2006). This produces inter-modal competition among several logistics chains, of which ports and terminals are just a node.

### 5.2 Port reform and trends

Most countries’ port sectors have been reformed since the 1990s. Although reform has taken different forms depending on each country’s particular conditions, a general trend toward limiting the public sector’s role and allowing ports to be managed under a business orientation has emerged.
Much has been written since the mid-1990s on the forces that are driving this reform: Haralambides (1995), Haralambides, Ma and Veenstra (1997), Juhel (2001), Oster and Strong (2000), Slack (2001) Van der Veer (2001), World Bank (2007), among others. It seems to be a consensus that globalization and the increasing containerization of general cargo are the main forces behind port reform. Other observed trends that are shaping today’s port industry, such as privatization or the increasing concentration of the container handling market, seem to be their most evident consequences.

5.2.1 Forces driving port reform

Historically, governments have had a large role in the provision and regulation of port infrastructure and services. Haralambides, Ma and Veenstra (1997) point out the following reasons for this policy:

a. Economic policy. As explained in section 5.1.1, the fact that ports produce public goods and externalities became one of the main justifications for government intervention, which was oriented at converting ports in “growth poles”. Moreover, since before containerization ports were large demanders of low-skilled labor, this argument was especially important in developing countries implementing inward-looking industrialization policies, where high port dues tariffs, long turnaround times and inefficient services constituted effective non-tariff barriers to foreign competition.

b. Natural monopolies. As seen in chapter 2, natural monopolies arise when technology imposes a cost function that makes it cheaper to produce a good or service with only one firm in the market. This situation is likely to occur when the total cost has a large fixed-cost component, as in ports, depending on the size of the demand they face.

In previous times, when port demand was constrained by high transport costs and the low value added by the commodities traded, ports were referred to as the classic example of natural monopoly. Direct provision of port services by the government was thought to be the best policy to deal with this type of market failure.

c. Financing. Ports require investing large quantities of capital which will have long recovery periods. In times where financial markets were not as deep and well developed as today, governments where the only ones capable of financing port investments, since public policies usually extend beyond considerations of short-term profitability and toward the maximization of long-term social profitability. Moreover, the port sector was thought to be
reluctant to invest in ports because institutional and regulatory frameworks could not guarantee positive financial returns.

d. **Military reasons.** Ports are strategically important in times of war and are especially vulnerable to attacks from the sea. Moreover, many commercial ports are located close to a country’s borders or nearby naval facilities –like Puerto Cabello (Venezuela), Murmansk (Russia) or La Spezia (Italy), for example. Reasons like this were often argued to justify a large role for government in ports.

Times have however changed and also technology and markets. Moreover, poor public sector performance and new ideas about its role in development have set the stage for reform in many infrastructure-based industries.

**Globalization**

In ports, one of the main forces driving reform is globalization. According to Haralambides (1995), globalization has brought greater mobility of goods, services and production factors due to three main factors:

a. Cultural changes that occurred as a result of telecommunications, mass media, advertising, the abolition of national barriers and other factors; have led to a convergence of consumption patterns and the creation of larger international markets.

b. The idea shared by many governments that economic integration will lead to a more efficient allocation of resources and thus stimulate growth and economic development.

c. The technological changes that have increased the speed and efficiency of transport and lowered the costs of communications.

The main consequences of globalization, along with the trend toward trade liberalization and the reduction of tariffs, have been to weaken the link between manufacturing and the location of production factors, the expansion of internal markets for goods and services and the relocation of manufacturing facilities to developing countries. Sources of raw materials and the markets where final goods are sold have become global.

For Kumar and Hoffman (2002), globalization means that trade is growing faster than the world’s GDP; and that this trade is not only in finished goods and services, but increasingly in components and services that are used within production processes of global scale (particularly, intra-firm trade).
According to the World Bank (2007), globalization has allowed four manufacturing processes to occur: vertical specialization, focused manufacturing, expansion of logistics' geographical scale and increased sourcing alternatives. The implementation of these processes, in turn, reinforces the increasing interrelation of national economies that characterizes globalization.

a. **Vertical specialization.** Firms have been increasingly concentrating on their core competences and subcontracting out non-core activities to contractors. Companies within the Italian apparel industry, for example, have been performing key activities such as design, cutting, finishing and quality control, while contracting out the assembly of their products. The customization of many Asian-manufactured products to be distributed in Europe is also performed in distribution centers located in Spain, Belgium or the Netherlands.

b. **Focused manufacturing.** Many firms have been changing the way they organize the manufacturing of their products. Instead of having local plants manufacturing a broad range of products for domestic markets, they have been concentrating production capacity in few plants that manufacture few products for global markets. European car producers like Volkswagen have been implementing this strategy, which allows the company to maximize economies of scale through specialization. However, it implies that products have to be transported, on average, over longer distances; thus making manufactured goods more transport-intensive.

c. **Expansion of logistics' geographical scale.** As a consequence of globalization, companies have expanded the geographical scale of their sourcing and distribution operations. It means that companies such as Toyota, Airbus or Starbucks acquire inputs and sell their products all over the world.

d. **Increasing sourcing alternatives.** Globalization allows producers in distant parts of the world to compete with each other for the same raw materials, intermediate or final product markets. Chinese companies, for example, can buy mineral ore from Africa, Australia or South America and European consumers can buy wine from France, Chile, Argentina or New Zealand from their convenient store.

The economies of developing countries have the distinctive characteristic that a large share of their exports consist of raw materials and semi-manufactured goods with low value-added. In a scenario of global and very competitive markets, inefficiencies such as those produced by inadequate infrastructure and poorly managed port services may hinder economic growth. This new reality requires cheap and reliable logistics chains of which ports are important components. Thus the need of transforming ports into efficient providers of logistics services.
**Containerization**

The second relevant trend behind port reform is the increasing containerization of general cargo. According to Vigariè (1999), before the advent of the container all general cargo was transported as break bulk, i.e., in sacks, boxes or crates of different size and volume. Most vessels were multi-purpose, carrying general cargo above and bulk cargo below. Turnaround times were slow. The author claims that in an efficient port as Antwerp, gangs could handle as much as 300 tons in two seven-and-a-half hour shifts. As a result, ships spent up to 30% of their time or more in ports, a fact that destroyed any economies of scale that could be reached by larger vessels.

The most important characteristic of pre-container ports was that much labor was needed. The irregularity of ships’ departure and arrival required maintaining a casual workforce to complement the number of longshoremen employed in a regular basis. (Haralambides, 1995). This required an organization of the workforce that facilitated unions achieving a position of importance in many ports of the world. Through strikes and confrontations, unionized port workers secured better work conditions for themselves (including much higher salaries than in other industries) but significantly raised the costs of port operations.

The container was introduced in the US in the early 1950s, for transporting freight between New York and Gulf ports. In 1956, the first container ship (a converted tanker) started operations, and ten years later the first regular container service between New York and Rotterdam took place. The improvements in cargo handling that were brought about by containerization dramatically reduced ship’s time in port and allowed achieving economies of scale by building larger ships. Other technical advancements in propulsion and ship design contributed to this trend by allowing these vessels to sail faster.

Although shippers and ship-owners gained much from containerization, port workers were the main losers. Instead of several gangs of low-skilled workers, containers require mechanized equipment and smaller numbers of trained laborers, thus changing the nature of port operations from labor-intensive to capital-intensive. Despite the best efforts of union leaders, the number of port workers has dramatically decreased around the world; and with them, the hours lost to strikes and other forms of service disruptions. This factor has also contributed to organize container shipping as a regular service with scheduled port calls, allowing better income predictability and the financing of ever larger ships. As Haralambides (2007) points out: “Operational practices have been streamlined; the element of uncertainty in cargo flows largely removed; forward planning has been facilitated; port labor regularized; and customs procedures simplified. These developments took place under the firm understanding of governments and local authorities that ports, now, constitute the most important link (node) in the overall door-to-door supply chain and thus inefficiencies

---

32 General cargo is a general term that refers to all cargo that is not transported as liquid or dry bulk.
(bottlenecks) in the port sector can easily whither away all benefits derived from economies of scale and scope in transportation and logistics”.

According to Kumar and Hoffman (2002), around two-thirds of the world trade (measured by weight) is transported by sea. Of this quantity, 10% corresponds to general cargo. Containerization reaches approximately 60% of world’s general cargo trade. The remaining 40% is still transported as break bulk. Figure 5.9 illustrates these proportions.

According to UNCTAD (2007), by the beginning of 2007 there were 3,904 fully cellular containerships with a total capacity of 9.4 million TEUs, which represents 12.3% of the world’s total fleet (measured by dwt). Ship sizes continue to increase, with average carrying capacity per ship growing from 2,324 TEUs in January 2006 to 2,417 TEUs in January 2007.

It is worth noting that containerization requires ports to invest sizeable amounts of capital in mechanized equipment and modern port facilities that allow the reduced turnaround times that make investing in large ships profitable. Haralambides (1995) argues that developing countries initially responded to this necessity with skepticism. Their worries concerned the suitability of capital-intensive techniques in countries with abundant and inexpensive labor, their lack of financial resources and the fact that the vast majority of their exports were not containerizable. However, the success of countries such as South Korea, Taiwan and Chile in implementing export-led

---

Figure 5.9: World trade by type of transport service

According to UNCTAD (2007), by the beginning of 2007 there were 3,904 fully cellular containerships with a total capacity of 9.4 million TEUs, which represents 12.3% of the world’s total fleet (measured by dwt). Ship sizes continue to increase, with average carrying capacity per ship growing from 2,324 TEUs in January 2006 to 2,417 TEUs in January 2007.

It is worth noting that containerization requires ports to invest sizeable amounts of capital in mechanized equipment and modern port facilities that allow the reduced turnaround times that make investing in large ships profitable. Haralambides (1995) argues that developing countries initially responded to this necessity with skepticism. Their worries concerned the suitability of capital-intensive techniques in countries with abundant and inexpensive labor, their lack of financial resources and the fact that the vast majority of their exports were not containerizable. However, the success of countries such as South Korea, Taiwan and Chile in implementing export-led

---

Percentage of metric tons. Source: Kumar and Hoffman (2002). Data from unctad.
strategies seems to have convinced many governments that job losses in the port sector could be more than compensated in other export-oriented labor-intensive industries such as agriculture, aquaculture or textiles.

Moreover, a number of principally Asian countries located near trading routes saw the opportunities that containerization brought by enabling them to export transshipment services, as explained in section 5.1.3. Transshipment increases cargo traffic beyond the limitations imposed by their own local economy and allows ports to achieve significant economies of scale that benefit the country’s external trade. Indeed, as a result of transshipment, shippers in Hong Kong, Singapore or Panama face lower logistics costs than their peers in other countries.

Haralambides (1995) also argues that transshipment traffic allows the development of feeder service networks for the regional distribution of containers. This would enable the port’s home country to be profitably involved in shipping and other value-added activities that otherwise would be lost to competing ports.

Privatization

The third main trend behind port reform is the private provision of port services. As said before, port services have been typically provided by the public sector. But the substantial changes occurring in the industry and the investments needed to meet them require efficient port operations and market-oriented organizations. In recent years, many governments have acknowledged that the private sector can provide the capital required to invest in equipment and infrastructure and the efficient management that their ports require.

The term “privatization” loosely refers to different types of processes involving the transfer of state-owned assets to the private sector. Although privatization can take several forms, two are the main ones (World Bank, 2007):

a. Under *divestitures*, the government transfers ownership of the business to private companies.

b. Under *concessions*, the responsibility for operating and maintaining the privatized facilities is passed to a private company, but not its ownership. Their key characteristic is that the concessionaire has the right or the obligation to modify the infrastructure. This is the reason why they are granted for long terms (20-30 years).

Divestures are common in non-infrastructure industries, such as mining, manufacturing, oil, shipping, air transport, etc. They were also the most common way to privatize telecommunication and energy networks. In transport, however, the large majority of privatized facilities (airports, toll roads and ports) have been concessioned (Guislain and Kerf, 1995). A less used option is the granting of *management*
contracts, under which a private firm operates a public company, but the public sector is still responsible for investments in infrastructure expansion and other important matters such as labor downsizing.

In the particular case of the port sector, some terminals have been transferred to private operation through *lease contracts*, under which the private operators are not required or allowed to modify the infrastructure. These agreements usually have shorter terms than concession contracts. In addition to the privatization of existing facilities, some governments have promoted the development of new ones through private investment, called *greenfield projects*. These projects typically include explicit subsidies or some form of government guarantee. The associated commercial risks tend to be the responsibility of the private party, while other risks (exchange or political ones) are shared with the government.

Table 5.5 provides an overview of the extension of privatization in the port industry. It can be seen that almost 300 port projects with private participation have been implemented between 1990 and 2006, half of them through concessions and one third through greenfield projects. More than 100 of them were developed in Latin America and the Caribbean. It can also be seen that divestitures are not common in this industry.

Table 5.5: Port projects with private participation by type (1990-2006)

<table>
<thead>
<tr>
<th>Region</th>
<th>Concession</th>
<th>Divestiture</th>
<th>Greenfield project</th>
<th>Management and lease contract</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Asia and Pacific</td>
<td>30</td>
<td>5</td>
<td>41</td>
<td>3</td>
<td>79</td>
</tr>
<tr>
<td>Europe and Central Asia</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>Latin America and the Caribbean</td>
<td>70</td>
<td>4</td>
<td>24</td>
<td>10</td>
<td>108</td>
</tr>
<tr>
<td>Middle East and North Africa</td>
<td>6</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>South Asia</td>
<td>6</td>
<td>0</td>
<td>18</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>33</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>43</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>151</strong></td>
<td><strong>18</strong></td>
<td><strong>107</strong></td>
<td><strong>23</strong></td>
<td><strong>299</strong></td>
</tr>
</tbody>
</table>

Source: PPIAF (2007)

Table 5.6 shows that the private sector invested more than US$ 33 billion in port projects between 1990 and 2006. Greenfield projects amounted to US$ 16 billion while concessions to US$ 15.5 billion.
According to the Asian Development Bank (ADB, 2000), the following are the main trends among PPPs in ports:

a. A large number of countries have adopted the landlord model;

b. Private common-user ports are mainly found in the UK and New Zealand;

c. Private investment within public-owned ports has been directed mainly to container terminals and break bulk facilities dedicated to specialized cargo: steel, coal, wood products, etc.; and,

d. The majority of privately-owned facilities are individual terminals operated by the owners of the cargo.

The shift toward the private provision of port services has not been smooth, neither in developed nor developing countries. Although there are indications that private supply has increased the efficiency of port services (Micco and Perez, 2001; Tongzon and Heng, 2005), several problems remain.

In an economically integrated area like the EU, where decisions of one Member State may have consequences on the others, interesting public policy issues emerged as a consequence of privatization, regarding subsidies, investments and pricing of port infrastructure and services (Haralambides, 1997a and 1997b). The first is the rationale for continuing subsidizing basic infrastructure (dredging, breakwaters, road/rail/canal connections) in an environment where ports are no longer viewed as growth poles. Indeed, if port terminals are viewed as commercial undertakings, why should they not pay for their infrastructure with the dues they collect? On the other hand, it is difficult for one country to stop subsidizing their ports unless countries home to competing ports also stop doing it.
An additional problem is that subsidies can be subtle. Haralambides (1997) argues that port land is rarely liable to market forces and thus its pricing does not reflect the opportunity cost of using it for other ends. Therefore, port charges will not reflect their opportunity cost either, which amounts to a subsidy. The fact that ports belong to a transport network add another level of complexity to the problem, since port investments cannot reflect commercial decisions when the rest of the transport infrastructure in which ports belong is publicly funded or subsidized.

In less developed countries, privatization has brought dramatic increases in efficiency due to better management and investments (see section 5.2.2), but private operators will not invest in socially (economically) profitable projects unless they are profitable from a private point of view as well. Because of this issue and the externalities caused by port operations (congestion and other environmentally-related problems), governments will continue having an important role in the port industry, regardless of the extension of the privatization programs.

After two decades of port privatization, three main problems have been observed. The first is the way the process itself took place. In many countries privatization was implemented in a haphazard way, raising questions about its real necessity, costs and the fairness of their results. The UK port privatization, for example, was heavily criticized it because allowed employees to buy shares at prices well below market value and sell them later at a large profit and because it also transferred some regulatory functions to the private sector (Baird, 1999).

A second problem related to port privatization processes is the one argued by Estache, Ellis and Trujillo (2007), that a large group of port and transport projects seems to be in the hands of no more that twenty operators, sponsors, bankers and investors with a large capacity to invest, access to financial markets and strong financial support from their parent companies. While local investors may participate in specific niches, these major organizations set the standards and practices in transport project finance, thus reducing even further the potential for competition in port markets.

A final major issue that emerges with privatization is the distribution of its benefits between operators and users. In markets where effective competition takes place, terminals will compete away any monopolistic rent, thus benefiting users in the process. But as discussed in section 5.1.1 port markets tend to be oligopolistic and natural monopolies may also arise, for which some form of regulation is warranted. The aim of this research is precisely to contribute to the development of such a regulatory body in the port industry.

Market concentration

An important consequence of containerization is the emergence of global players. Indeed, the latest years have witnessed a wave of mergers and acquisitions in the
container terminal sector, which has resulted in fewer competitors handling a larger share of the world’s throughput.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Operator</th>
<th>2006</th>
<th>Share</th>
<th>2002</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HPH</td>
<td>60.9</td>
<td>13.80%</td>
<td>HPH</td>
<td>36.7</td>
</tr>
<tr>
<td>2</td>
<td>APM Terminals</td>
<td>52</td>
<td>11.80%</td>
<td>PSA</td>
<td>26.2</td>
</tr>
<tr>
<td>3</td>
<td>PSA</td>
<td>47.4</td>
<td>10.70%</td>
<td>APM Terminals</td>
<td>17.2</td>
</tr>
<tr>
<td>4</td>
<td>DP World</td>
<td>41.6</td>
<td>9.40%</td>
<td>P&amp;O Ports</td>
<td>12.8</td>
</tr>
<tr>
<td>5</td>
<td>Cosco</td>
<td>22</td>
<td>5.00%</td>
<td>Eurogate</td>
<td>9.5</td>
</tr>
<tr>
<td>6</td>
<td>Eurogate</td>
<td>11.7</td>
<td>2.70%</td>
<td>Cosco</td>
<td>4.7</td>
</tr>
<tr>
<td>7</td>
<td>Evergreen</td>
<td>9.4</td>
<td>2.10%</td>
<td>Evergreen</td>
<td>5.7</td>
</tr>
<tr>
<td>8</td>
<td>MSC</td>
<td>8.8</td>
<td>2.00%</td>
<td>DPA</td>
<td>5.3</td>
</tr>
<tr>
<td>9</td>
<td>SSA Marine</td>
<td>7.6</td>
<td>1.70%</td>
<td>SSA Marine</td>
<td>4.4</td>
</tr>
<tr>
<td>10</td>
<td>HHLA</td>
<td>6.6</td>
<td>1.50%</td>
<td>APL</td>
<td>4.3</td>
</tr>
<tr>
<td>11</td>
<td>APL</td>
<td>5.9</td>
<td>1.30%</td>
<td>HHLA</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>Hanjin</td>
<td>5.4</td>
<td>1.20%</td>
<td>Hanjin</td>
<td>3.7</td>
</tr>
<tr>
<td>13</td>
<td>OOCL</td>
<td>4.8</td>
<td>1.10%</td>
<td>MSC</td>
<td>2.2</td>
</tr>
<tr>
<td>14</td>
<td>Dragados</td>
<td>4.7</td>
<td>1.10%</td>
<td>NYK Line</td>
<td>3.5</td>
</tr>
<tr>
<td>15</td>
<td>CMA CGM</td>
<td>4.5</td>
<td>1.00%</td>
<td>OOCL</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>NYK Line</td>
<td>4.1</td>
<td>0.90%</td>
<td>CSXWT</td>
<td>2.7</td>
</tr>
<tr>
<td>17</td>
<td>MOL</td>
<td>3.3</td>
<td>0.80%</td>
<td>MOL</td>
<td>2.7</td>
</tr>
<tr>
<td>18</td>
<td>K Line</td>
<td>3.1</td>
<td>0.70%</td>
<td>Dragados</td>
<td>2.2</td>
</tr>
<tr>
<td>19</td>
<td>TCB</td>
<td>2.9</td>
<td>0.60%</td>
<td>K Line</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>ICTSI</td>
<td>2.2</td>
<td>0.50%</td>
<td>TCB</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>308.9</strong></td>
<td><strong>69.9%</strong></td>
<td><strong>154.6</strong></td>
<td><strong>56.0%</strong></td>
</tr>
</tbody>
</table>

Source: Drewry Shipping Consultants (2004 and 2007)

This trend can be seen in table 5.7, which shows the throughput of the 20 largest container terminal operators. In 2002, they handled 56% of the world’s throughput. In 2006 that figure had increased to 69.9%.

The consequences of this consolidation phenomenon on port competition are yet to be seen. On the one hand, consolidation reduces the number of players in the market, which may lead to less competition and higher prices for the shippers. On the other hand, the presence of economies of scale indicates larger operations tend to be more efficient, for which the consolidation of the industry may lead to a more intensive rivalry. So far, the second effect appears to be the largest, since intense competition is observed and there are no signs of higher prices for port services.
5.2.2 Port reform in developing countries

As mentioned before, a general trend in port reform is toward limiting the public sector’s role and allowing ports to be managed under a business orientation. According to Juhel (2001), the key topics addressed under port sector reform programs include:

a. The reformulation of national port system development strategies;

b. Reforms of legislative, institutional and procedural provisions for the planning and regulation of port systems;

c. The participation of private parties in the provision and management of port services; and,

d. The introduction of innovative financing and cost recovery schemes.

The following sections describe how port reform was implemented in four developing countries. It can be seen that in most cases reform resulted in a dramatic increase of operational efficiency, reduced prices and increased investments.

Argentina

Argentina has forty active ports, one under the jurisdiction of the federal government and the remaining ones under the provincial governments. The most important is Buenos Aires, which handles more than 40% of the country’s tonnage. Buenos Aires includes two areas: Puerto Nuevo and Dock Sud.

Argentina began reforming its port sector in 1990. Until this year, the system was operated under a complex institutional framework, its infrastructure was inadequate, tariffs were high and traffic was declining. Indeed, traffic declined 10% between 1970 and 1989 and tariffs raised 250% in real terms between 1980 and 1991. In addition, labor practices were severely restrictive and the system faced major corruption problems. Between 1990 and 1993, the system was reformed. With the exception of Buenos Aires’ Puerto Nuevo, all major ports (including the area called Dock Sud) were transferred to the provinces to be administered by Sociedades de Administración Portuaria (SAP), which operate under certain parameters. The provincial and municipal governments as well as users, operators and workers have a chair in the SAP’s board. The remaining ports were also transferred to the provinces, which could operate them directly, lease them or shut them down. In addition, stevedoring, pilotage and tug assist services were deregulated and operators were allowed to set their own prices. Restrictions on vessel registration and short-sea shipping were also released and restrictive labor practices were abolished (Trujillo and Serebrisky, 2004).
Free entry to the industry was established, which means that any agent can build and operate a port, subject to basic safety, environmental and custom regulations. It was also established that users would be protected against anti-competitive practices by general antitrust laws. Regulation remained in the areas of safety, security and environmental protection.

In 1994, Buenos Aires’ Puerto Nuevo was divided into six terminals and concessioned to the private sector on the basis of the highest annual payment to the government. Bidders were allowed to present offers for more than one terminal and winners would compete for containers and general cargo. Duration of the contracts was set according to their investment needs (between 18 and 25 years). Global players started to operate in the port. P&O Ports was awarded terminals 1 and 2, and APM Terminals acquired terminal 4 in 2001. Soon after Puerto Nuevo concessions were awarded, Dock Sud was concessioned by the Province of Buenos Aires. This operation was largely criticized, since the concession was granted without competitive bidding and payments to the Province were lower than what Puerto Nuevo’s operators had to disburse (although the required investments were higher). Moreover, Dock Sud concession was granted after Puerto Nuevo’s, which gave the new concessionaire an informational advantage over its competitors.

The reform of the Argentinean port system is generally considered a success. According to Hoffman (2001), it has achieved the desired goals of increased throughput, lower costs and an almost fourfold increase in productivity. Private investment reached about US$ 200 million in five years, the number of cranes increased from 3 to 22 and throughput capacity increased from 400,000 to 2.5 million TEU a year. Stevedoring prices fell 8% per container, 11% for grain and 22% for general cargo. Container terminal charges reduced from US$ 450 before the reform to approximately US$ 220 in 1999. It is important to mention that even though there is a price cap for the tariffs of some container-related services, market rates are between 30% and 10% below these marks, making them irrelevant. Productivity also increased dramatically. The number of handled tonnes per worker grew from 900 in 1991 to 5,400 in 2001. Average ship stay was reduced from 70 to 15 hours between 1995 and 2002 (Trujillo and Serebrisky, 2004).

Colombia

Most of Colombian cargo is handled by four ports: Buenaventura (in the Pacific coast), Cartagena, Barranquilla and Santa Marta (in the Atlantic Coast). Until the reform, all common-user ports in Colombia were centrally managed by Colpuertos, a monopolistic state-owned company. Private companies were allowed to own their own terminals but not to supply services to third parties.

Colpuertos management was characterized by high-priced low quality service, technological backwardness and corruption. Vessels had to wait more than 10 days to
enter the port, and Colpuertos had more than 11,000 employees. Despite high tariffs, the institution ran operation deficits of more than US$ 20 million (Gaviria, 1998).

Colombia began reforming its port industry with the issue of Law 01 of 1991. This law sought to liquidate Colpuertos and to separate regulatory from operational activities. It created the Superintendencia General de Puertos to regulate the system and awarded concessions to specially-created entities called “Sociedades Portuarias Regionales” (SPR). SPR’s are private-public partnerships where the public sector (national, state and municipal governments) retain ownership and participates with 30% of the shares. SPR are in charge of the infrastructure, own the superstructure and are not allowed to provide services directly. Concessionaires can set tariffs within regulated guidelines and coordinate with the government future investment plans. Port services are provided by competing operators which contract with the SPR to use the facilities. The new system also encourages private ports to provide services to third parties. Private ports do not have the restrictions imposed on SPR to provide stevedoring services, for example (Hoffman, 2001).

The process of concessioning Colombian ports was finalized in 1993. The concessions were awarded for 20 years to the highest lease offer. According to Gaviria (1998), uncertainty produced during the concessioning process regarding the terms of the concession contracts and the role of the Superintendencia discouraged experienced foreign operators.

Since their concession, the Atlantic ports of Cartagena, Barranquilla and Santa Marta have engaged in a fierce inter-port competition, which is exacerbated by the presence of two privately-owned container terminals in Cartagena. Traffic increased so dramatically in Cartagena and Buenaventura that their respective SPR were obliged to seek an extension of their concession contracts in order to carry on unforeseen but required investments in infrastructure (Nathan Associates, 2004).

The results of the reform were positive. In a few years, ship time in port was drastically reduced and productivity yields multiplied in all ports, as shown in table 5.8. Productivity increase was translated into price reductions of up to 50%. In the case of containers, cargo-handling rates fell from US$ 600 to US$ 150 in 2001. SPR’s equity grew on average 2.2 times between 1994 and 1998 (Hoffman, 2001).

<table>
<thead>
<tr>
<th>SPR</th>
<th>General Cargo</th>
<th>Solid Bulk</th>
<th>Containers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barranquilla</td>
<td>8.03</td>
<td>21.53</td>
<td>6.02</td>
</tr>
<tr>
<td>Buenaventura</td>
<td>18.70</td>
<td>47.41</td>
<td>18.70</td>
</tr>
<tr>
<td>Cartagena</td>
<td>7.97</td>
<td>n.a.</td>
<td>24.75</td>
</tr>
<tr>
<td>Santa Marta</td>
<td>26.98</td>
<td>24.72</td>
<td>8.50</td>
</tr>
<tr>
<td>Colpuertos 1990</td>
<td>6.85</td>
<td>7.04</td>
<td>5.05</td>
</tr>
</tbody>
</table>

Source: Kent and Fox (2004)

*34 This row presents the yields Colpuertos achieved in 1990.*
Table 5.9 presents performance indicators for the Port of Cartagena:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Containership waiting time</td>
<td>10 days</td>
<td>0 days</td>
</tr>
<tr>
<td>Containership turnaround time</td>
<td>72 hours</td>
<td>7 hours</td>
</tr>
<tr>
<td>Gross productivity/hour</td>
<td>7 moves/ship hour</td>
<td>52 moves/ship/hour</td>
</tr>
<tr>
<td>Berth occupancy</td>
<td>90 percent</td>
<td>50 percent</td>
</tr>
<tr>
<td>Bulk Cargo Productivity</td>
<td>500 ton/vessel/day</td>
<td>3,900-4,500 tons/vessel/day</td>
</tr>
<tr>
<td>Hours worked per day</td>
<td>16</td>
<td>24</td>
</tr>
<tr>
<td>Cargo dwell time</td>
<td>30+ days</td>
<td>2 days</td>
</tr>
<tr>
<td>Port costs</td>
<td>US$ 984 per move</td>
<td>US$ 222 per move</td>
</tr>
</tbody>
</table>

Source: Kent and Fox (2004)

Mexico

The Mexican port system is composed of 108 ports and terminals located in the country’s Pacific and Atlantic coasts. Most of the cargo is handled by 27 commercial, industrial and tourist ports and 10 terminals specialized in oil and mineral ore. The most important ports are Veracruz, Tampico and Altamira in the Atlantic coast; and Manzanillo and Lazaro Cardenas in the Pacific side. Manzanillo and Veracruz alone handle 75% of Mexican containerized cargo. Before its reform, the Mexican port system was managed by the state-owned agency Puertos Mexicanos, the only one authorized to build port infrastructure and provide port services. As in previously analyzed cases, the system was inefficient, overmanned and did not generate enough revenues to cover its costs (Paredes, 2007).

The reform of the Mexican port system started in 1993, when a new Ports Law was passed. The law created the Administraciones Portuarias Integrales (API); autonomous, self-financed agencies in charge of a port or several small terminals. The APIs are state-owned bodies whose boards include representatives of the state and municipal governments and private sector. APIs have rights over port assets and may grant them in concession to the private sector. Although exceptions apply, APIs are not allowed to supply port services directly. The federal government, through the Secretaría de Transporte (Mexico’s Ministry of Transport) supervises 16 APIs controlling the most important ports, while provincial governments are responsible for the oversight of 5 provincial ones. The Secretaría also acts as regulator in cases where competition is absent or not strong enough. This, however, has to be determined by the competition authority. Although port tariffs were liberalized, the tariffs charged by the APIs for the use of infrastructure are regulated using a price cap regime.

Private participation has been introduced in almost every port through concessions. Initially, competition authorities ruled that no operator could manage more than one concession in each coast, but this restriction was later modified. As a result, Hutchison Port Holdings, which has controlling interests in container terminals in Manzanillo and Veracruz, was able to form a joint venture to operate a similar terminal at Lazaro Cardenas. The markets for services such as pilotage and tug assist
were also deregulated and since then, entry is only subject to qualifications. Labor restrictions were also relaxed. Collective bargaining is only allowed at firm-level, thus allowing negotiation according to local conditions. Only Acapulco’s API, specialized in tourist services, has been privatized (Estache, Tovar and Trujillo, 2004).

The introduction of competition brought important investments to the port system. Between 1993 and 2003, annual handling capacity almost doubled, passenger traffic more than doubled and container traffic quadrupled. In Manzanillo, for example, average port time for container ships reduced from 1.4 days in 1991 to 0.6 days in 1994 (Hoffman, 2001). Moreover, between 1995 and 1998, cargo handling rates decreased between 34.5% and 21.7% for dry bulk and palletized goods, respectively; and 5.6% for container traffic.

Two studies suggest that the reform brought overall efficiency gains to the Mexican port system. Estache, Gonzales and Trujillo (2001) found that efficiency gains for the industry reached between 6% and 8% per annum between 1996 and 1999. Moreover, the findings of Estache, Tovar and Trujillo (2004) suggest that the main contributors of efficiency gains in the Mexican case were the adoption of new technologies and capacity increases resulting from investments made right after the reform began.

Peru

Peru has seven common-user ports, spread along a coast of 2,500 km. The most important is Callao, which handles 70% of the total and 90% of the containerized cargo mobilized common-user ports. Since 1970, Peruvian ports have been operated by Enapu, a state-owned company dependent from the Ministry of Transport. As with other state monopolies, Enapu is overmanned, its tariffs are among the highest in the region and the service is poor (Defilippi, 2004).

Since Peruvian ports constitute natural monopolies, a regulator, Ositran, was created in 1999. Ositran’s mission is to supervise and regulate not only ports, but airports, railways and highways as well.

Attempts to reform the port system started in the early 1990s, when restrictive labor schemes were abolished. Although the first port privatization committees were organized in 1992, the first concession was awarded in only 1999, when the port of Matarani was transferred to the private sector. The concession of the remaining state-run ports was scheduled for 2000, but political problems forced the suspension of the process and since then, only a greenfield project for a container terminal at Callao has been granted. The remaining ports are still operated by Enapu.

The only terminal at Matarani Port was concessioned for a period of 30 years, to the bidder offering the highest initial payment to the government. In addition, the concessionaire has to pay the government monthly 5% of its gross revenues. There are no restrictions to the services the concessionaire can provide within the port, but it has
to grant free access to other companies providing stevedoring, pilotage, and tug assist services. Uncertainties about Ositran’s functions and last-minute changes to the concession contract discouraged two of the three pre-qualified bidders from participating in the final auction for Matarani. As a result, the concession was awarded to the only remaining bidder, which offered a payment barely above the required minimum (Alcazar and Lovaton, 2005).

Unlike the ports that remain under state management, Matarani shows signs of increased efficiency. The investments made by the concessionaire, a professional management and an aggressive marketing campaign allowed increasing the port’s annual throughput by almost 50% between 1998 and 2003; as shown in figure 5.10 (Ositran, 2005).

![Figure 5.10: Annual Throughput in the Port of Matarani, 1998-2003](image)

During the same period, the number of workers in the port was roughly kept at the same level, for which labor productivity also increased by almost 50% . This trend is shown in figure 5.11.

After the first five years of the concession, regulated prices were subject to revision by Ositran. The aim of price revisions is to pass on the consumers the efficiency gains obtained by the concessionaire between revision periods. The analysis determined that the concessionaire had obtained productivity gains of around 4.16% a year, for which the price cap of the port’s regulated prices (wharfage, berthing, mooring, and

storage), had to decrease annually at that rate. Moreover, increased competition in related port services has also resulted in lower handling rates for users (Defilippi and Flor, 2008).

5.3 Access, competition and regulation in the port industry

In competitive markets, firms compete with each other for selling their goods or services to customers, offering lower prices or improving their quality. In a competitive environment, companies lack “market power”, which is the ability of firms to raise price above the competitive level for a sustained period (Buccirossi, 2008). There are, however, several reasons why markets may not operate properly. The first is agents in the market behaving in ways that foreclose competition. To deal with this problem, governments apply antitrust policies. The second is market structures that facilitate anti-competitive policies or impede competition to arise at all. To deal with this problem, governments implement regulatory policies.

Antitrust policies are typically enforced by a specialized agency, while regulatory policies are usually enforced by a sectorial regulator. It is worth mentioning that some countries have both an antitrust agency and a sectorial regulator, while some countries

---

have one or neither. In some countries, such Australia and Papua New Guinea, both functions are performed by the same authority.

Table 5.10 shows the typical role distribution between antitrust agency and sectoral regulator.

<table>
<thead>
<tr>
<th>Agency</th>
<th>Policy</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antitrust agency</td>
<td>Prevention of collusive practices</td>
<td>Behavior control</td>
</tr>
<tr>
<td></td>
<td>Prevention of exclusionary practices</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prevention of discriminatory practices</td>
<td></td>
</tr>
<tr>
<td>Antitrust agency/Regulator</td>
<td>Merger Control</td>
<td>Preventing excessive concentration harming competition</td>
</tr>
<tr>
<td>Regulator</td>
<td>Price Regulation</td>
<td>Mimic competitive outcomes</td>
</tr>
<tr>
<td></td>
<td>Access Regulation</td>
<td></td>
</tr>
</tbody>
</table>


As mentioned before, in the particular case of the port industry, diverse factors such as the magnitude of the investments required to build a new terminal and the scarcity of suitable shores result in markets served by only a handful or firms. This situation may reduce the degree of rivalry in the market and thus, facilitate anti-competitive behavior.

5.3.1 Antitrust policies in the port sector

Antitrust policies are tools governments use to prevent market agents behaving in a way that deviates from fair competition. Many potentially anti-competitive practices in the port sector can be prevented using antitrust policies, although dealing with others might require sectoral regulation. Anti-competitive practices can be categorized as collusive, exclusionary and discriminatory (Motta, 2004).

Collusive practices

Collusive practices are agreements among firms who should be competitors to reduce their output to agreed upon levels; or to sell their output at an agreed upon price. This is known as “cartelization”, which can be more or less difficult to occur depending on certain market conditions. The following ones facilitate cartelization (Ivaldi et al, 2003):

a. Reduced number of firms;

b. Possibility of monitoring each other’s behavior;
c. Existence of similar firms;

d. Homogeneous products;

e. Existence of substantial entry barriers;

f. Inexistence of substitutes; and,

g. Inelastic demand.

It can be seen that some of the characteristics of the port industry facilitate cartelization: port services are fairly homogeneous, the number of competitors tend to be reduced and there are substantial barriers to enter the market. Thus, cartelization is a likely behavior to occur in the port industry. However, unless special circumstances arise, cartels tend to be unsustainable in the long-run because the incentives to cheat are usually very large (Posner, 2001).

Exclusionary practices

Exclusionary practices are behaviors that attempt to exclude competitors from the market. Practices that only involve firms participating in the same market are called “horizontal”. Others that involve exclusion through discrimination of clients or suppliers are called “vertical” (Hovenkamp, 1999).

Two are the most common horizontal exclusionary practices. The first one consists of attempting to monopolize the market by foreclosing competitors access to inputs that are necessary to compete. This problem will be analyzed later with more detail.

The second common horizontal exclusionary practice is predatory pricing. It consists on a firm reducing prices below their costs with the intention of driving competitors out of the market. The idea is to raise prices to monopolistic levels once the competitors have left the market. Understandably, for this strategy to be profitable the value of losses incurred during the first period has to be lower than the profits obtained once the competitors have left the market. Nevertheless, for this to occur substantial barriers to enter the market must exist; otherwise, the monopolistic prices set during the second period might attract new competitors who would undermine the profitability of the strategy (Ivaldi et al, 2003). Considering that substantial entry barriers exist, predatory pricing is an anti-competitive practice that might occur in the port sector. This is what in fact occurred in Cartagena (Colombia), where Terminal Marítimo Muelles El Bosque, an integrated terminal operator providing all-in-one services, was accused of predatory behavior by its competitors in the pilotage and towage markets (World Bank, 2007).

The kind of vertical restriction that might occur in the industry is called “exclusive dealing”. It arises when a consumer agrees to only acquire services supplied by a
determined provider or, at least, not to do it from competitors (Buccirossi, 2008). Port terminals might try to extend their market power to other markets or to raise their rival’s costs by requiring their clients to deal exclusively with certain shipping companies or providers of complementary services. The fact that port services are complementary (vessels must be supplied by both piloting and tug assist to enter the port, for example) facilitates this behavior.

**Discriminatory practices**

These are practices that allow companies to obtain unfair benefits or to harm competitors by using their market power (Hovenkamp, 1999). Two are the ones that may occur in the port sector: price discrimination and tying.

Price discrimination arises when (i) identical products are sold at different prices under identical cost conditions; or (ii), identical products are sold at common prices under different cost conditions. For example, port terminals may attempt to charge different prices to captive and transshipment cargo even if the cost of producing services for both kinds of clients is the same. Another option is terminals using their market power to avoid involving in standard commercial practices, such as granting clients discounts for volume transactions or early payment.

Tying consists of “bundling” services to oblige consumers to purchase services that need with those not requested or required (Buccirossi, 2008). For example, terminal operators might condition the use of their infrastructure to the supply of other port services as well, such as warehousing or container stuffing/striping. This practice has the additional negative consequence of foreclosing competition in other markets.

**Merger control**

In addition, many antitrust agencies are in charge of imposing safeguards to prevent industry concentration harming consumers. The rationale behind merger control is that is preferable to avoid companies obtaining excessive market power than to avoid its abuse once it exists.

Mergers and acquisitions are useful to the economy because they take advantage of scale economies and synergy, and promote innovation. However, these operations also reduce the number of competitors in the market and have the potential of reducing rivalry among remaining firms, which may lead to the occurrence of anti-competitive practices (Whinston, 2006).

The potential negative consequences of concentration (excessive prices, supply not meeting the needs of consumers) are exacerbated in industries such as ports, where high entry barriers may dissuade competitors to enter the market. For these reasons, port terminal mergers are commonly subject to previous authorization or operators are prohibited to possess other companies’ stock above certain level (as the case of
Chile). In some countries, regulators are even empowered to mandate divestiture when concentration reaches excessive levels.

There are several remedies available to competition authorities to prevent the potential negative consequences of mergers (OECD, 2004):

a. *Merger prohibition.* Consists on preventing a merger;

b. *Partial divestiture.* Consists on allowing firms to proceed with a merger if they divest the assets or business units that may cause anti-competitive behavior;

c. *Conditional approval.* Consists on allowing a merger to take place subject to monitoring of its post-merger behavior. The regulator can later order a complete divestiture if anti-competitive behavior is observed.

5.3.2 Regulatory policies in the port sector

There are certain markets where competition is unlikely to develop or it is undesirable for technical, safety or security reasons. In these circumstances a monopoly will arise. Antitrust policies were designed to deal with anti-competitive behavior in markets where competition exists, but they are not adequate to deal with situation were competition does not exist. In these circumstances, economic regulation is needed.

As discussed earlier, port competition occurs in several ways. But it has also been said that the port industry has several economic characteristics that may generate oligopolistic or monopolistic markets. According to the definitions given in chapter 2, ports terminals may become natural monopolies when demand is insufficient to exhaust their economies of scale or scope. In these circumstances, it will be cheaper to supply the existing demand with just one provider, since competition from a second terminal would lead to higher prices for all users. Therefore, competition would not develop and the monopolist would be able to extract rents from shippers or carriers, harming them and hindering the development of the country’s international trade.

It is very important to underline that naturally monopolistic port terminals only arise in very particular circumstances: areas with relatively low economic activity where traffic is limited, inter-modal connections are bad or border crossings are problematical, such as developing countries or small islands. Although oligopolistic structures are common in the industry, most terminals do face some degree of competition. As explained before, competition can arise from other terminals located in the same port (intra-port competition), in a different one (inter-port competition), as well as with other logistics chains (inter-modal competition). Although antitrust suffices to deal with oligopolistic markets, economic regulation is necessary to deal with monopolistic ones.
Naturally monopolistic port terminals are located all over the world. For example, a survey by UNCTAD (2003) revealed that several African ports and their respective terminals have monopoly status. Van Niekerk (2005) also argues that many South African port terminals constitute effective monopolies. These are also found, for example, in Eastern Europe, where Odessa port plant (Ukraine) has a monopoly status in the shipment of ammonia and nitrogen fertilizers (ICPS, 2007). In Asia, Aden port terminal (Yemen) handles all the country’s containerized traffic. According to the World Bank (2007), Réunion Island generates traffic for only one container terminal, for which the one located at its main port can also be considered a natural monopoly. Likewise, examples of naturally monopolistic terminals are found in Latin America. Tamayo, Paredes and Flor (1999) analyzed the main common-user Peruvian port terminals and concluded that all of them have a monopoly over one type of cargo or another. Moreover, Kent and Hochstein (1998), argue that Colombia, Nicaragua and Costa Rica have at least a dominant terminal in a near or total monopoly position. Even in a developed country like Australia, port terminals are subject to monitoring or regulation because of their monopolistic position (SAIIR, 2002).

As discussed previously, the existence of monopolies warrants economic regulation, which aims to mimic prices and conditions of service to those that would have prevailed if markets were competitive. Economic regulation is based on two policies (Defilippi and Flor, 2008):

a. Regulation of prices and quality of service (price regulation); and,

b. Regulation of how firms access the facilities they need to compete in the market (access regulation).

As Defilippi and Flor (2008) argue, price regulation is more costly. It requires the regulator to possess large amounts of information that is hard to collect, and to use unbiased decision-making processes that are difficult to implement, especially in a context of weak institutions. A more preferable policy is then to try introducing competition in as many markets as possible and only regulate the remaining ones. Depending on the particular situation, it may be possible for governments to induce some form of competition in their ports. For container terminals, for example, Kent and Hochstein (1998) provide a rule-of-thumb to assess the type of competition that can be introduced at different traffic levels.

<table>
<thead>
<tr>
<th>Type of competition</th>
<th>Traffic level (TEUs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-terminal</td>
<td>30,000</td>
</tr>
<tr>
<td>Inter-terminal</td>
<td>100,000</td>
</tr>
<tr>
<td>Inter port</td>
<td>300,000</td>
</tr>
</tbody>
</table>

Source: Kent and Hochstein (1998)
In situations where monopolistic and competitive markets are vertically-bundled, monopolistic terminal operators will seek to recover the profits foregone by regulation by controlling a related market (Paredes, 1997). This could be possible because international trade requires logistics services that are provided through supply chains. Since ports are crucial nodes in these chains, the party that controls a monopolistic terminal would be in position of excluding rivals from the provision of the other services that conform that chain. This can be done, for example, by giving a more favorable treatment to firms belonging to the same business group (granting exclusive rights to use terminal’s assets or allowing preferential entry to its premises), discriminating against non-related companies (not providing services with the same expediency or quality) or increasing entry barriers (requiring competitors to have over-qualified personnel, equipment with specific characteristics or disproportionately large insurance cover). OECD (2006) provides examples of such behavior in Europe.

The limited traffic generated by small ports may facilitate this behavior, because it causes many port services to be provided by a handful of firms. This means that the few companies that act as shipping agencies also provide pilotage, towage, mooring and purveyance to ship-owners, and may act as customs brokers and cargo consignees too. In the abovementioned case of Matarani Port, for example, the same business conglomerate that holds the concession of the port’s single terminal provides all other mentioned services through related companies. Moreover, there have been complaints that the concessionaire has asked non-related firms for disproportionately large financial guarantees as a requisite to use the terminal’s facilities (Alcazar and Lavation, 2005). The case is similar in other ports, not necessarily as small. In the port of Cartagena (Colombia), for example, the concessionaire Sociedad Portuaria Regional de Cartagena has been accused of using its advantageous position (it is the exclusive provider of infrastructure and superstructure) to drive competing stevedoring companies out of the market (World Bank, 2007). In these situations, regulation is needed not to determine prices but to guarantee access of competing firms to non-monopolistic markets (access regulation).

5.3.3 Access policies for naturally monopolistic port terminals

Simply stated, an effective access regime should impede a monopolist to discriminate against non-related firms in the provision of competitive services. To achieve this, the regime should ensure that any firm is able to access terminal’s facilities under “fair” conditions, i.e., similar to those requested in competitive markets. However, the many factors policy makers should consider when designing such regimes converts access into one of the most complex regulatory policies (OECD, 2006).

The first issue that arises is to what and for what to grant access. Not every port asset is required to provide port services nor every service requires safeguards for firms to enter the market. Since regulation is more costly to society and is more likely to spawn inefficiencies than allowing markets to work freely, regulators should be very
careful not to intervene when it is not warranted. One can recall from chapter 3 that the EFD is a useful criterion to decide when to intervene and to what facilities grant access, but also that its use carries large risks and trade-offs for which is necessary to avoid misuse.

The second issue that arises when designing access policies is how much entry to allow. In theory, an optimal entry is the number of competitors that would enter the market if the terminal was not monopolistic. Unfortunately, this is not observable by the regulator and, as explained before, a regulated monopolist has incentives to restrict entry in order to favor itself or related firms. Moreover, since the regulator has less information about the true cost of providing infrastructure than the terminal operator, the probability of regulatory failure is high and its consequences dear. In fact, if access prices and conditions are set below the optimal, the terminal operator would not have the incentives to adequately maintain and expand the infrastructure, which could lead to congestion, delays and other inefficiencies. If access terms are set too high, a less-than-optimal number of firms will enter the market, for which economic rents will not be competed away and operational efficiencies will not be achieved (Laffont and Tirole, 1994).

As we will see in the following chapters, formulating access policies involves making decisions or setting rules regarding four main issues:

a. Vertical structure;

b. Pricing;

c. Non-price terms and conditions; and,

d. The mechanism to expand the infrastructure.

Vertical structure

It refers to the provisions allowing or prohibiting terminal operators, for example, to integrate with shipping companies or to supply related services (storage, towage or bunkering). Under vertical integration, a single firm is allowed to supply several services, while vertical separation implies that several firms would be needed to provide different services (OECD, 2001).

Vertical separation can be legal or operational. Legal separation implies that activities have to be performed by different companies; while operational separation only requires activities to be performed by different business units within the same company. To enforce operational separation, measures such as accounting separation are commonly implemented.
The importance of regulating vertical structure is that incumbents have radically different sets of incentives under integration or separation. When vertical separation is enforced the incumbent has no incentives to discriminate, since it cannot provide the service directly and there is no related company to favor with this behavior. It would be more profitable to charge for access to as many companies as possible, regardless of their affiliation. Vertical separation could then resolve the access problem outright if there were not economies of scope and transaction costs. Indeed, this policy would require services to be provided by different companies even if it is more efficient to supply them jointly, a practice that also involves transactions costs (coordination and administrative costs derived from having additional suppliers). As a consequence, vertical separation could cause higher prices than integration.

One can notice the importance of economies of scope and transaction costs for the transport industry by observing its trend toward integration. Indeed, several shipping companies also operate container terminals (Maersk, MSC, APL, Evergreen, among others) and seem to be increasingly interested in land-side operations. The growth of inter-modalism and the fact that in many ports few firms provide several services also constitute examples of this trend (ADB, 2000).

Another factor to consider when deciding about vertical structure is that reduced traffic also implies reduced businesses. In small ports, not allowing an operator to provide other services would reduce even more the port or terminal’s potential to generate profits and with that, the private sector’s interest in taking a concession (Flor and Defilippi, 2003).

Pricing
Designing an access regime requires deciding how access prices will be determined. There are two options: (i) regulation, with the already-discussed shortcomings caused by the information asymmetry between regulator and regulated firm; and (ii), negotiation by the incumbent and the access seeker with regulatory intervention limited to cases when the parties do not reach an agreement. Regulators also face the problem of deciding what methodologies to use to set prices under option (i), or to settle disputes under option (ii).

Section 3.2 discussed several pricing methodologies: ECPR, FDC, LRIC and global price caps. It showed that every methodology has advantages and shortcomings, for which is difficult to decide a priori which one to use in the port industry. However, when implementing access policies, a regulator should state clearly which methodology it would use under particular situations.

Non-price terms and conditions
Alternative policies to determine terms and conditions are negotiation, regulation or intermediate approaches, such as the publication of standard reference terms that can
be subject to further negotiation by the parties. Regulating non-price access conditions is especially difficult because it involves subtle issues that offer opportunities for discrimination. OECD (2006) gives several examples of non-price access deterrence in European ports:

- In the UK, an incumbent (Sealink) owning a common-user port and providing a ferry boat service undermined an entrant’s (B&I) ability to operate by scheduling its own service to disrupt the loading and unloading of passengers by the entrant.

- The port of Rødby, Denmark, is owned by the state-owned company DSB, which also provides ferry services. A company interested in entering the ferry market (Stena) was denied access to the port on the grounds that it would prevent companies already operating in the port from expanding their activities. Moreover, Stena’s efforts to build a new terminal were blocked by the government on behalf of the state-owned incumbent.

- In the port of Varna, Bulgaria, a catering firm complained that although it had acquired all of the permits required to cater and to dispose of waste from maritime vessels, it was unable to operate because the port management refused to dispose the waste brought by the caterer and also refused to allow its trucks entering port property.

An important non-price issue when designing access policies is how to distribute berths’ available time among the ships calling at the port. Indeed, a shipping company may want to schedule a service where the ship arrives at a time when the berth is being used by another ship. Unless access rules have been adequately established (in this case, the terminal’s queuing policy), an integrated terminal operator might try to discriminate by favoring itself (if it is integrated with the shipping company) or the client that uses more of the other services it provides. One option to set these rules is to allocate the time slot to the company that was already using it, but this raises the question of what would happen if the shipping company wants to rent or sell the slot to a third party. On one hand, this would be desirable, since the scarce resource (the slot) would be ultimately used by the company that values it most. On the other hand, this might not be fair with the terminal operator, since it would not benefit from earnings generated by terminal’s assets. A second option is to auction the slot, which raises the question of what to do with the proceeds. If they are considered earnings, the monopolistic operator may end up obtaining supra-normal profits, thus evading regulation (Parker, 2000).

The typical non-price terms included in access agreements are:

- Services that the access seeker will provide using incumbent’s assets;

- Assets that the access seeker will be allowed to use;
c. Minimal equipments, and their technical characteristics;

d. Personnel, number and their qualifications;

e. Schedule under which the services will be provided;

f. Procedures to coordinate activities between incumbents and the service provider;

g. Responsibilities of the parties;

h. Investments that need to be undertaken by either party;

i. Guarantees that the parties will comply with the agreed terms. This may include insurance policies or other financial instruments such as letters of credit.

The mechanism to expand the infrastructure

The fourth element that designing an access regime involves is the set of incentives used to expand the infrastructure.

The first option is through regulation. The regulator may monitor congestion and mandate the port operator to expand the infrastructure when it reaches certain levels. This option has, however, two main shortcomings. The first is caused by the abovementioned information asymmetry between regulator and regulated firm. The regulator may ignore that a shift in demand has occurred, or the existence of cheaper options to expand infrastructure or new technologies used to ease the congestion without investing in expensive infrastructure. It is important to consider that mandating unwarranted infrastructure expansions harms consumers, since regulated tariffs are set taking into account these investments. The second shortcoming is that the mandate to expand infrastructure may be motivated for political, not technical, considerations. As explained in preceding chapters, one of the main regulatory failures is caused by the regulator’s capture from the monopolist. However, the regulator may also be captured by the government and willing to serve its political agenda. Therefore, allowing regulators to make this decision might increase regulatory risks and the private sector’s interest in taking out a concession.

A second option is to give the incumbent economic incentives to expand the infrastructure when necessary. Under this approach regulators have several policy options:

a. Giving financial rewards for keeping congestion under acceptable levels. If the rewards are profitable, the incumbent will be enticed to expand the
infrastructure when necessary. As we will see, this is the approach used in the UK’s electric industry. The problem with this option is that by paying the incumbent for doing what similar firms in competitive markets would anyway do, it allows obtaining monopolistic rents.

b. *Using market incentives.* The problem with this approach, as seen in chapter 2, is that a monopolist has incentives to reduce supply, for which it might be more profitable not to expand the infrastructure. This option seems to be better suited for incumbents that face some degree of competition, either in the same market or in related ones. For example, a monopolistic terminal operator that faces inter-modal competition. As we will see, this is the option used in natural gas and railways industries in the US.

c. *Awarding subsidies.* Under this approach, the government gives incumbents subsidies to undertake otherwise unprofitable investments or to provide otherwise unprofitable services. In network industries, this mechanism is used to cover poor, rural or isolated areas (in telecommunications these are called “universal service obligations”).

The mechanism used for the concession of the Matarani port terminal (Peru) was to establish a threshold. The concessionaire has the duty of building a new berth once any of the actual ones reach an occupancy rate of 70% (Alcazar and Lavation, 2005). The concession of the Buevaventura port terminal (Colombia) did not foresee a mechanism to expand the terminal’s capacity, which caused congestion when traffic increased more that expected. Moreover, because of limitations on the duration of the concession contract, the concessionaire was reluctant to incur additional investment costs unless it could be assured of a reasonable cost recovery period for their added investments (Nathan Associates, 2004)
Part II: Access policies in regulated network industries
6. Network economics

This chapter describes the main characteristics of network industries and their relation with diverse market structures. This discussion allows understanding how the complementarity between the components of a network may allow a monopolist controlling an essential input to extend its market power to related competitive markets.

The chapter also analyses how these characteristics might lead to market structures that do not necessarily result from lack of competition or anti-competitive behavior.

6.1 Architecture of networks

A network is an interconnected system of links and nodes. Network industries are those which possess physical or electronic linkages that create networks. Industries with such characteristics that have raised public policy concern are telephone, broadcasting, cable television, electricity, water pipelines, sewage systems, oil pipelines, natural gas pipelines, road and highway systems, bus transport, truck transport, airlines, inland water transport, liner shipping, postal services, package delivery systems, refuse pick-up systems, airline computer reservation systems, bank automated_teller machine systems, bank and non-bank credit card systems, bank debit card systems, bank check and payment clearance systems, local real estate broker multiple listings services and the internet (White, 1999).

Figure 6.1 shows a simple star network that could represent a hub-and-spoke system. Cargo from either node is routed through a central node, the node S, which in this case may be a hub port as Gioia Tauro (Italy) or Balboa (Panama). This structure economizes on the number and length of links necessary to provide all possible node-to-node transactions, but requires central nodes to have the capacity to handle all transactions among nodes.
The structure shown in figure 6.2 economizes on the number of links, making the central node unnecessary, but some transactions would have to travel longer distances than in the previous case. Additionally, the links would need greater capacity, because they also provide third-party transport.

Figure 6.3 shows a structure in which all points are directly connected. This architecture minimizes the distance to complete a transaction, but requires the maximum number of links. Highway networks connecting main cities follow this layout.
Figure 6.4 shows two star networks and a trunk link (S_A S_B) that connects their central nodes. S_A and S_B could be two large ports such as Rotterdam and Singapore and both star networks could be the hub-and-spoke systems developed around them. In this context S_A S_B may represent a main maritime route. Connections among stand-alone networks create larger networks.

Networks where transactions AB and BA are different are named “two-way” networks. These are air and maritime transport, road and railroad systems, telephone and postal services, email, etc. “One-way” networks arise when AB and BA are identical or one of them is unfeasible, like in electricity supply or natural gas distribution, where transactions can flow only in one direction. In a typical one-way network, consumers demand composite goods formed by the combination of several goods or services.
It can be seen that the most distinctive feature of both type of networks is the complementarity between links and nodes. In fact, although there are many non-network industries where final goods are composed of complementary components (such as electronics or equipment manufacturing), goods in network industries are produced using both links and nodes as production inputs.

In networks where links and nodes are owned by different firms, the issues of interconnection, compatibility, and coordination become crucial for the production of composite goods.

6.2 Economic characteristics of networks

6.2.1 Externalities and complementarities

Networks exhibit positive consumption and production externalities. This means that the value of a unit of a good increases with the expected number of units to be sold. For this reason, the demand curve slopes downward, but shifts upwards with increases in the number of units sold. The key reason of the appearance of network externalities is the complementarity between the components of a network (Farrell and Saloner, 1985). For example, for a “n” components two-way network like the one depicted in figure 6.1, there are “n(n-1)” potential goods. An additional customer provides direct externalities to all other customers in the network by adding 2n potential goods through the addition of a new link to the existing links”.

In typical one-way networks, however, the externality is indirect. In figure 6.5, for example, if all A-type goods are compatible with B-type ones, there are “mn” potential composite goods. An additional customer yields indirect externalities to other customers of the network by increasing the demand for components of types A and B, thus potentially increasing the number of varieties of each component that are available in the market.

A different way to see this effect is through unexploited economies of scale. For example, if there are unexploited economies of scale in the provision of potable water, a new customer reduces the costs of serving all other customers of the network, thus generating positive externalities.

6.2.2 Economies of scale, scope or density

The links and nodes of a network may exhibit significant economies of scale, scope or density for some technologies and for low volumes of transactions. As a matter of

---

37 Economies of scale and scope were defined in section 2.1.2. Economies of density arise when is cheaper to add a new link to an existing network than building a new one (Fernandez-Baca, 2006).
fact, either economies of scale or constant returns to scale can support the externalities of two-way networks and, as said before, the presence of unexploited economies of scale may explain the occurrence of externalities in one-way networks.

Economies of scale, however, do not pervade all aspects of all links or nodes at all times. For some technologies, a sufficiently large volume of transactions may exhaust economies of scale, thus maintaining the average and marginal costs roughly constant, or even increase them for the relevant demand interval. In these cases, multiple competitors for some or all of the links and nodes of a network are likely to arise, as in shipping, airlines, mobile and long-distance telephony, etc. (Economides, 1996). As it can be seen from the examples given, the exhaustion of economies of scale, scope or density may occur for an entire network or for some parts of it. For example, the volume of traffic between two important ports like Rotterdam and Singapore are much greater than between Rotterdam and any minor port. Therefore, it is possible that the economies of scale of serving the pair of large port is exhausted with the larger volume, while unexploited economies of scale could still be available in services between Rotterdam and the minor port.

It is worth noting that even if competition exists in most of the components of a network, the links and nodes where economies of scale, scope or density are not exhausted may become monopolies. Moreover, due to the feature of complementarity among the components of a network, the existence of monopoly in one of the links or nodes may be sufficient to capture all the potential rents from the transactions that use that component (Paredes, 1997). This can be illustrated by using figure 6.4. As we can see, the nodes $S_A$ and $S_B$ are needed to complete any transaction between $A_i$ and $B_j$. The monopoly ownership of $S_A$ is sufficient to capture any rents that are available from the transaction. But if node $S_B$ is also owned by a monopolist, its intention to maximize profits may lead to even higher prices, lower output and greater inefficiency in the network. In this case, vertical integration (the single ownership of both $S_A$ and $S_B$) may be desirable and may even increase efficiency by eliminating the double-monopoly problem.

The existence of a monopoly in a link or node of a network creates further problems when an incumbent competes in a related market with other firms in the provision of goods that use the monopolized link or node as an input. In that case, the monopolist has incentives to foreclose or refuse to deal with its competitors, or perform a price-squeeze by overcharging them for the use of the node or link it controls exclusively (Kessides, 2004). This is the cause of the access problem.

### 6.2.3 Compatibility and standards

If a transaction within a network is to be successful, the various links and nodes must be compatible with each other. Depending on the industry, compatibility may involve not only formal technology, such as software or electronics, but also physical
standards like rail tracks, containers and runways. However, compatibility needs not to be perfect. Some decrease in quality or speed maybe acceptable as long as the transaction costs are not significantly increased. The way members of a network achieve compatibility is by agreeing in the set of standards on which transactions will be based upon. The problem arises in cases where the network is owned by different firms, since the adoption of one standard over another favors one participant over the others. To enhance their market power, incumbents may agree on a technology that discourages potential entrants or harms actual competitors (Economides and White, 2004).

In this case, an explicit agreement or implicit understanding will be necessary. Such an agreement can be negotiated, imposed by the government or surge through imitation. Stover (1961) illustrates this point by describing how the railway industry in the US adopted a standard gauge during the nineteenth century:

“The diversity of [track] gauge, especially in the South, (...) made impossible the cross-country shipment of freight without break of bulk. (...) Standard automatic couplers and air brakes could do little to speed or facilitate freight traffic as long as a diversity of [track] gauge persisted. In 1861 more than 46 per cent of the nation’s rail mileage was other than the 4 feet 8½ inch standard gauge (…).

Several expedients were used in the sixties and seventies to permit the interchange of equipment between lines of different gauge. The “compromise car” was the simplest. Having wheels whose tread was five inches wide, the cars could be used on either standard-gauge track or track as wide as 4 feet 10 inches. However, careful railroad operators frowned upon the use of such cars because they claimed that many accidents could be traced to them. A second innovation, the car with wheels that could be made to slide along the axle, was no safer and was never widely adopted. Car hoists, or “elevating machines,” with the cars lifted to a set of trucks of different gauge, were much safer and were used extensively. . . . A number of lines also went to “double gauge,” the addition of a third rail, permitting the use of equipment of different gauge.

There was no substitute for the adoption of a single gauge by the whole nation. In the early eighties most of the gauge divergence was found in the narrow-gauge lines of the mountain West and in the Old South, where the five-foot-gauge mileage had actually increased from 7,300 miles during the Civil War to more than 12,000 miles in 1880. The Chesapeake & Ohio, the Illinois Central, and the Mobile & Ohio had all changed to the narrower standard gauge by the middle eighties. James C. Clark, general manager of the Illinois Central lines south of Cairo, spent weeks of careful preparation for his change of gauge in the spring and early summer of 1881. On July 29, 1881, between dawn and 3:00 P.M., 3,000 workers shifted the gauge on the entire 550-mile line.

The rest of the South soon gave in. Representatives of southern lines totaling more than 13,000 miles agreed early in February 1886, to change their gauge the following May 31 and June 1. During the weeks before the day of the change, part of the southern rolling stock and motive power was changed to the narrower gauge and track gangs moved alternate inside spikes on one rail to the new position. . . . Using track gangs of from three to five men, ten roads west of the mountains shifted their rail on the last day of May. The remaining roads shifted June first. On both days the
work was accomplished between 3:30 A.M. and 4:00 P.M., during which all traffic was stopped. Southern railroads had truly become part of the national network, and passenger or freight trains could move from any southern depot to any part of the nation without change of trucks or bulk (...).

An immediate dividend of increased operating efficiency resulted from gauge standardization as a system of more extensive car interchange developed among the railroad companies.  

6.3 Consequences of network effects

The characteristics of network industries discussed in the previous sections have several important implications over the structure of markets within network industries. The first one is that monopoly may maximize social surplus. This is because in the presence of robust network externalities, a large market share boosts economies of scale, scope or density, thus generating productive efficiency (goods or services will be produced using the least quantity of inputs). The problem, of course, is how to ensure that allocative efficiency (prices are set as close as possible to costs) is also reached, since a monopolist has no incentives to do it. A relevant characteristic of markets within network industries is that firms that can set the industry’s standards tend to have market shares several times larger than their followers. And this, in turn, tends to have a much larger share than the one of the third, and so on. This geometric sequence of market shares implies that in equilibrium, there is extreme market share and profit inequality (Economides and White, 1994).

This characteristic can be the consequence of the fact that firms setting the market standard have also higher sales of complementary goods and, therefore, their goods are more valuable to consumers (Gabszewicz, Sonnac and Wauthy, 2001). This feedback results in even higher sales. On the contrary, firms with small market shares have lower sales of complementary goods, the feedback resulting in even lower sales. For example, it is easier for a fixed-telephony incumbent to sell a related service (for example, a mobile subscription) to an existing client than to a new firm that has to enter the market. One can argue many reasons for this result: the user already knows the incumbent, there is an additional cost of dealing with more providers, fears that compatibility may not be complete, etc. Nevertheless, since the position of the high sales firm allows it to charge high prices, low sales producers may not be driven out of business. The existence of large liner shipping companies along with relatively small ones may illustrate this point. Although the shipping industry has been immersed in a path toward consolidation (companies tend to be larger), small companies have not been driven out of business.

Another important implication of the network effects is that additional users are not rewarded for the positive externalities they cause to others, although they may discriminate to favor large providers in order to maximize the network effect

---

38 Source: Stover (1961), pp. 149-156.
(Spulber, 2008). This can be the case of liner shipping or parcel services, in which new users would prefer to use the services of the firm with the largest coverage.

A further implication of the existence of network effects is the importance of “path-dependence”\(^\text{39}\), i.e., investments made in the past condition investments to be made in the future (Rena and Herani, 2007). For example, since expansion projects in the rail industry use the existing infrastructure, they tend to perpetuate actual track standards, even though they or the network’s layout reflect economic conditions from the time when it was initially developed.

### 6.4 Network effects and public policy\(^\text{40}\)

As we have seen, one can expect competitive network markets in equilibrium to have a “winner-takes-most” structure, with significant inequalities in shares and profits. However, this outcome does not mean that competition is weak. On the contrary, striving to become the top firm in the market and thus obtain most of the benefits makes competition very intense. Therefore, the existence of network effects may cause competition “for the market” to take precedence over competition “in the market”.

Another important lesson for governments is that there should be no presumption that the inequality in profits among competitors is the necessary consequence of anti-competitive actions by incumbent firms. As we have seen, market structures in this type of markets are expected to be unequal.

The third important consequence of the existence of network effects is that the breakup of a monopoly into competing firms with incompatible standards may raise production costs and thus reduce social surplus. In this context, standardization is valuable even if it is achieved de facto by a monopolist.

Lastly, policy makers should notice that free entry to markets within network industries does not lead to perfect competition. In these industries, once few firms are in operation, the surge of new competitors does not change the market structure significantly. Therefore, governments should not expect to affect the market structure in a significant way by encouraging competition through the elimination of entry barriers.

---

\(^{39}\) “Path-dependence” is the dependence of a network on past decisions of producers and consumers. See Ellig and Lin (2000)

\(^{40}\) White (1999)
7. Access regulation in the telecommunications industry

A large part of theoretical body available today for regulating network industries was initially developed for telecommunications. Indeed, the problems posed by the British government’s decision of privatizing British Telecom encouraged economists to devise new ways to regulate private monopolies. Access regulation was practically created to allow the development of technical innovations as mobile telephony and the internet. Since then, a large number of countries have opted for restructuring their telecommunication industries.

The following sections describe the economics of the telecommunications industry and the alternative policies followed to deal with the access problem. The particular approaches taken in the UK, US and Australia are analyzed. The last section summarizes the lessons for ports.

7.1 Economic characteristics of the telecommunications industry

Telecommunications is a network industry with the characteristics described in chapter 6. Telecommunication networks require complementary and compatible inputs to produce one call, and they exhibit economies of scale, scope and density. Moreover, the inclusion of a new subscriber increases the value of the network for all existing users. These externalities make some services to be produced cheaper if the market is provided by only one firm.

Although a variety of technologies can be employed to provide telecommunications services, the basic architecture of telephone networks is similar to the one shown in figure 7.1. End users are connected to local switches. In order to complete a call, operators connect the local switches of both caller and receiver using alternative technologies.
The telecommunications industry shares with the port industry the characteristic that both produce public and private services. Indeed, the benefits produced by this industry are much larger than those captured by private transactions. Telecommunications constitute one of the basic components necessary to ensure access to information, considered a strategic factor to reduce poverty and enhance development (Chowdhury, 2002). As in ports, cost recovery from end users is possible and provision can be limited to those who pay for it.

Telecommunication networks also share with ports the characteristic that their infrastructure supports several markets. Table 7.2 shows that telecommunications operators compete in five different markets: local, long distance (domestic and international) and mobile telephony, as well as in the provision of internet and data services.

Table 7.1: Telecommunications markets: typical competitive structure

<table>
<thead>
<tr>
<th>Market</th>
<th>Typical market structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local telephony</td>
<td>Monopoly</td>
</tr>
<tr>
<td>Long distance</td>
<td>Competitive</td>
</tr>
<tr>
<td>Mobile telephony</td>
<td>Competitive</td>
</tr>
<tr>
<td>Internet services</td>
<td>Competitive</td>
</tr>
<tr>
<td>Data services</td>
<td>Competitive</td>
</tr>
</tbody>
</table>

41 Source: Productivity Commission (1997)
It is important to note that in order to compete in these markets, operators should be interconnected, i.e., their networks should be linked in such a way that voice or data flow seamlessly between networks owned by different operators. In telecommunication contexts, interconnection agreements allow one operator to use parts of the networks of other operators, including both physical infrastructure and software systems. These agreements permit operators to give users of their own networks access to the users or services of other networks.

Interconnection is required, for example, for a call made from a mobile network to be completed (terminated) in the fixed-line one owned by the incumbent (or another mobile operator). In the network depicted in figure 7.1, this would be implemented by linking the cellular switch of the mobile operator with the local switch of the incumbent.

It is worth explaining how operators charge for their services in this industry. Although there are several options, the most used is the one known as “caller pays”. Networks set the tariffs for their subscribers. When a caller subscribed to company A calls another subscribed to company B, company A charges the caller and pays access charges to company B for the right of terminating the call in its network (also called “termination charges”). This situation creates incentives for companies to subsidize outgoing calls as a way to attract subscribers and later charge expensive access fees. This behavior prompted the European Commission to open an investigation into pricing between fixed and mobile networks in 1998 and to establish price caps for roaming charges in 2007 (European Parliament, 2007).

Since local telephony almost universally constitutes a monopoly due to the described characteristics of distribution activities, its operator is typically regulated. As in ports, this incumbent has incentives to recover profits foregone by regulation by monopolizing any of the remaining markets (Paredes, 1997). This can be done by refusing to subscribe interconnection agreements with other operators or by discriminating in favor of itself or a related firm. Therefore, access regulation is required to safeguard competition.

Incumbent’s ability to monopolize services in other markets comes from what is known as the “consumer lock-in problem”. This problem derives from the fact that in most countries telephone services are priced as a two-part tariff, comprising a fixed charge for subscription plus a variable one for usage. The existence of a fixed subscription charge has two main effects that reinforce the incumbent’s market power (Lapuerta, Benavides and Jorge, 2003):

a. Fixed subscription charges lock in users to only one distribution network of each type for a period of time.

42 In other industries, these would be called access agreements.
b. It causes users to be interested in the range of services offered by the company they are already subscribed (since subscribing to a second company would involve paying another fixed charge), thus creating a large entry barrier for new entrants.

7.2 Markets and services

As in similar industries, it is possible to decompose a telecommunications network into transmission and distribution activities.

7.2.1 Transmission

Although terminology varies among regions, transmission refers to the high-capacity (high-bandwidth) networks that connect the large switches of various operators (in terms of figure 7.1, these would be the transit or secondary switches).

A number of technologies can be used to provide transmission services, such as cooper wires, coaxial cable, digital lines and wireless technologies, among others. The main cost characteristics of the most important technologies used for transmission activities are shown in table 7.1.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Initial Investment</th>
<th>Sunk Costs</th>
<th>Economies of Scale</th>
<th>Economies of Scope</th>
<th>Economies of Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional Local Loop</td>
<td>Large</td>
<td>High</td>
<td>Large</td>
<td>Small</td>
<td>Large</td>
</tr>
<tr>
<td>xDSL</td>
<td>Small</td>
<td>Low</td>
<td>Small (or negative)</td>
<td>Medium</td>
<td>Negative</td>
</tr>
<tr>
<td>Improved Local Loop</td>
<td>Large</td>
<td>High</td>
<td>Large</td>
<td>Large</td>
<td>Large</td>
</tr>
<tr>
<td>CATV</td>
<td>Large</td>
<td>High</td>
<td>Large</td>
<td>Large</td>
<td>Large</td>
</tr>
<tr>
<td>Digital Powerline</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
<td>Small</td>
<td>Small</td>
</tr>
<tr>
<td>Fixed Wireless Narrowband</td>
<td>Small</td>
<td>Low</td>
<td>Small</td>
<td>Small</td>
<td>Small (or negative)</td>
</tr>
<tr>
<td>Fixed Wireless Broadband</td>
<td>Small</td>
<td>Low</td>
<td>Small</td>
<td>Medium</td>
<td>Small (or negative)</td>
</tr>
<tr>
<td>Mobile Wireless</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Small</td>
<td>Small</td>
</tr>
<tr>
<td>Mobile Satellite</td>
<td>Large</td>
<td>High</td>
<td>Medium</td>
<td>Small</td>
<td>Small</td>
</tr>
</tbody>
</table>

Source: Cas (1999)
According to Cas (1999), the high level of fixed and sunk costs of wire-based networks and the limited availability of frequencies for wireless networks, constitute entry barriers. However, given the variety of technologies and the large number of potential players involved, these barriers seem not to be as large as to prevent competition. Indeed, satellite services, for example, cover large regions and can be accessed from many parts of the world, allowing the aggregation of small demands that might not be large enough to sustain a local network. Likewise, many poor areas in Asia, Africa and Latin America are covered by mobile providers even if they are not serviced by fixed-wire operators.

7.2.2 Distribution

Distribution refers to the low-bandwidth services provided through a high geographic-density network that connects local switches to final consumers, whether fixed-wire or wireless. In fixed-wire systems, the physical connection between the local switch with the final consumer is known as the “Local Loop” (see figure 7.1). As shown later, the local loop constitutes a key element for providing access.

Fixed-wire distribution shows sizeable economies of density, for which the potential for competition for local telephony services is much more limited than in transmission. Economies of density arise from the fact that network costs per connection decrease rapidly as the number of connections increase, mostly because of shorter lines facilitate an efficient use of the network. Studies on interconnection costs report that population density and average line length together typically explain over 80% of the variation in network costs (Edwards and Waverman, 2004).

In addition to economies of density in distribution services, technologies such as fiber-optic cables also involve substantial installation costs and large economies of scale. This fact contributes to limit the potential for competition in fixed-wire distribution, even in metropolitan areas. In fact, several econometric models developed for the US suggest that only the most densely populated parts of large cities could expect to have two competing distribution networks (OECD, 2002a).

In theory, wireless technologies could facilitate competition in local telephony, but their cost is still high compared to fixed-wire ones. However, this may change in the future. Indeed, in spite of its costs, mobile services are gradually becoming substitutes of fixed services, especially in low-populated areas (like rural zones) where deploying fixed-wire networks would result even more costly. Furthermore, cable networks are increasingly delivering telephony and high-speed internet services and new technologies are being tested to deliver voice, data and video products within the same line. It is worth noting that as substitutability increases; economic regulation of the industry will be less warranted.
This process may follow a similar path to the one occurred to European ports, where the disappearance of border controls, better transport infrastructure, containerization and other technological changes enlarged their historical hinterlands and allowed competition where previously was absent.

7.3 Reform and regulation in the telecommunications industry

The telecommunications industry has been heavily regulated almost since the invention of the telephone. Natural monopoly conditions and other network effects discouraged governments to rely on competitive markets for the organization of their national telecommunication industries. Therefore, prices, entry, ownership and availability of services were decided through administrative procedures (Melod, 1997).

Services were generally provided by state-owned companies legally protected from competition, which set prices and structured their supply to meet social goals and universal service obligations. This panorama changed radically from the 1980s. Since that time, a large number of governments have opted for restructuring their telecommunications industries, open their markets to competition and privatize their Public Telecommunications Operators (PTOs).

Both technological advances and market changes motivated these reforms. For example, the development of technologies that allow reaching final users bypassing fixed-wire networks questioned the natural monopoly status of some services (Boylaud and Nicoletti, 2001). In addition, the gradual increase in the demand for telecommunications services have made possible several operators to operate simultaneously while recovering their sunk costs.

Intver, Olivier and Sepúlveda (2000) argue that other factors might have also motivated the restructuring of the industry:

a. Government’s need to attract private sector capital to finance the expansion of existing networks;

b. The growth of the internet, which increased data traffic and made apparent the need for upgrading existing networks;

c. The growth of wireless services, that provided alternatives to fixed-wire networks and introduced new competitors; and,

d. The development of an international telecommunications market, which changed the traditional structure of the industry and allowed private
companies to undertake investments that were previously only available to governments.

It is worth noting that unlike ports, where privatization has not excluded public port authorities from continuing to provide port services (dredging, VTS, and security services, among others), PTOs have been privatized by transferring their control to private investors; thus excluding government from the direct provision of telecommunication services (Rozas, 2005).

However, privatization has not implied that telecommunication services are not longer considered as public services; i.e., those that have to be provided even if it is not economic to do so (United Nations, 2001). On the contrary, one of the goals of regulation is precisely to ensure that the provision of telecommunication services is consistent with broader public policies.

Considering the costs that regulation imposes on society and the possibility of regulatory failures, reforms implemented in the industry have been oriented at introducing competition in as many markets as possible; implementing economic regulation only in those that remain monopolistic. In this sense, privatization is not considered a goal per se (at least not from an economic viewpoint), but a reform tool to allow competition to be introduced in telecommunication markets (Rozas, 2005).

Most countries that have restructured their industries have also created regulatory bodies or invested existing ministries with regulatory powers to oversight the privatized PTOs. Most regulatory agencies regulate principally three economic issues: price, quality and the conditions of access telecommunication markets (Melod, 1997).

Table 7.3 shows how regulation is implemented in OECD countries after reform and privatization. It can be seen that in local telephony markets, regulation tend to be asymmetric (applies mostly to PTOs or dominant operators). It can also be seen that retail prices (those charged to subscribers) tend to be set using incentive-based methodologies (price-caps estimated using RPI-X or similar ones); while access charges are set using cost-based methodologies (ECPR, FDC or LRIC). It is worth noting that a number of countries do not regulate access charges to incumbent’s networks.
As for mobile markets, table 7.3 shows that most countries do not regulate retail prices (presumably because of existing competition), but those who do it apply different approaches. It also shows that access charges tend to be set using cost-based methodologies.

The alternative to sector-specific regulation to set access prices and terms is to rely on general antitrust law. The experience with this option is nonetheless discouraging. New Zealand, for example, fully liberalized its telecommunications market in the late 1980s and relied primarily on antitrust instruments to regulate access. But when a competitor already operating in the long distance market attempted to provide local telephone services the incumbent refused reaching an agreement on access terms, for which the dispute had to be appealed to the final arbiter, the UK’s Privy Council of the House of Lords (New Zealand's Supreme Court). The court took four years to settle the dispute. The main controversy was the setting of access prices, which lower courts were unable to do. In 2001, a telecommunications regulation unit was created within the antitrust authority, and a law passed requiring access prices to be set according to the LRIC methodology (Kerf, Nieto and Géradin, 2005a).

7.4 Access arrangements in the telecommunications industry

As in the port industry, incumbents lack of incentives to agree on interconnection terms with access seekers. In the absence of a regulatory framework, they may refuse to agree on interconnection terms (as the New Zealand case illustrates), delay agreements or set expensive access charges. They can do so because they control the

---

Table 7.3: Regulation of telecommunications in OECD countries

<table>
<thead>
<tr>
<th></th>
<th>Retail prices</th>
<th>Access charges</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Local (fixed-wire) telephony</td>
<td>Mobile markets</td>
</tr>
<tr>
<td>Regulation applies to:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All operators</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Dominant operators</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Only PTOs</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Type of regulation:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incentive-based</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td>Cost-based</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Discretionary</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>No regulation</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Countries surveyed</td>
<td>27</td>
<td>19</td>
</tr>
</tbody>
</table>

Source: Boylaud and Nicoletti, (2001)
fixed-wire network, including the local loop that links the network with the subscriber’s premises. Without access to the local loop, providers of long distance, mobile, internet or data services would not be able to enter these markets.

Two are the main policies to overcome the incumbent’s market power: Interconnection Agreements and Local Loop Unbundling (OECD, 2002a). As explained earlier, interconnection agreements allow operators to reach other operator’s subscribers by connecting their networks. It is worth noting that even though this policy is aimed primarily at reducing the incumbent’s market power, it is applied to all telephone service providers too. This ensures that all networks are interconnected, allowing any user to be reached from any telephone regardless the network they are subscribed to.

According to Intven, Oliver and Sepúlveda, (2000), many telecommunication regulators use the EFD is the guiding principle to decide to which network components operators should provide access. In this industry, essential facilities include local loops, public right-of-ways, telephone numbers, frequency spectrum and support structures such as poles and ducts.

In many countries access terms are negotiated by incumbents and access seekers (Jamison, 1998a). However, given the natural reluctance of incumbents to interconnect with competitors, regulatory supervision is required to establish reasonable interconnection agreements on a timely basis. To ensure agreement on terms and prices, regulators have adopted combinations of the following policies:

a. *Establishing guidelines that reduce the scope of negotiations to issues where the information asymmetry between regulator and incumbent is still large.* These guidelines are implemented to avoid incumbents trying to delay agreements by negotiating conditions that have become a standard in the industry or have previously been granted to other operators. The increasing availability of interconnection agreements (from the same or similar countries) facilitates the establishment of these guidelines.

The main problem with this approach is that regulators may fail to recognize topics that should be left open to negotiation. As in ports, non-price access terms in telecommunications involve subtle issues that are difficult to detect, measure and regulate, such as the hearing of noises during calls or the number of calls that not reach their destination. These can be attributed to normal failures or sabotage from incumbents.

b. *Setting default interconnection agreements that would be implemented if negotiations fail.* If the terms of these default agreements are less convenient for the incumbent than those it expects to reach through negotiation, it will be encouraged to negotiate. This is the approach implemented by the Federal Communications Commission of the US since 1996.
One problem with this approach is that some default terms may result favorable for the incumbent, thus reducing its incentives to negotiate an agreement. The information asymmetry between regulator and regulated firm may induce regulators to make this mistake. On the contrary, if the terms of the default agreement are set too low access seekers would not have incentives to negotiate. In this case, implementing the agreement might result in excess entry and may disincentive the incumbent to maintain and expand the infrastructure.

c. Establish deadlines for various stages of the negotiation process. This policy has the advantage of limiting the incumbent’s ability to delay negotiations. The problem is that regulated negotiation periods tend to be lengthy, because regulators need to ensure that periods in between deadlines are large enough as to allow and effective negotiation of interconnection terms.

d. Establish industry technical committees to deal with technical issues or the adoption of standards. The main advantage of this policy is that reduces the information asymmetry between regulator and regulated firm. The main inconvenience is that in developing countries is hard to find experts with deep knowledge of the industry not related to any of the network operators. Hiring international consultants constitutes an option but it tends to make the negotiation process more expensive to the access seeker.

e. Create incentives to complete interconnection agreements. This is the approach used by US legislation. The 1996 Telecommunications Act establishes that local operators will not be able to enter the long distance market unless they open their networks to competition by establishing interconnection agreements with other operators.

In recent years, many countries have further encouraged competition by implementing a policy known as “local loop unbundling” (LLU) aimed at facilitating competition in internet provision markets. While interconnection arrangements mandate incumbents to provide competitors access to the local loop, while under LLU incumbents must lease them the local loop and related facilities. This operations allows competing operators to substitute the services provided by the incumbent.

LLU requires incumbents to provide competitors access to components of their networks on a stand-alone basis. The EFD is also used as a criterion to determine exactly which components of the network need to be leased. According to the OECD (2002a), LLU policies have the following main benefits:

a. This policy incentives incumbents to innovate constantly, since lack of innovation may result in loss of market share to competitors.

b. LLU also prevents inefficient investment, since competitors do not have to invest in duplicating the local loop.
The implementation of LLU, however, has not been exempted from criticisms. The main ones are the following:

a. LLU reduces incentives for building new networks, since it allows competitors to enter the market using the one owned by the incumbent. This may maintain incumbents' dominance even if natural monopoly conditions do not prevail anymore.

b. If access is priced wrongly, excess entry will occur. Since the collection of these charges will not suffice for maintaining and expanding existing networks, the incumbent will not have incentives to do it.

c. It requires detailed technical coordination between operators, which may result in higher levels of technical failures.

d. This policy also requires detailed regulatory intervention, therefore increasing regulatory risks.

These disadvantages seem to have deterred regulators to implement this policy. Indeed, even though a number of countries have implemented local loop unbundling, new entrants have found extremely difficult to compete with incumbent operators. As a consequence, penetration has been slow. In the United States, the country with the highest number of unbundled loops, these represent only 5.5% of the total lines. In Japan, the rate is 2.78% and in Denmark, 1.47%. In other countries, the penetration is even lower (Umino, 2003). Besides the lock-in problem discussed earlier, other factors might explain this slow penetration. For example, inherited advantages from the time when PTOs were public-owned monopolies reinforce incumbents’ market power. Indeed, state-owned PTOs had little difficulty in obtaining rights-of-way to build their networks, which new entrants are unlikely to obtain.

It is worth mentioning that something similar occurred with ports. When they were operated under the premise of constituting “growth poles”, port authorities could expect to be transferred a critical asset such as land at a price that did not reflect its opportunity cost (Haralambides, 1997a). Private entrants are unlikely to obtain a similar treatment, which reinforces incumbents’ market power.

Since LLU requires incumbents to provide access to components of their networks on a stand-alone basis, the implementation of this policy has impacted prices in two ways. First, it has encouraged the elimination of the typical cross-subsidization of subscriber line charges through usage charges. Second, it has forced the de-

---

43 Telecommunication services are priced as a two-part tariff: a fixed charge for subscription plus a variable one for usage. To make these services affordable to most users, state-owned incumbents had the tendency of cross-subsidizing the cost of subscription charges with usage ones. The existence of cross-subsidies is not desirable under a policy that promotes competition in as many markets as
averaging of geographically-averaged charges, i.e., services that are priced evenly within a geographic area regardless the fact that costs may differ for particular users (Umino, 2003). This effect enhances operational efficiency, since services are charged according to their real cost.

It is worth mentioning that low-income users and others located in rural or low-density areas (those that use to benefit from geographically-averaged charges) still receive subsidies, although their source has changed. Indeed, PTOs in most countries are expected to meet “universal service obligations”, which are requirements aimed at assuring that consumers with special needs receive telecommunications services (Tardiff and Taylor, 2003).

The costs incurred by providing universal service obligations are funded in different ways. In the UK, costs are met by the PTOs themselves. In the US, companies obtain support from a federal fund to which all operators must contribute. In Australia, costs are calculated annually and apportioned by all operators (OECD, 2002a). Other countries finance universal service obligations with general tax revenues (Chile) or with revenues from spectrum auctions (Guatemala). In Peru, rural telephony is financed by a fund constituted by 1% of all telephone bills (Rozas, 2005).

7.5 Reform, regulation and access regimes in selected countries

7.5.1 United Kingdom

Until 1981 British Telecom (BT) enjoyed a monopoly on all aspects of network operation and equipment supply. Its origins date back to the establishment of the Electric Telegraph Company in 1846 and the emergence of other private firms that developed the first telephone networks in UK in the 19th century. These companies later merged, were taken over or collapsed until they were eventually transferred to the Post Office in 1912 (with just one exception). In 1969, the Post Office Telecommunications became established as a public corporation, holding the exclusive right of providing telecommunication services. Its name was changed to British Telecom in 1980, remaining under control of the Post Office. BT was privatized in 1984 without restructuring, as the sole owner of the existing network.

The Telecommunications Act of 1984 set the framework for BT’s privatization. The Act abolished BT’s exclusive rights to provide certain services, created the Office of Telecommunications (Oftel) to regulate the industry and established new duties for
the Secretary of State for Trade and Industry. Both agencies have the duty to promote
the interests of consumers and maintain and promote effective competition.44

In preparation for the incumbent’s privatization, a second operator, Mercury, was
licensed in 1982 to compete initially in local and domestic long-distance markets,
although later the company was authorized to compete in the international long-
distance as well. Until 1991, the government followed a duopoly policy, for which no
other company was allowed to compete with BT and Mercury. During the 1980s, the
latter provided services only to city-based business customers and high-value
residential users, representing only marginal competition to BT. This duopoly policy
continued until 1991, when the regulator awarded licenses to other companies to
provide services in virtually all markets with the exception on international long
distance, where BT and Mercury retained exclusive rights until the end of 1996
(Laffont and Tirole, 2000).

At the time of privatization, there was a cross-subsidy from international to local and
rural calls. In this scenario, complete liberalization would have resulted in prices for
local calls which would not allow BT to recover the cost of providing the service. The
problem was dealt by restricting the rate of adjustment of the fixed charge while
capping the overall weighted price, and setting access charges to include a deficit
charge. This deficit charge was nonetheless waived until an entrant Mercury had a
market share of 10 percent or BT’s market share fell below 85 percent. This situation
ended in 1996 when BT was authorized to rebalance its rates (Valletti, 1998). A
similar duopoly policy was followed for cellular services. The only two licenses were
awarded to Cellnet (partly owned by BT) and Racal-Vodafone. These companies
offered mobile service on an exclusive basis until two further mobile operators were
licensed in 1990.

Under the UK’s interconnection regime, all operators are required to interconnect with
all other operators. Operators designated to have “significant market power” have
special obligations:

a. To set transparent and cost-oriented access charges;

b. To provide interconnection to other operators on the same terms as for their
   own services;

c. To offer access at any technically feasible point; and,

d. To keep separate accounts for different services.

---

44 In December 2003, the Office of Communications (Ofcom) assumed the duties of Oftel, The
Independent Television Commission, The Broadcasting Standards Commission, The
Radiocommunications Agency and The Radio Authority.
The access regime has changed through the years. When it was first implemented, a larger role was given to negotiation for setting access price and non-price issues. Nowadays, most of these terms are regulated by Ofcom. Operators wishing to interconnect with BT must agree to the conditions established in a document called the “Standard Interconnect Agreement”, produced and published by the incumbent after discussions other operators. This document covers basic terms and conditions, billing, technical matters, and over 100 schedules for particular services; and is regularly updated under Ofcom’s supervision. Parties are not allowed to negotiate individual terms. Operators seeking access must be registered with the regulator and undertake a series of steps, such as supplying call routing plans, before a standard interconnect agreement can be signed. Access charges are also published. These charges have been developed jointly by the regulator and the incumbent, in consultation with other operators and representatives of the consumers. Access charges are set for four-year periods and indexed on the basis of an RPI-X formula, using the LRIC methodology and current cost accounting asset valuations. Disputes over access charges are referred to the regulator (OECD, 2002b).

An interesting case arose in 1997, when two mobile operators (Vodafone and O2) were determined as having significant market power in the mobile market. Unlike fixed incumbent operators, they were not obliged to offer cost-oriented access charges. An Oftel review concluded that call termination charges (access charges paid by a caller’s network for the right of terminate the calls on the receiver’s one) for fixed-line users were excessive, for which the regulator referred the companies to the Monopolies and Mergers Commission (MMC)45. The MMC ruled in favor of Oftel, for which the two operators were required to make a significant reduction in their call termination charges and accept price caps. In 2001, Oftel carried out a further review of competition in the mobile market. The review concluded that the pricing of call termination on mobile networks was excessive and to the detriment of customers. In light of these findings, the regulator decided to continue with the price caps over the next four years and apply it to all mobile operators (Oftel, 2003).

LLU was not adopted in the UK until 2001. Oftel established that the methodology to estimate charges for unbundled components should be based in LRIC plus a mark-up to allow the incumbent to recover common costs. The regulator also set the following pricing principles:

a. The price of the local loop and other necessary inputs will be cost-oriented;

b. BT should be able to recover all the costs incurred in providing local loops; and,

c. The charges will be initially geographically averaged, but BT may request its de-averaging if this is justified.

45 At the time the agency carrying out investigations regarding abuse of market power. The MMC was later replaced by the Competition Commission.
Universal service obligations are placed on the operators with significant market power, which also have to meet their costs. They must provide telephone services at uniform tariffs, free access to emergency services, public telephony, operator assistance, directory services, and special tariffs for disabled people and low income users, among others (OECD, 2002b).

7.5.2 United States

The telecommunications industry in the US is regulated jointly by the Federal Communications Commission (FCC) and the state Public Utility Commissions (PUCs). The FCC has jurisdiction over all inter-state matters and some intra-state matters where the federal law prevails. State PUCs have jurisdiction over intra-state matters such as local long-distance, prices and entry conditions into local markets.

Until 1984, AT&T was the largest provider of telecommunications services in the country. It was integrated not only in the long-distance and local telephony markets, but also into equipment manufacturing and research and development. AT&T was organized into several local operating companies and a long-distance carrier. The local operating entities were regulated at state level by PUCs, while the long distance carrier was regulated by the FCC. AT&T’s monopoly was initially challenged in 1972, when MCI was granted permission to build and operate a microwave link between Chicago and St. Louis. This link also allowed other operators to compete in major city-pair routes. The behavior of AT&T toward its rivals induced the Department of Justice to bring an antitrust case against the firm in 1974. The case was finally settled in 1982, requiring the breakup of the company into seven regional companies. As a result, the country was partitioned into 192 local areas each served by a local company. The companies resulting from AT&T’s break up (called “baby Bells”) were banned from supplying long-distance services, which were to be provided by competing long-distance carriers (Armstrong, Cowan and Vickers, 1994).

In 1996, the US congress passed the Telecommunications Act, aimed at promoting competition and reducing the regulatory burden in telecommunications markets. It was implemented by three major FCC orders on local competition, universal service; and access charging. The Telecommunications Act set the framework under which competitive access providers could enter local markets and established that the regional Bells could enter the long-distance market if they opened their local markets for competition. In particular, the Act requires incumbents to undertake the following actions:

a. Interconnect their networks with those of other carriers at just, reasonable and non-discriminatory rates;
b. To lease unbundled network elements, at just, reasonable and non-
   discriminatory rates;

c. Sell retail services at wholesale rates for resale by competitors to end users.

The Telecommunications Act regards local loops as essential facilities controlled by
the incumbent, mandating that interconnection should be provided at least at four pre-
established points (Kahn, 2004).

LLU was implemented in 2003, when the FCC ordered PTOs to provide unbundled
access to a list of elements. Charges for interconnection, termination fees and
unbundled network elements are set at state level by the PUCs, although parties are
free to negotiate other rates, terms and conditions.

Besides the efforts of US regulators, the implementation of the Telecommunications
Act has not been easy. Incumbents legally challenged the new access rules in the
courts, arguing that they constituted regulatory expropriation. In May 2002, however,
the Supreme Court upheld the FCC's authority to choose LRIC as methodology to
estimate costs, although the uncertainty produced by these legal actions might explain
the minimal penetration of unbundled local loops discussed in the previous section. It
is also worth noting that despite the possibility of entering the long-distance market,
many local companies have decided not to open their markets for competition (Kahn,
2004). The Act did not establish any deadline to do so.

Universal service obligations in the US have traditionally aimed at providing
consumers in rural and insular areas with telecommunications services comparable to
those in urban areas. Carriers satisfying certain conditions can obtain support from the
federal Universal Service Fund, which is financed with contributions from all carriers
(including wireless) according to their revenues. Other universal service programs
subsidize the monthly telephone bills of low income customers and provide discounts
to schools, libraries and rural health care centers to connect to the internet (Jacobs,
1999).

7.5.3 Australia

Australia liberalized its telecommunications industry in 1997. At the time, only three
operators were licensed in Australia: Telstra (the incumbent), Optus and Vodafone.
The two former companies provided both fixed and mobile telephone services, while
Vodafone participated in the mobile market only.

The industry is regulated by the Australian Consumer and Competition Commission
(ACCC), which applies incentive regulation in the form of price-caps (using the RPI-X
methodology) to Telstra only. In establishing an access regime for the industry, the
government considered that the particular characteristics of telecommunications
require a regime that includes additional features to those contained in the National Access Regime, described in chapter 3. The access regime applicable to the telecommunications industry is contained in the Part XIC of the Trade Practices Act 1974. According to this legislation, only eligible services previously declared by the ACCC are covered by the regime. Eligible services are: i) a telecommunications service between two or more points, at least one of which is in Australia, or; ii) a service that facilitates supply of such a carriage service. The ACCC can declare eligible services in one of two ways: (i) in accordance with a recommendation of the industry self-regulatory body, or; (ii) after holding a public inquiry, which can be requested by an access seeker or anyone providing supporting information (Productivity Commission, 1997).

A service can be effectively declared to be subject to access regulation if such an arrangement is found to be in the long-term interest of the users. The ACCC is required to assess this issue using the following criteria:

a. The objective of promoting competition in the final market.

b. The objective of achieving universal connectivity in relation to services that involve telecommunications between final users.

c. The objective of encouraging the efficient use and the efficient investment in telecommunications infrastructure.

The ACCC may grant certain exemptions, but it can only revoke a declaration after a public inquiry has been held. Once a service has been declared, “standard access obligations” apply to those parties supplying the service, including the incumbent and non-dominant firms. These obligations address issues such as technical and operational standards, fault detection and rectification, accounts and access to services provided by specific customer equipment, such as set-top boxes for cable TV.

The standard access obligations do not specify neither the price nor many other terms and conditions for the supply of a declared service. These may be determined by three means:

a. Commercial negotiation.

b. An undertaking provided by the incumbent to the ACCC specifying the terms and conditions upon which it would provide access to a declared service.

c. Arbitration. When negotiations fail either party can request the ACCC arbitration of their dispute.

The ACCC may only accept undertakings if their terms promote the long-term interests of end users. If accepted, access undertakings are enforceable in the courts and serve
as the base for any subsequent arbitration of access disputes. In those cases where the ACCC is required to estimate access charges, the methodology used is LRIC.

The ACCC ordered the implementation of LLU in 1999. In 2002, it took a further step by mandating the incumbent to share its lines with competitors. This sharing allows specialized broadband operators to supply data services without providing voice services. The implementation of LLU in Australia was reached in a relatively short time and with little controversy, for which coordination between the ACCC and the industry proved to be useful. Almost all issues concerning the provision of technical resources and equipment interconnection were solved through bilateral negotiation between the incumbent and new entrants (Umino, 2003).

Universal services obligations in Australia are aimed at assuring that telecommunications services are reasonably accessible to all people, regardless of their location. These include standard telephone services, special services for the disabled and public telephony, among others. The policy is implemented and regulated at state level. The costs of providing these services are calculated annually, and all carriers are required to contribute in proportion to their revenues. In most states, telephone companies tender to provide these services (OECD, 2002a).

7.6 Lessons for a port access regime

It has been shown that ports and telecommunication networks share several economic characteristics, such as the presence of economies of scale, scope and density that create monopolies that need to be regulated. Public goods are also produced within these industries, and in both their supply can be limited to those that pay for them. A further similarity is that advantages inherited from the time when port and telephone services were provided solely by public operators limit the contestability of several markets of both industries.

A more important similarity is that services produced in both industries require complimentary inputs that may be produced by just one firm (the incumbent) or by several competing operators. Indeed, as explained in chapter 6, the completion of the port logistics chain requires, at least, the provision of several services feasible of being produced by only one firm or business group supplying integrated services (as it occurs in Matarani, Barranquilla or Cartagena) or by several competing companies. In telecommunications, long-distance, mobile, internet and data services can be produced by just one firm (the incumbent) or several competing operators. In both industries, the firm controlling a “key” facility (a monopolistic terminal or the local loop) will have the ability to exclude competitors and monopolize related markets unless access rules are applied.
Although reform in both industries involved privatization, it is still common to find public port authorities managing public-user ports. In telecommunications, however, privatization practically excluded the public sector from the provision of services.

An important lesson for port regulators is the verification that access regulation cannot rely on general antitrust law. As shown by the New Zealand experience, courts are unable to set access charges terms, which allows incumbents to delay entry by engaging in lengthy procedures. Moreover, many regulators in this industry have adopted policies to ensure that the incumbent does not delay negotiations, such as establishing deadlines for various stages of the process, reducing the number of topics subject to negotiation, or setting default interconnection agreements.

Table 7.4 presents the characteristics of the access regimes implemented in the UK, US and Australian telecommunications industries.

<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>US</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vertical structure</strong></td>
<td>Integration allowed for long-distance; legal separation required for mobile; LLU for internet services</td>
<td>Integration allowed for long-distance to incumbents opening their networks; legal separation required for mobile; LLU for internet services</td>
<td>Integration allowed for long-distance and mobile; LLU for internet services; sharing for data services</td>
</tr>
<tr>
<td><strong>Access pricing</strong></td>
<td>Set jointly by the regulator and the incumbent in consultation with other operators</td>
<td>Set at state level, although parties are free to negotiate other rates</td>
<td>Negotiated</td>
</tr>
<tr>
<td><strong>Non-price terms and conditions</strong></td>
<td>Regulated via “standard interconnect agreement”</td>
<td>Set at state level, but parties are free to renegotiate them.</td>
<td>Main ones regulated via “standard access obligations”. Others are negotiated.</td>
</tr>
<tr>
<td><strong>Mechanism to expand the infrastructure</strong></td>
<td>Universal service obligations</td>
<td>Universal service obligations</td>
<td>Universal service obligations</td>
</tr>
</tbody>
</table>

The first lesson for port regulators is that integration with open access constitutes a workable option to organize an industry. This is not a trivial result, considering the substantial market power that private incumbents enjoy, large enough to prompt regulators to intervene even in competitive markets (European Parliament, 2007); the discussed “consumer lock-in problem” (Lapuerta, Benavides and Jorge, 2003), and the slow penetration of competitors even after LLU policies have been implemented (Umino, 2003).

As for access pricing, negotiation seems to play an important role in this industry. In the UK, charges were developed through negotiation between the incumbent and the regulator, in consultation with operators and consumers. In the US, access prices are set at state level by the PUCs, although parties are free to negotiate other rates. In Australia, the determination of access prices is subject to a negotiation-arbitration framework consistent with the National Access Regime. The main lesson for port
regulators is that negotiation constitutes an efficient mechanism to set access charges, for which they should only intervene when negotiations are lengthy or parties are unlikely to reach an agreement. Regarding pricing methodologies, it has been seen that telecommunication regulators favor the use of incentive regulation through price-caps for retail prices (rates to be paid by final users) and cost-based methodologies (ECPR, FDC or LRIC) for setting access prices (rates to be paid by operators that need to use the infrastructure controlled by the monopolist). In ports, this would be similar to use price caps to set wharfage or berthing, but a cost-based methodology to set the rates to be charged to providers of nautical or ancillary services for the use of a monopolistic terminal’s infrastructure.

Port regulators should take note that it seems to be less room for negotiation in setting non-price access terms. Indeed, in the US basic non-price conditions are set by state regulators or by the Telecommunications Act, which requires access to be provided in at least four pre-determined points). Parties are free to negotiate only complementary terms. The case is similar in Australia, where some conditions are already set via “standard access obligations”. Others can be set by negotiation, arbitration or a pre-approved access undertaking. In the UK, non-price access conditions have been established in a “Standard Interconnect Agreement”, discussed by the incumbent and other operators and approved by the regulator. Further negotiation of these terms is discouraged. It is also important to highlight that in all countries analyzed, the assets that are considered essential facilities have been specifically described in the legislation establishing the country’s access regime. This facilitates negotiations and reduces regulatory risks. The lesson port regulators can draw is that even though the regulation asymmetry between them and regulated firms can be large, there are some access terms that can be set outright. The dilemma in this case does not seem to be between setting terms trough an administrative decision or by allowing parties to negotiate, but to determine which mechanism is better suited to determine a particular access term.

As for the set of incentives used to expand the infrastructure, it is customary in the industry to impose universal service obligations upon their incumbents to assure that services are provided to all users, including population segments with special needs and those located in high cost areas. In the port sector, this policy would be equivalent to the regulator mandating the terminal operator to expand the infrastructure when it reaches certain congestion levels.
8. Access regulation in the electricity supply industry

The electricity supply industry (ESI) is the second main network industry having experienced substantial restructuring since the early 1980s. Chile was the first country to introduce competition in electricity markets, and since then, many other countries have introduced market-oriented reforms in their respective industries (OECD/IEA, 2001).

The following sections describe the main characteristics of the ESI and how the access problem has been approached in different countries. The specific cases of UK, US and Australia are also analyzed.

8.1 Economic characteristics of the electricity supply industry

An electric system is made of interconnected generating plants, high-voltage transmission systems and low-voltage distribution networks that supply power to an entire area (see figure 8.1). The activities of an electric system are coordinated by an operator that can be independent or perform also other activities.
The next section will explain that reform requires activities formerly performed by and integrated company to be carried out by separate firms. However, the physical characteristics of electricity and those of its supply and demand make these activities difficult to coordinate (OECD, 2003). One important physical characteristic of electricity is that it cannot be stored. Indeed, an electric system has to be physically balanced all the time; i.e., supply has to match demand continually. However, the demand for electricity is seasonal, cyclical and subject to random variations. These features, and the fact that electricity has to be supplied at a constant voltage to be considered reliable, imply that certain capacity should be maintained in reserve to meet surges in demand (This is an economic characteristic that the ESI and the port industry have in common, and one that has important implications for the pricing of electricity and port services.

In the ESI, the continuous balance of the system is achieved by having few generating plants delivering a base load when the demand is low and having the remaining ones in stand-by. As demand increases, plants in stand-by are gradually asked to dispatch electricity by the system operator (OECD/IEA, 2001). Moreover, congestion in the transmission network may impede generators with idle capacity to dispatch electricity when needed, thus making the short-term supply of electricity highly inelastic (as in ports). However, since electricity travels at high speed, close and centralized coordination may reduce short-time variability and allow the full utilization of the installed capacity.

Similarly to port facilities, the supply of electricity shows economies of scale. Large networks have fewer requirements to maintain reserve capacity. Indeed, since different customer types (industrial, commercial and residential) have different demand patterns, a larger network allows an increase in demand from one type of

Figure 8.1: A typical electricity network
customer to be compensated with a reduction of demand from other, thus reducing the need for reserve capacity and diminishing operational costs. The presence of externalities also reinforces the natural monopoly characteristics of the ESI. Electricity does not flow along a defined path, but will divide and flow over all possible paths in quantities that are inversely proportional to the resistance of the path. Therefore, the ability of one plant to transmit electricity between two points of the system will depend on the activities of the remaining plants transmitting at the same time. Moreover, failures in any node or link of the network, such as line congestion or the shut down of a generation plant, affect the operation of the whole system. These characteristics require a centralized management of the system. Centralized coordination between generation and transmission also results in further economies of scale, such as the following ones (OECD, 2001):

a. Relatively small demands can be aggregated to take advantage of economies of scale at plant level.

b. The aggregation of segments with different consumption patterns reduces demand variability, thus allowing a more efficient use of installed capacity.

c. The requirements of capacity to be kept in reserve are lower in an integrated system than in the case of several systems operating independently.

d. Integrated coordination allows the full use of cheaper generation technologies before the more expensive ones are deployed.

8.2 Markets and services

Electricity supply entails three main activities: generation, transmission and distribution. In later years, the activity of retailing has also arisen as a consequence of reforms in the ESI. Two of these activities (transmission and distribution), possess characteristics of a natural monopoly; while the remaining two (generation and retailing) are typically considered competitive. As in ports, all of these markets produce inputs that are required to deliver electricity to industries and households; i.e., to produce one unit of the industry’s product. Controlling one market would allow an integrated company to gain control of the other markets.

8.2.1 Generation

Generation involves the transformation of a different form of energy into electricity. Electricity can be produced by burning oil, natural gas and coal; or using nuclear power, hydro power or wind turbines, among others. Fossil fuels and nuclear power produce steam to propel electricity-generation turbines. Falling water and wind do the same in the remaining technologies.
Each of these technologies has a different cost structure, as shown in table 8.1. Nuclear and hydro technologies have large capital costs and low variable ones. Indeed, the use of nuclear plants involves long construction times and large decommissioning costs, but operational costs are relatively low and constant during the economic life of the plant. Hydro power typically requires large investments and long construction times, but their variable costs are almost negligible. On the other hand, the use of oil and gas involves lower capital costs than nuclear and hydro power, but their variable costs are subject to variations in the market price of these fuels, which can be substantial. The use of coal is also facing rising environmental opposition, and its use is increasingly restricted.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Capital costs</th>
<th>Operational costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydro power</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Nuclear power</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Wind power</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Oil-fired</td>
<td>Low</td>
<td>Variable (subject to market price variability)</td>
</tr>
<tr>
<td>Gas-fired</td>
<td>Low</td>
<td>Variable (subject to market prices)</td>
</tr>
<tr>
<td>Coal-fired</td>
<td>Low</td>
<td>Low (but increasing environmental opposition)</td>
</tr>
</tbody>
</table>

Source: Steiner (2000)

The different cost structures of generation technologies allow plants to operate according to a least-cost merit order. As explained before, physical balance of an electric system is achieved by having few generating plants delivering a base load and leaving the remaining ones in stand-by until they are needed. As demand increases, most electricity networks use nuclear and hydro power as base load, dispatching further generators in a merit order based on their variable cost.

It is important to mention that generation plants sell both capacity and energy. Capacity is the guarantee that a client can demand electricity up to certain level and is measured in watts (W). Energy is the electricity flow and is measured in watts per hour (Wh).

8.2.2 Transmission

Transmission refers to the transport of electricity from the point where generators deliver it to the network, to the point (called “node”) where distributors withdraw it. A transmission network can be constituted by several interconnected lines or sub-
networks, owned by one or several companies. Since electricity losses increase with the distance and decrease with the voltage at which it is transmitted, electricity is transported at high-voltages.

Traditionally, the operation of the system has been considered a transmission activity. System operation involves the management of the system to coordinate activities between generation, transmission and distribution companies; maintain the voltage and prevent breakdowns. As it will be seen later, some reform options imply the unbundling of system operation from transmission activities.

Transmission presents sizeable economies of scale. According to OECD (2001), the cost per MW per mile of a 765kV transmission line is at least 30% less than a 500kV line, and 85% less that a 138kV line. Other studies suggest that transmission capacity increases 1.6 times the cost of transmitting one extra MW (Hjalmarsson, 1996). Furthermore, since electricity losses decrease with the voltage at which it is transported, higher-capacity transmission lines suffer less losses than lower-capacity ones, thus strengthening the activity’s economies of scale.

Transmission charges are typically structured as two-part tariffs, with a fixed component for capacity (kw per period of time) and a variable one (wh) for energy. Since the initial investments constitute the larger part of the total costs, the fixed one is usually the largest component of the tariff (Rothwell and Gomez, 2003).

Two are the main methodologies used to estimate the fixed component of transmission charges: “postage stamp” and network use. Under postage stamp, all fixed costs are allocated among the users using a simple parameter, such as the installed capacity of a generator or the contracted capacity of a distributor; independently of their actual use of the network. Under the second methodology, costs are apportioned to the users proportionally to their physical use if the system, which can be, in turn, estimated using several alternative methods (De La Cruz and García, 2003).

There are several approaches to deal with network congestion. The first is called “zonal pricing” and divides the network into several separate zones. Each zone represents a different market and transmission prices are set to reflect the price-difference between zones. “Nodal pricing” is a variant of this approach in which prices are calculated for each point of connection. Alternatively, system operators may auction available capacity, granting tradable capacity rights. This is the option used in Norway and in several US regional markets. Another approach is “counter trading”. In a counter-trade system there is only one market. If market participants require more transmission capacity than what is available, the system operator has to pay operators to free capacity. The financial loss that this transaction implies is meant to encourage the transmission owner to enlarge the capacity of the network. This is the system used in the UK (OECD, 2003).
Many countries (several Latin American among them) do not require their owners to expand the transmission network. It is usually proposed and financed by users with the approval of the regulator, the system operator or both (Fischer and Serra, 2000). In countries where the system is still operated by an integrated monopoly, such as Mexico or France, central planning is still used to determine network expansion.

8.2.3 Distribution

Distribution refers to the transport of electricity, from the node where it is withdrawn from the transmission system, to the place where is finally consumed. This activity is performed over high-voltage, medium voltage and low voltage systems.

Distribution is also considered a natural monopoly, due to the presence of considerable economies of scale and density, as well as of sunk costs. One of the sources of economies of scale are the costs of obtaining rights-of-way, which represent a large part of the initial investments and are insensitive to the amount of energy distributed by the network. A right-of-way, in this context, is a strip of land that a distribution company uses to build an electricity line. It usually allows companies to clear the land from trees, buildings and other structures that could interfere with the operation of the network (Rothwell and Gomez, 2003).

In the presence of idle capacity, both transmission and distribution present economies of density, since it is more efficient to transmit using existing lines or to connect a new user to the existing wire system, than building new grids. Idle capacity is likely to arise in these activities because of indivisibilities in the scale of the plant and due to the use of standardized voltages. Furthermore, both activities present large sunk costs due to the low residual value of many of their investments, such as civil works, wires and towers, for example.

8.2.4 Retailing

Retailing is the purchase of electricity from generators and re-sale to final users. It involves metering, marketing and billing. Before reform, this activity was not performed separately from distribution. When retailing is allowed, retailers buy electricity at the exchange (at wholesale prices) and sell it (at retail prices) to final customers. Competition among retailers keeps prices close to costs (Hunt, 2002).

Retailing can be, however, subject to regulation when customers under certain consumption threshold are not allowed to choose their supplier, as in Belgium or Ireland. In these cases, users buy electricity at regulated prices set by the regulator. Even though vertical separation of retailing from distribution is usually enforced, in some countries, as the UK, for example; generators are allowed to perform retailing.
8.3 Reform and regulation in the electricity supply industry

Electricity is used as input for the production of most goods and services, and its importance for the working of an economy is evident. However, due to the characteristics of electricity discussed in the previous section, its production requires constant coordination and monitoring to ensure the reliability of the system. In almost every country, this coordination is achieved by delegating the exclusive supply of electricity to integrated monopolies.

The argument for an integrated monopoly was strengthened by the large economies of scale and density present in the generation, transmission and distribution of electricity. Under this system, a single electricity company owned and operated all generation and transmission assets, also acting as system operator. With the notable exceptions of the US, Japan, Germany and Spain, all of these companies were publicly owned. Distribution networks could also be owned by this company or others, usually municipally-owned. Since competition was absent, prices under this organizational arrangement were set through centrally-planned processes.

During the early 1980s, this traditional structure began to change. According to Kessides (2004), the main forces behind the reform of the ESI differed between developed and developing countries. In developed countries, the oil crises of the 1970s induce them to try reducing their dependence on oil for electricity generation, which led to invest in alternative and expensive-to-build options such as nuclear power and large coal-fired generation plants. However, budget pressures and attempts by governments to contain inflation reduced electricity revenues, delayed necessary investments and undermined the public’s confidence in state-owned integrated utilities. Moreover, the increasing availability of cheap gas in Western Europe and the US and the development of combined-cycle gas turbines further reduced the minimal scale of generation plants, weakening the natural-monopoly argument. In this context, well-integrated electricity networks, an array of power plants and excess capacity made competition in generation feasible and attractive. In developing countries, the main driver for reform has been the poor performance of public-run electricity companies, in a context of increasing demand. Political considerations tended to push prices down, impeding the adequate maintenance and expansion of the network. Inadequate coverage and unreliable supply were considered one on the main hindrances for industrial development.

The reforms implemented in the industry aimed at creating a market for electricity. This is achieved by fostering competition among generation companies and allowing electricity prices to be set by supply and demand interaction. The final price for electricity would then be the exchange (market) price plus the cost of using (regulated) transmission and distribution networks. However, since both networks are
needed to deliver electricity to final users and they are almost universally naturally monopolistic, regulation of these activities is required (Chao and Wilson, 1999).

It is important to note that open access is essential for this model to work (Newbery, 2001a). Indeed, if network owners are integrated and have the ability to discriminate against competitors in the generation market, they might try to do so by establishing expensive access terms or by denying access to new generation plants. As explained before, an open access regime limits integrated network owners to recoup profits foregone by regulation.

Although specific measures differ from country to country, the typical electric system arising after reform has the following characteristics (Guasch and Spiller, 1999; Steiner, 2000; OECD/IEA, 2001):

a. The industry has been vertically separated\(^{46}\) into generation, transmission, and distribution activities. Some countries allow retailing as a separate activity. Vertical and horizontal integration is limited.

b. Electricity is generated by competing companies. Each company can own one or several generation plants.

c. The transmission network remains centrally operated, even though different companies may own parts of it. An open-access regime is implemented to allow new generation capacity to connect to the network and to promote its expansion.

d. Distribution networks have been allocated to franchised companies\(^{47}\), each covering a determined area. Since these networks do not overlap, no competition occurs at this level.

e. A system operator has been appointed to coordinate the activities of the electric system. Depending on the reform model, the appointed system operator can be transmission owner or an entirely independent body.

f. Prices are set in an electricity exchange, created to allow interaction between supply (generators) and demand (distribution or retailing companies). The rules of how these exchanges work vary from country to country. Section 8.4 shows how they work in UK, US and Australia.

g. A regulatory authority has been created to oversee the electricity exchange, regulate prices for transmission and distribution and supervise the functioning of access regimes.

\(^{46}\) Either operationally or legally separated (see section 3.1)

\(^{47}\) Through privatization or not
h. Large users are allowed to buy electricity directly from the generators at market rates. Electricity prices for small users are typically regulated.

i. Some countries have further separated the functions of distribution network owner and electricity retailer.

It is worth noting that even though reform has implied some form of privatization, it has not been necessarily the case. In the US, for example, private utilities were allowed before reform and some companies remain municipal or state-owned afterwards.

8.4 Access arrangements in the electricity supply industry

Three have been the basic models under which the ESI have been restructured: single buyer, wholesale and retail competition. They imply different access arrangements.

Under the single-buyer model, a single agency buys electricity from competing generators and sells it to distribution companies and large users. In this model, the price at which the single buyer sells electricity is the price of purchase plus transmission costs, which are regulated. The single buyer can be an integrated company, a transmission company acting as system operator, or an independent system operator. In its extreme form, the single buyer buys all energy in the market and has the monopoly to sell it, deciding how much and from whom to buy. Another option is the single buyer acting only as the aggregator of demands from final users, not deciding how much of from whom to buy. This is the model adopted by countries such as Indonesia, Pakistan, Thailand and Hungary. Under this system, access price and non-price terms are regulated (Lovei, 2000).

Under wholesale competition, distribution companies buy electricity from competing generators, use the transmission network (paying the regulated price) and deliver it to their franchised areas. Transmission networks operate under open access arrangements and only final users consuming over certain threshold are able to contract directly with generators (Bacon and Besant-Jones, 2002).

Under retail competition, independent companies are allowed to buy electricity from competing generators and deliver it directly to final users, using both transmission and distribution networks owned by other companies. Unlike the previous case, distribution companies only maintain the network, since users are retailer’s clients. In this model, both transmission and distribution networks operate under open access arrangements (OECD/IEA, 2001).

Although there have been a debate on the costs and benefits of introducing retail competition (Newbery, 2001a), countries such as the UK and Australia have already
implemented it. In the US, where distribution regulation lies within state regulators, the degree of retail competition varies from state to state. In the EU, Directive 2003/54/EC establishes retail competition for all consumers not later than July 2007.

Some countries, such as Argentina, Chile, Brazil, Australia and some states in the US, have further unbundled their ESIs separating the function of system operator from transmission ownership. The aim of this policy is to avoid transmission operators managing the system in a way that favor their own commercial interests. For example, they might try to avoid future congestion by pursuing transmission-based solutions (such as building new transmission lines) even if less costly alternatives exist, such as building generation in another location. Critics of this arrangement, however, claim that it is inherently inefficient, since independent system operators are non-profit organizations with complex governance structures which do not bear the costs of their decisions (Arizu, Dunn and Tenenbaum, 2001).

It can be seen that all three restructuring models require open access to transmission or distribution grids. Under the single-buyer model, both price and non-price terms for access are explicitly regulated. Under the two latter models, wholesale and retail competition, some or all of the terms, as well as the access charges might be subject to negotiation.

8.5 Reform, regulation and access regimes in selected countries

8.5.1 United Kingdom

The reform of the UK’s electricity system followed different paths in each of its constituent countries: England, Wales, Scotland and Northern Ireland. In Scotland, the electricity system was privatized as integrated regulated utilities in 1991 (with the exception of two nuclear plants privatized in 1996). Nowadays, there are two vertically integrated companies that compete for customers. Their activities are regulated. In Northern Ireland, generation plants were vertically separated from transmission, distribution and retailing, which remain as a regulated monopoly. All of these companies were privatized by 1993. In England and Wales, the restructuring of the ESI involved the vertical separation and privatization of the state-owned generation, transmission and distribution assets and the creation of a power exchange, the Electricity Pool.

The Electricity Act of 1989 divided the generation plants of the state-owned Central Electricity Generating Board of England and Wales, into three companies: National Power, with 40 generation plants; PowerGen, with 23 generation plants; and, Nuclear Electric, with 12 nuclear plants. All companies but Nuclear Electric were privatized by 1991. The most modern nuclear facilities were privatized in 1996 and the remaining ones still remain state-owned to date.
Electricity generation comprises more than forty generation companies, including several owning just one plant. Transmission in Great Britain remains a monopoly, where the grid located in England and Wales is owned by the National Grid Electricity Transmission (NGET). NGET also fulfils the function of system operator. Two other companies own the lines located in northern and southern Scotland, respectively. The transmission network of Great Britain is shown in figure 8.2.

![Figure 8.2: Great Britain’s Transmission Network](image)

Distribution also remains a monopolistic activity. Seven distribution companies operate twelve separate areas. These companies hold a license whose terms establish their obligations.

The Electricity Pool was an exchange created for trading electricity in England and Wales. Between 1990 and 2001, it allowed retailers and large users to buy electricity from generators. The pool operated by the transmission operator as a daily, day-ahead, sealed bid auction. Every morning, generators had to declare which of their generating plants were to be available the next day, and announced five prices that will hold for the following day. At the same time, retailers submitted their demand

---

48 Source: IEA (2002a)
estimates. The pool determined which generation plants were to be used to meet expected demand. The price was determined by the most expensive generator used during each half-hour slot and applied to all dispatched generators. Generators in stand-by received a capacity payment according to the degree to which capacity was needed (Littlechild, 2000).

This system design had the main flaw of facilitating collusion. Indeed, even though there were more than 40 generation companies, two of them (PowerGen and National Power) set the price most of the time. The problem subsided despite the regulator asked both companies to divest some generation capacity in 1993. As a consequence, Offer regulator was forced to set a price cap between 1994 and 1996. Another important problem was the governance structure of the pool. Voting rights were distributed among generators and retailers according to their market shares and changes had to be approved by 65% of the voting rights. This structure limited the regulator’s powers and thus reinforced the status quo (Newbery, 1999).

In 2000, through the Utilities Act, the government changed radically the exchange rules by introducing NETA (New Electricity Trading Arrangements), which replaced the pool in March 2001. The Act also merged the regulators of electricity and gas into Ofgem, the Office of Gas and Electricity Markets. The main feature of NETA is that electricity trading is no longer done through a centralized market. Electricity prices are negotiated directly between buyer and seller, for which there is no single price. In 2005, the government extended these arrangements to include generation plants located in Scotland; creating an integrated market in the whole Great Britain. This measure, however, also increased demand for new transmission infrastructure (Parker, 2000).

Transmission is operated under an open access regime, for which any new generation plant can connect to the system subject to capacity availability. The main non-price terms of access are also regulated and set in the “Connection and Use of System Code”. These are supplemented by bilateral agreements which establish the works required for the connection to be possible. New users are obliged to pay a transmission connection charge, bearing the direct costs of connecting to the transmission network. This charge is negotiated between the parties.

The method used to set access prices and non-price terms has changed through the years. Nowadays, the revenues of both transmission and distribution network companies are set by the regulator using the RPI-X methodology for a five-year period. These controls restrict the amount of money companies can earn from their regulated businesses. Transmission charges are set annually by the regulator using the LRMC methodology and paid by both generation and distribution companies (Ofgem, 2007).

The Utility Act of 2000 separated completely the functions of retailer and distribution network owner. Nevertheless, no vertical separation is enforced between retailing and generation, and in fact several generators have entered the retailing market. Retailers are allowed to operate nationwide, paying distributors for the use of their networks.
To manage short-term congestion, NGET operates a balancing mechanism based on the difference between contractual positions and NGET’s own demand forecasts. It provides financial incentives to all market participants to avoid congestion. Incentive schemes are given each year to NGET to perform its role as system operator in an efficient and economic manner. If actual costs are below the target NGET is expected to achieve, the company receives a payment. Otherwise, it receives a penalty. This incentive scheme is designed by the regulator in consultation with other industry parties.

Network expansion plans of transmission and distribution companies are scrutinized by the regulator at the beginning of each price review. The price caps for the following periods take into account the projected investments and the regulator remunerates only those that are considered prudent, at the estimated cost of capital. According to Parker (2002), this methodology is nonetheless prone to regulatory failures given the information asymmetries between the regulator and the regulated firm.

8.5.2 United States

The electric industry in the US is vast, diverse and geographically segmented. According to the International Energy Agency, it accounts for about 10% of the physical capital investment of the country, and has annual sales exceeding US$200 billion (IEA, 2002b).

There are three main interconnected electricity systems in the US, comprising both transmission networks and pools of generation plants: the Eastern Interconnected System, comprising two-thirds of the states; the Western Interconnected System, comprising the states west of the Rocky Mountains; and the Texas Interconnected System, comprising mostly the state of Texas.

Regulatory powers in this industry are shared among federal, state, and some municipal regulators. The regulator at national level is the Federal Energy Regulatory Commission (Ferc). The Ferc has authority over all privately-owned lines involved in interstate transmission. Public Utility Commissions (PUCs), at state level, have jurisdiction over local generation transmission and distribution assets, but usually not over municipal companies.

Participants in the US electricity industry can be grouped into five broad categories:

a. Vertically integrated, privately-owned utilities;

b. State and municipal utilities and authorities. Due to their public-owned status, they are not regulated;
c. Federally-owned utilities that generate and transmit electricity;

d. Privately-owned generators not affiliated to an integrated utility; and,

e. Retailers, called “brokers” or “marketers” in the US.

Since the early 1990s, the US government has been implementing reforms aimed to increase competition in the ESI. In 1992, the Congress passed the “Federal Power and Energy Policy Act”, which promoted open and non-discriminatory access to the transmission network. Further steps toward the adoption of an open-access regime were taken by the Ferc in 1996, with the issue of orders 888 and 889 (IEA, 2002b).

Order 888 required all companies owning transmission assets to publish open-access tariffs separately from others. Transmission tariffs would be set by the transmission owner, but according to cost-based methodologies and taking into account a set of minimum terms and conditions indicated by the Order. Since the Ferc does not have legal powers to order divestiture of privately-owned companies, Order 888 could only mandate operational separation. Nevertheless, it encouraged the formation of “independent system operators” (ISOs). ISOs are formed by companies in a particular region which transfer control (but not ownership) of their transmission assets to an ISO responsible for managing the electricity exchange and the transmission network. It is important to mention that even though the formation of ISOs was encouraged, its adoption was not mandatory. To be authorized, ISOs must satisfy a series of requirements to assure their autonomy, such as fair and non-discriminatory governance, independence from financial interests of participants, open-access under one tariff, among others. ISOs do not necessarily have responsibility for the expansion of the transmission system (Brown and Sedano, 2003).

Order 889 aimed to standardize how operationally-separated utilities should carry on transmission and other activities. It established the “Open-Access Same-Time Information System” that regulate the disclosure of information that transmission companies must provide to market participants. Operational unbundling, nonetheless, seemed not to be working. Complaints of discriminatory behavior among integrated utilities persisted. Furthermore, according to Arizu, Dunn and Tenenbaum (2001), the Ferc concluded in 1999 that operational unbundling was “inefficient, unfair, and difficult to enforce”. Moreover, some state PUCs, such as California’s and Arizona’s, opted for providing financial incentives to integrated utilities to divest their generation assets.

In 2000 the commission issued Order 2000, encouraging the formation of “regional transmission organizations” (RTOs), electricity systems covering a wide area with a governance structure similar to that of the ISOs. Both ISOs and RTOs are called “organized markets”. They operate day-ahead and real-time electricity exchanges. Up to date, there are five organized markets: ISO-NE (New England region), NYISO (New York Independent System Operator), MISO (Midcontinent Independent System Operator), SPP (Southwest Power Pool), and CAISO (California Independent System Operator).
York State), PJM (Pennsylvania, New Jersey, and Maryland) ERCOT (most of Texas) and CAISO (most of California). Their location can be seen in figure 8.3.

Figure 8.3: Organized Markets in the United States

Order 888 and 889 were modified by Order 890, which modified the minimum access terms and conditions indicated in Order 888 and required more information to be disclosed under the system devised in Order 889. This reform was necessary because opportunities for discrimination continued to exist even though operational separation was imposed. For this reason, the Ferc issued in 2007 a new order establishing financial incentives to encourage investment in and expansion of transmission networks. This order gives incentive rates of return on equity for new investment, higher rates of return for companies that are members of ISOs or RTOs, accelerated depreciation, and deferred income taxes, among others.

Since regulation of distribution and supply lies within the jurisdiction of state regulators, the introduction of retail competition in the US has not been uniform. States accounting for approximately half of the population have implemented some degree of retail competition, but the process is underway at both federal and state levels. Some states have decided not to implement retail competition based on studies that show that marker power might arise due to high-concentration levels in transmission (OECD, 2003).

8.5.3 Australia

Although Australia began restructuring its ESI in the early 1990s, an interstate market for electricity was not established until 1998. As in the previously analyzed countries, the reform also involved the vertical separation of integrated companies, the
implementation of open-access transmission, and the gradual introduction of retail competition.

The Australian electricity system is not integrated country-wide due to its physical and demographical characteristics: long distances, scarce population concentrated in the east and southwest of the country, few large cities, etc. (see in figure 8.4.).

Before 1990, each Australian state owned its own electric system, either through an integrated utility or a combination of several agencies. Decisions about price and network expansions were made by governments through these utilities. Interstate trade was limited and regulatory powers lay within state regulators. Reform started in 1990-91, when the executive branches of the commonwealth and state governments approved plans to develop a single wholesale electricity market in eastern and southern Australia. These plans involved the restructuring of the ESI and the development of a new regulatory framework over a period of eight years.

In 1993-94, a council created to devise the development of an interstate electricity market proposed the National Electricity Code (the Code), consistent with the National Access Regime described in chapter 3. The Code, enforced by an ad-hoc administrator51, was adopted by the states participating in the NEM in 1996. The Code contains three separate but related elements:

50 Source: IEA (2001)
51 The National Electricity Code Administrator
a. Established market rules governing how electricity buyers and sellers trade and how the system is operated.

b. An access code that establishes non-price access terms and delineates principles for setting access prices.

c. Administrative arrangements that state how disputes are settled and how the Code is changed, among others.

The Code also established that the National Electricity Market Management Company (NEMMCO) will be the system operator.

Simultaneously, the Australian ESI was vertically separated in all jurisdictions participating in the NEM. Generation assets were further divided into competing business and “ring-fencing” arrangements were implemented to prevent information flows and discriminatory behavior. The industry was completely privatized in Victoria in 1994, but remains publicly owned in Tasmania, Queensland and New South Wales. The South Australian ESI was concessioned to the private sector in 2000. States also handed over some of their regulatory roles to independent regulators. The Australian Energy Regulator (AER) is the national agency regulating transmission and wholesale markets, while state agencies regulate distribution and supply.

In 1998, an electricity exchange was created called the National Electricity Market (NEM), operated by NEMMCO. Although it is called “national”, it only comprises eastern and southern Australia: Queensland, South Australia, New South Wales, Victoria, the Australian Capital Territory and Tasmania. There are additional regional systems in the remaining states. The NEM accounted for 93% of total Australian generation in 2002 (ACCC, 2004).

The NEM is a mandatory auction market in which generators of more than 30 MW and customers trade electricity. Generators submit bids specifying the amounts they are willing to supply at certain prices. These bids are used by NEMMCO to build a dispatch schedule, although scheduling can be constrained by congestion in the interconnectors linking two regions, thus prompting inter-regional price differences. Prices for each region are calculated *ex post*. There is a cap on spot prices.

Prices of transmission and distribution activities are set by the regulator through price-caps estimated using the RPI-X methodology. Charges are paid entirely by final users. The allocation of these charges among consumer groups varies among states. In South Australia, for example, transmission charges are averaged among all users, while in Victoria it depends on the location of buyers and sellers.

52 This function was performed by the ACCC until July 2005, when the AER was created to regulate both electricity and energy markets. The AER is a constituent part of the ACCC, but a separate legal entity.
Transmission companies also receive revenues from congestion, although in an indirect way. When congestion occurs, the price in the importing region is higher than in the exporting region and part of the surplus revenues generated by these differences are used to reduce the transmission costs paid by final users (Gans and King, 2000). In Australia, the owners of the transmission assets in each region of the NEM are also the network planners, with the exception of Victoria. In this state, a separate body devises the plans for network expansion that the owners of transmission assets will have to undertake.

In order to ensure that transmission or distribution network expansions are effectively needed, the investments will only be remunerated through regulated tariffs if they pass a “regulatory test”, consisting of the following:

a. If the new investment minimizes the net present value of the cost of meeting the service standards set out in the Code; or,

b. If the net present value of the profit generated by the expansion is the highest among a number of alternative projects, timings and market development scenarios.

Alternatively, investors can undertake unregulated investments in transmission assets (IEA, 2001). Unregulated investments are made at the investor’s risk and there are no restrictions in network expansions within a region.

Retail competition has been progressively introduced the states that conform the NEM since 1994. Although the pace and thresholds were different among states, all end-users are able to choose their electricity retailer since January 2003 (ACCC, 2004).

8.6 Lessons for a port access regime

It can be seen that several features of the ESI, such as the need to maintain the system balanced continually, make the organization structure of this industry unique. Electricity networks share with ports the characteristics of economies of scale, presenting externalities and producing public goods. These characteristics, among others, were used as arguments for allowing only integrated utilities to supply electricity.

Another characteristic of the ESI shared with ports is that the delivery of electricity to final users requires inputs produced in several markets. Moreover, some of these markets (transmission and distribution) are monopolistic and the remaining ones (generation and retailing) competitive. If left unregulated, monopolists (the owners of the transmission or distribution grids) will seek to obtain rents by restricting access to their networks (thus restricting supply) or by integrating up- or downstream and then discriminating against non-related companies. In the maritime industry, as explained
in Chapter 5, the completion of the logistics chain also requires activities that are produced sequentially and in several markets. A port terminal operator that holds an unregulated monopoly would obtain economic rents by restricting access; or, if it is an integrated monopolist, by granting access only to related companies.

The are also similarities in the reform path followed in the ESI and in the port industry. Indeed, reform did not exclude totally the public sector from the provision of services. In the US, for example, while transmission assets are mostly privately-owned, generation, and distribution present a mix of private and public ownership at federal, state and municipal levels. The ESI also remains state-owned in several Australian states. As explained in section 5.1.2, port services are provided under several organizational models. According to the classification made by Baird (1999), only one out of four organizational models totally excludes the public sector, and most ports are organized according to one of the remaining ones. Moreover, in many ports of the EU, the public sector continues to provide port services directly (European Commission, 2006).

As explained in chapter 5, port reform was driven by the need to adapt the industry’s organization to new economic and technical developments, such as globalization and containerization. In the case of the ESI, reform drivers were similar. Technological advances, such as the development of combined-cycle turbines, and market developments, such as the availability of cheap gas; weakened the natural-monopoly argument and made competition possible, at least in generation.

Table 8.2 shows the main characteristics of access regimes in the ESIs of UK, US and Australia.

<table>
<thead>
<tr>
<th>Table 8.2: Main characteristics of ESI access regimes</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
</tr>
<tr>
<td>Vertical structure</td>
</tr>
<tr>
<td>Access pricing</td>
</tr>
<tr>
<td>Non-price terms and conditions</td>
</tr>
<tr>
<td>Mechanism to expand the infrastructure</td>
</tr>
</tbody>
</table>
One can recall from section 3.1 that deciding over the vertical structure of an industry implied a trade-off between allowing some discrimination to occur (by allowing integration) and increasing transaction costs (by ordering separation). From the structures adopted after reform in the analyzed countries, one can draw the conclusion that regulators in the ESI consider that discrimination is very likely to occur, since vertical separation seems to be the norm. Indeed, as shown in table 8.2, the UK’s transmission owner is not allowed to participate in the remaining activities even though integration is allowed between generation and retailing. In the US, integration is only allowed because regulators lack of powers to order divestiture; and in Australia integration only subsists among public-owned utilities. This trend is probably related to the fact that reforms in the ESI were aimed at creating markets for electricity that require vertically-separated utilities to operate properly. Indeed, as shown by the US experience, any arrangement short of a complete vertical separation seems to be ineffective in preventing discrimination from occurring (Arizu, Dunn and Tenenbaum, 2001).

The first lesson for port regulators is that, when deciding over vertical structure of the industry, regulators need to compare the expected costs of discrimination with those of ordering vertical separation. As the experience in the ESI shows, it might be preferable to incur extra transaction costs than allowing discrimination to occur.

One can also see from table 8.2 that the room for negotiated access prices is small in the ESI. Indeed, negotiation is used in a single country (UK), and only to determine one-time charges to be paid for connecting to the transmission network.

As for the methodologies used to price access, lessons for ports are not as clear as in the telecommunications industry. As one can recall from the previous chapter, in that industry retail prices are estimated using price caps while access prices are set using cost-based methodologies. In the ESI, however, methodologies vary according to the country. In the UK, transmission charges are calculated using LRMC (a cost-based methodology). RPI-X is complementarily used to cap overall revenues for transmission and distribution companies. In the US, the methodology ordered by the Ferc is cost-based, while in Australia the regulator uses price caps.

Although there is not information regarding the criteria used by lawmakers and regulators to decide over using price caps or cost-based methodologies, these decisions were probably linked to the incentives given to the regulated companies by each pricing methodology. As discussed in chapter 2, the most effective regulation schemes are those, such as price caps, that reward regulated firms €1 for each €1 they save (Laffont and Tirole, 1993). Under this scheme, the regulated firm enjoys the cost reductions until the next price review. For this mechanism to work, however, prices have to be set for periods long enough to incentive firms to reduce costs. For this reason, it does not seem well suited for open access regimes designed to incentive potential entrants to negotiate their access terms (like the British one). The lesson for ports is that price caps seems to be better suited for setting access prices in regulated
regimes, while cost-based methodologies such as FDC or LRIC seem to be better for regimes that favor the use of negotiation for pricing access.

The lesson for ports regarding non-price access conditions is that, unlike telecommunications, the mechanisms to determine them are similar to those used to determine access pricing. Indeed, the UK uses an intermediate approach: main terms are regulated (set in the “Connection and Use of System Code”), but negotiation is used for setting complementary conditions. In the US, regulation is used for integrated utilities but unbundled ones (“Organized Markets”) can set their own following guidelines set by the regulator. In Australia, non-price terms and conditions for access are regulated (set in the National Electricity Code).

As for the mechanism to expand the infrastructure, one can see that the approach in the ESI seem to be an intermediate between regulation and market incentives. Indeed, in both UK and Australia, even though expansions are proposed by the asset’s owner (arguably following market incentives), these need to be approved by the regulator. In reformed US utilities (Organized Markets), system operators act as planners, ordering expansions when they consider are needed.

There are two main lessons for port regulators concerning the mechanism to expand the infrastructure. The first is that the decision regarding the expansion of the infrastructure cannot be left to the monopolist. As discussed in chapter 1, unregulated monopolists have incentives to create artificial scarcity by not expanding the infrastructure, while regulated ones have incentives to invest excessively (to earn returns over capital investments). The second lesson regards the two options regulators seem to have to deal with this problem, each with important shortcomings. The first is to allow an independent planner to decide over infrastructure expansions, as in the US ESI. Arguably, independent system operators have enough information to make opportune and demand-adjusted decisions, but, as discussed in section 8.4, they may lack of incentives to make the right decisions (they do not bear the cost of their mistakes) and tend to have complex governance structures. The second option is to allow the asset owner to propose expansions but only approve those that are considered necessary, as in UK and Australia. As argued by Parker (2002), this option is prone to regulatory failures given the information asymmetries between the regulator and the regulated firm.

Although the need to create a market is unique to this and the natural gas industries, important lessons for ports could be drawn from this experience. It could be seen that even though the approaches are similar, they vary from country to country. The UK was the first developed country to restructure its ESI, and as such, was also the first to suffer the consequences of a flawed regulatory design. As corroborated later, privatization failed to create enough rivalry among generators, and errors in the design of the electricity market allowed some of them to obtain economic rents by exercising their market power. Moreover, flaws in the design of the market’s governance structure further limited the ability of the regulator to introduce the necessary changes. In the US, the fact that most integrated utilities were already
private and that regulators lack authority to order their divestiture made the reform difficult to carry on. As explained, complaints of discriminatory behavior from integrated utilities persisted even after the implementation of an open-access regime. In this context, the only chance to regulators to introduce adequate levels of competition was to encourage the formation of “organized markets” such as ISOs and RTOs, even by giving financial incentives.

Australia, on the other hand, reformed its ESI after the other analyzed countries and was able to reap from their experiences. The design of the NEM, for example, is similar to that of the England and Wales pool; but measures were taken to avoid market power abuse, the feature that forced UK regulators to replace the pool with NETA. As in the US, the Australian ESI is regulated at federal and state levels, but since states agreed to hand over some of their regulatory powers to independent agencies, the resulting regulatory framework is simpler.
9. Access regulation in the natural gas industry

The natural gas industry has also been under restructuring since the early 1980s. It will be shown, however, that the pace of the reform has been slower than in telecommunications or electricity.

The following sections describe the economic characteristics of the natural gas industry and the access arrangements that allow competition to occur in certain activities with sub-additive costs. As in previous chapters, the particular cases of the UK, US and Australia are discussed.

9.1 Economic characteristics of the natural gas industry

Natural gas is an energy source. In most of its applications, it is burned as a fuel to produce heat. Households consume it mainly for heating and cooking, while industries use it to generate electricity and in a number of industrial processes. In some countries, natural gas is also used as car fuel.

Two physical characteristics of natural gas determine to a large extent the economic structure of the industry. First, natural gas is found in geological basins that are concentrated in certain parts of the world (North Sea, North America, Russia, North Africa, Middle East, etc). This favors disproportionately geographical areas located close to them, in a similar way as shippers located close to large ports benefit from better connectivity and a broader supply of services. However, unlike ports, basins cannot be duplicated, giving their countries (or the companies which exploit these areas) an enormous bargaining power over customers in importing countries. Second, natural gas needs to be extracted and transported to consumption centers, which are usually located far from production basins. As in ports, these activities require large
capital investments (much of them sunk) that generate entry barriers and reduce competition (Gallick, 1993).

The chain of activities needed to extract and deliver natural gas is illustrated in figure 9.1. The product is extracted from underground wells, sometimes as a by-product of oil production. Once it is extracted, the gas is transported to a nearby plant where it is cleaned and processed. Then, the gas is injected into a high-pressure pipeline and transported to a facility called “city gate”, where it is injected into a low-pressure network to its final destination. Gas can also be transported by other means as Liquified Natural Gas (LNG), for which it has to be cooled down, and re-gasified later. Japan, for example, imports LNG from Malaysia, Indonesia and other countries because the country does not have sufficient domestic production (APERC, 2003).

![Figure 9.1: The natural gas supply chain](image)

High-pressure networks are conventionally regarded as transport and the low-pressure one as distribution. As in other previously studied industries, transport and distribution networks exhibit economies of scale and density.

---

It is worth noting that, as an energy source, natural gas competes with other energy sources, such as coal, oil and electricity. The price of these commodities acts as a ceiling on the price of gas (Newbery, 1999). The extent to which these commodities represent an effective competition vary from country to country. In the US, for example, where the price of electricity is relatively low compared to natural gas, inter-fuel competition is intense. But in the UK, where gas is relatively cheap, inter-fuel competition is more limited. Another important characteristic of natural gas is that its demand is highly seasonal. Demand for electricity production is higher in summer due to higher air-conditioning, while household demand peaks in winter because of heating requirements.

9.2 Markets and services

The supply of natural gas can be decomposed in four activities: production, transmission, distribution, storage and retailing.

9.2.1 Production

Production comprises the activities of drilling, extraction and processing of natural gas. The construction of the facilities required to perform these activities (sometimes in the middle of the sea) involves a series of relatively large, sunk, investments that may incentive opportunistic behavior among gas buyers once the investments are made. For this reason, a large part of the investments is carried out under long-term contracts (up to 30 years) agreed beforehand that contain “take-or-pay” clauses. These clauses specify a price that has to be paid even if it the gas is not taken, thus giving assurances to producers that their investments will be remunerated (Newbery, 1999).

Although economies of scale in gas production exist, these seem not to be as large as to impede competition. Indeed, although production fields are unevenly distributed around the world, there are in many countries enough fields to sustain competition among producers. Moreover, there can be competition with producers located in different areas, such as Russia, Algeria and the North Sea, which supply the continental European market (IEA, 1994).

9.2.2 Transport

Transport (or transmission) involves transporting natural gas from the production well to the “city gate”, where it is delivered to high-volume consumers such as retailers, distribution companies or electricity producers. Transporting natural gas requires a network composed of pipelines and compression stations that exhibit sizeable economies of scale. According to Lawrey (1998), the sources of economies of scale in natural gas transport networks are:
a. The cost of obtaining rights-of-way, which are insensitive to the volume of gas that is finally traded.

b. The fact that investments in pipeline networks are indivisible. As in the construction of a new port or terminal, they can only be undertaken in a large scale to be economical.

c. The cost of transport capacity decreases rapidly when wider pipelines are used. For example, capacity for transporting one billion cubic meters a year falls from US$270,000 a mile for a 15-inch pipeline to US$100,000 a mile for a 36-inch pipeline (IEA, 1994).

d. The cost of compression also decreases rapidly when more powerful stations are used. Indeed, the cost of an 80-bar compression station is lower than the cost of two 40-bar compression stations.

Although specific arrangements vary from country to country, transport rates are typically structured as a cost-based, two-part tariff. The first component is usually charged for the capacity that is reserved, while the second is charged for the actual amount of gas that is transported. In Austria, France, Germany and The Netherlands, transport prices are charged according to distance; but other countries such as Denmark, Sweden, the UK and Finland follow a “postage stamp” approach, which is a flat fee charged independently of the gas’ entry and exit points.

In this industry, transport capacity is usually commercialized on “firm” and “interruptible” basis. Under the former, capacity is reserved for the client. However, since demand is seasonal, pipelines usually have spare capacity that can be sold on a spot or “interruptible” basis. Users under “interruptible” contracts only receive services if capacity is available. Customers willing to contract interruptible services are those with higher demand elasticity; i.e., those that can switch from one fuel to another according to short-term price variations, such as electricity generators using combined-cycle gas turbines (FERC, 2004).

In most cases, transport networks are considered natural monopolies. Nonetheless, the capability of their owners of exercising market power can be constrained by several factors, such as the size and concentration of users or the availability of substitutes. Indeed, if users are few or some are relatively large, they may have enough bargaining power to overcome transport owners’ market power. Likewise, market power can also be constrained by competition, which can occur depending on the location of the production fields. For example, in Australia, Sydney is served from two different gas basins located northwest and south from the city: Moomba-Sydney and Longford-Sidney, respectively. The location of these fields allows competition between the pipelines transporting gas from these basins to the city (Productivity Commission, 2004). A similar situation occurs in Austria, where the country is supplied by two competing gas companies that own their transport networks (OECD, 2000).
Competition in transport can also occur if demand is sufficiently large relative to the capacity of individual pipelines. In this case, competition may arise even if producers are located relatively close to one another.

One issue that facilitates competition among transport networks is that high-pressure systems are normally built at a standard size. This makes transport costs more uniform and distance-related than otherwise.

### 9.2.3 Distribution

This activity is carried out by distribution companies, which receive natural gas at “city gates” and distribute it to small and medium gas users through low-pressure networks (see figure 9.1). Unless this activity has been separated from retailing, distribution involves buying the gas, transporting it through the distribution network and selling it to the customers (Gallick, 1993).

As in other network industries, distribution exhibits sizable economies of density that make competition difficult to occur. As discussed in chapter 6, economies of density arise from the fact that once a trunk pipeline has been installed, the marginal cost of making further connections is relatively small. For these reasons, gas distribution is generally considered a natural monopoly, for which distribution companies are regulated.

The degree of public/private ownership of distribution companies varies across countries. In the Netherlands, for example, there are 35 distribution companies, owned by regional and local authorities. Eleven of them also distribute electricity and heat. In Belgium, even though 19 of the 23 distribution companies have private shareholders, public shareholders hold the majority of votes, even when they hold a minority of stocks (IEA, 2000).

### 9.2.4 Storage

As explained before, demand for natural gas is highly seasonal. Since it costs more to build a new network (or expand an existing one) to cope with peak demands, there is a demand for services smoothing the flow of gas through the network. Smoothing is done by increasing the flow at off-peak times and reducing it at peak times.

Figure 8.2 shows natural gas consumption in the US in 2002, where winter demand almost doubles summer consumption levels (Ferc, 2004). It illustrates how storage allows the network to cope with demand seasonality and maintain production levels relatively unchanged.
Gas can be stored in different types of facilities, such as LNG plants, depleted gas reservoirs, or in the pipeline network itself (by increasing pressure).

Even though there is scope for competition in this activity, it may remain regulated, as in the UK. In the US, most gas storage facilities are owned by pipeline companies or distribution companies, for which its terms and pricing policies are subject to regulation either by the Ferc or by the states.

9.2.5 Retailing

Retailers interact between buyers and sellers of natural gas and transport services, making some of all of the arrangements, from gas purchase to final delivery, on behalf of their customers. As in the ESI, retailing arose as a result of the industry reforms.

Retailing is a competitive activity. There are over 260 independent retailers in the US and almost 200 in the UK. They can operate independently or affiliated to gas producing company (NERA, 2002).

9.3 Reform and regulation of the natural gas industry

The traditional structure of the natural gas industry has been that of an integrated monopoly, in which a public-owned company owns the production wells, the pipeline system, and the gas transported through the system. In the US, the structure was somewhat different, since the production wells and the transport networks have been...
usually privately owned. The argument for an integrated monopoly was strengthened by the large economies of scale or density present in the industry’s composing activities.

The restructuring process of the natural gas industry started with the passage of the Natural Gas Policy Act by the US Congress, in 1978. It gained momentum after the privatization of British Gas in 1986 and since then, similar policies have been adopted by other countries in Europe, Asia and Latin America (APERC, 2003). Notable among the deregulation efforts are the EU Gas Directives 98/30/EC and 2003/55/EC, aimed at establishing a European-wide gas market. These directives establish an open access regime for storage facilities and state that storage operators must be legally separated from transport operators.

As in the ESI, reform in the natural gas industry required the creation of a market where gas could be priced independently to transport and distribution. For this market to develop, rival gas producers must have access to pipeline networks and there must be a mechanism for allowing consumers to choose among gas producers. In the US, for example, where a large number of private producers existed, competition was encouraged when prices at the wellhead were liberalized in the late 1970s. As we will see, integrated monopolists were later encouraged to separate their activities and offer transport services to third parties. In the UK, the incumbent was privatized as an integrated monopoly but under open access provisions. Other countries followed a similar path (NERA, 2000).

Some countries have further unbundled their natural gas industries by separating distribution from retailing. As explained before, distribution companies acted both as the owner of the distribution network and as a retailer, buying gas from producers and selling it to consumers. However, since the price at which the gas is purchased is passed on to consumers, distribution companies have weak incentives to buy efficiently. Separating both activities places the purchase decision in the hands of those who have strong incentives to buy gas efficiently, thereby encouraging competition in upstream markets (production and transport). Separating both activities also improves the quality of regulation, because instead of monitoring a competitive market (retailing) and a monopolistic one (distribution), regulators can focus only in the latter.

It is worth noting that even though there is a tendency toward privatization, not all the countries reforming their natural gas industries have decided to do so. This is most evident in Europe, where some of the countries are complying with their legal obligation of implementing EU gas liberalization directives, but have not gone farther in this process (France, for example).

Reforms carried out in the natural gas industry have three common characteristics (Kessides, 1994):

---

54 The Directive 2003/55/EC replaced the Directive 98/30/EC
a. Vertical separation has been implemented between transport and other activities. The degree of integration among the other activities varies according to the country. Depending on the country, separation may be legal or operational.

b. An open-access regime applies to transport pipelines that were historically controlled by incumbent monopolists.

c. Prices for those having enough bargaining power (retailers and large customers, for example) are not regulated.

d. The industry is regulated by an overall energy regulator which also oversights electricity markets.\(^{55}\)

9.4 Access arrangements in the natural gas industry

Although the model of natural gas delivery through an integrated monopolist started to evolve in the late 1970s, it took a series of court and regulatory decisions and the threat of bankruptcy for a workable access regime to develop in the US. As it will be explained in the next section, the path taken in the UK was different but not less tortuous, including several investigations from antitrust authorities and an Act from parliament. Other countries such as Argentina, and Peru, having learned from these previous experiences, decided to restructure its gas industry adopting an open access regime from the beginning.

Open access, however, is not universally accepted as the most beneficial regime. During the debate previous to the adoption of the EU Gas Directive, for example, some countries argued that the unbundling required to implement open access would make domestic companies lose bargaining power \(\text{vis-à-vis}\) off-shore companies with substantial market power, as those located in Russia and Algeria. Other arguments against open access is that reduces the time horizon over which agents are willing to contract, incrementing the risk of investments and therefore increasing supply costs. There are three types of access arrangements under which pipelines traditionally operate (Newbery, 1999):

a. The pipeline may act as a **private carrier**, when it is owned by the same company that owns the producing well and the gas transported through the pipeline.

---

\(^{55}\) In the UK, where Ofgas was initially created with the single purpose of regulating the privatized British Gas, this agency was later merged with Offer to create Ofgem. In Australia, the Australian Energy Regulator (AER) was created within the ACCC to fulfill this task.
b. The pipeline may act as a *contract carrier*, when only provides transport services to shippers. It is free to determine whether or not to accept contracts.

c. The pipeline may act as a *common carrier*, when it is required to offer transport services on a non-discriminatory basis.

In various countries, reference conditions to access pipelines (both price and non-price terms) are proposed by the pipeline owner and approved by the regulator, although access seekers are usually free to negotiate different terms for their particular requirements. This is the case of US and Australia, for example.

In the EU, the common access regime has changed since it was initially proposed. The first Gas Directive (98/30/EC) distinguished two forms of open access: regulated “Third Party Access” (TPA) and negotiated TPA. Under regulated TPA, users obtain the right to access the system on the basis of terms, conditions and rates set by regulators. Under negotiated TPA, only main terms conditions are pre-established; the rest are negotiated individually. The second Gas Directive (2003/55/EC), however, ruled out negotiated access, establishing the regulated option as the standard model for accessing transport and distribution networks. The negotiated option remains open for storage and upstream facilities such as off-shore pipelines.

9.5 Reform, regulation and access regimes in selected countries

9.5.1 United Kingdom

According to the OECD, the UK is the third largest gas market in the world, after the US and Russia (OECD, 2002b). Most of the gas comes from offshore gas fields located in the North Sea and injected into the system at several beachhead facilities, of which St. Fergus, in Scotland, is the main one. The transport system is connected to the rest of Europe via a link to Belgium as seen in figure 8.3. A second link to Norway is under construction.
The UK natural gas industry was publicly owned until it was privatized under the Gas Act of 1986. Nowadays, a single company (Transco) operates the UK’s high-pressure transport pipeline and two main storage facilities. There is a strong competition in gas trade, with 100 wholesale and 90 retail traders.

In 1986, British Gas (BG) was privatized as an integrated monopoly; involved in gas production, transport, distribution, storage, and retail sale of gas appliances. In parallel, the Office of Gas Supply (Ofgas) was created to regulate BG’s competitive behavior. Under the conditions of BG’s privatization, the company had the exclusive right to distribute gas to small customers\(^{57}\), which represented 70% of the demand. Prices in this regulated market had to be set by Ofgas using the RPI-X methodology. The price of gas for large customers was not subject to regulation and access for entering this market was encouraged. Access price and non-price terms for transport had to be negotiated between BG and access seekers, but Ofgas could intervene and set access price and conditions if parties could not agree (Juris, 1998)

\(^{56}\) Source: Parker (2002)
\(^{57}\) Those consuming up to 25,000 therms per year. One therm equals 100,000 British Thermal Units (BTUs).
At the time of privatization, policy makers did not expect integration to deter competition, since producers could sell gas to large users and access the transport network for its delivery at negotiated terms. However, competition could only be introduced after a series of inquiries and procedures by antitrust authorities that ended with the vertical separation of BG.

BG was referred for the first time to the Monopolies and Mergers Commission (MMC) in 1987. The Office of Fair Trade (OFT) argued that the company was discriminating among consumers and that no entry had produced as a consequence of BG’s anticompetitive behavior. The MMC concluded that certain company policies were effectively against the public interest, such as price discrimination, lack of transparency in transport pricing, and a monopsonistic position as the only purchaser of the gas produced in the North Sea. The MMC recommended enhancing the transparency of BG’s activities, ending its discriminatory behavior and limiting its purchases of North Sea gas (Weir, 1999).

After another inquiry, the OFT concluded in 1991 that the implementation of MMC’s recommendations had not been sufficient to promote effective competition in the industry. According to the OFT, three factors were impeding competition: i) BG’s monopolistic position in small-consumer market; ii) BG’s dominance in the large-user market made difficult for competitors to obtain enough gas supplies; and iii), BG’s ownership of the transport system. The OFT also found that BG was able to deter entry by discriminating against competitors and by cross-subsidizing non-regulated activities with regulated revenues. In 1992, the government, following OFT’s recommendations, made a further attempt to promote competition in the retail sector by substantially reducing the consumption threshold under which BG held supply exclusivity. In 1993, BG was referred for the second time to the MMC, and this agency found again that BG’s practices were against public interest. MMC recommended legal separation between BG’s transport and other activities and to remove its exclusive right to supply small customers. These recommendations were not fully adopted before the passing of the Gas Act of 1995, which ordered transport under non-discriminatory conditions, separate price caps for transport and storage, and price caps for gas distribution. As a consequence, BG’s activities were legally separated in the following years. Production and retailing activities were transferred to a company called Centrica; while transport, storage and distribution were transferred to BG Transco. Both companies remained part of the same holding until 2000. In 1997, a dispute between BG and Ofgas about the adequate estimation of the X factor for transport and storage motivated a further referral of BG to the MMC. The MMC decided that storage rights at the two main storage facilities (which granted BG significant market power) had to be auctioned, leaving RPI-X regulation for the remaining storage facilities (Weir, 1999).

The Gas Act of 1995 also required the formulation of a Network Code that set price and non-price access terms to transport and storage facilities. According to the Network Code, transport charges were to be subject to price caps using the RPI-X
methodology. The Network Code worked relatively well until 1998, when capacity constraints began to appear. To allocate capacity rights under these new circumstances, the regulator authorized the implementation of short-term auctions. The auctions were held every six months allowing secondary trading (i.e., shippers could sell not-used contracted firm capacity to other companies). However, if the actual capacity was lower than the rights auctioned, Transco had to buy them back at market prices. This mechanism was intended to give the company economic incentives to expand the transport network. This solution, however, introduced two distortions. First, since transport is a regulated business, Transco’s revenues are capped. In this context, the auctions resulted in an over-collection of revenues. Indeed, in March 2001 the company recovered £370 million against a target of £54 million. It was not clear what to do with these revenues. Second, the short-term character of the auctions and the fact that Transco could keep at least part of the revenues distorted the company’s incentives to expand capacity. To address these problems, Ofgem proposed in 2002 the implementation of long-term capacity auctions (5 years) and a new system of incentives for Transco. Under the new regime, the regulator would set targets for capacity expansions and Transco would be more exposed to buy-back costs (Parker, 2000).

9.5.2 United States

The US gas market is the largest of the world. It is composed by 6,800 producers that sell gas to 1,300 distribution companies and other shippers through a 250,000 km high-pressure pipeline system. The most important production area is the Gulf of Mexico, while imports from Canada account for 17% of the US. Many users have access to more than one pipeline and a series of hubs and markets centers facilitate transactions (Ferc, 2004).

Figure 8.4 shows the flow patterns of the US natural gas trade.
Until the 1930s, most gas production and consumption took place at local level, for which the activity was regulated by the states. In 1938, the US Congress passed the Natural Gas Act, which granted the federal government powers over the parts of the industry involved in interstate trade. In 1954, a decision by the Supreme Court expanded oil federal regulation to include gas production. As a result, prices at wellhead for gas to be sold out of the state were to be regulated by the Ferc, while those for in-state consumption were to be set by local regulators. This decision also stated that interstate prices should be calculated using historic costs, which led to much lower prices than in intrastate markets. The logical consequence of this decision was that producers were discouraged to sell gas out of the states.

This flawed regulation led to a disconnected network of pipelines. As a result of the lack of connection, each city was served using its dedicated gas supplies. The situation gradually worsened until 1978, when the Congress passed the Natural Gas Policy Act which proposed the gradual deregulation of wellhead prices. After the passing of the Act, many pipelines committed to purchase large quantities of gas under contracts containing take-or-pay clauses that required them to pay for gas even if it was not consumed. However, the Act had also raised the price of gas for industrial users and prohibited its use for electricity generation and other industrial uses. Both factors threatened bankruptcy, since demand reduced dramatically and pipelines were left with idle capacity and expensive take-or-pay contracts (NERA, 2002).

In 1987, the Ferc forced pipeline owners to adopt an open access regime. The regulator established two mechanisms to allow passing the costs of the stranded take-or-pay contracts to consumers: companies could add a surcharge to their rates, or
adopt an open access regime and collect up to 50% of past take-or-pay costs in the form of a fixed monthly amount, absorbing themselves an equal proportion of the costs. Most pipelines opted for the second mechanism, which provided them with a guaranteed stable cash flow and secured them a profit. By 1988, all the major transporters had become open-access pipelines.

A series of court decisions questioning led the regulator to issue Order 636, which established an open access regime for gas transport. This order, that still regulates the activity, requires all pipelines to provide unbundled transport and storage services under non-discriminatory basis. Pipeline owners can be involved in gas production or retailing, but have to be separated from these activities and offer other users the same quality of service as the one provided to related companies. The Ferc also established further standards of conduct to prevent preferential treatment. Under this system, access terms and conditions are set by each pipeline, but they have to be consistent with the guidelines issued by the Ferc. These conditions are public and open to inspection by the regulator. Minimum and maximum charges are set by the pipeline themselves following Ferc guidelines58. In certain circumstances, the Ferc allows market-based and negotiated rates. Market-based rates are authorized when the pipeline demonstrates it lacks market power and have been mostly authorized for storage services. Negotiated rates apply for specific requirements and may exceed maximum rates, but should not preclude shippers opting for services under regulated rates and conditions.

Pipelines may use several methods to allocate capacity. Available capacity has to be disclosed on the company’s electronic boards or web sites. Existing firm capacity is allocated on a first come, first serve basis but not-used firm capacity can be allocated on interruptible basis even if it has been fully committed. Holders of firm capacity rights can resell it to other users. For new developments, the Ferc generally requires pipelines to hold open seasons. Authorizations are granted when there is sufficient interest from shippers.

According to OECD (2000), this access regime greatly benefited the industry. As open access spread through the pipeline network, gas producers in fields with low prices demanded transport connections to access customers in areas with high prices, leading to increased interconnection and the development of spot markets at interconnection points. The maximum flow between most points has also increased because open access has created paths around former bottlenecks. There are nowadays 21 major interstate pipelines having in average 70 holders of transport rights.

58 Rates for firm service are structured as a two-part tariff using the “Straight Fixed-Variable” method. Under this method, a fixed “reservation charge” covers the fixed costs of the pipelines; while a variable “usage cost”, covers all variable costs. Since customers on the interruptible service do not possess reserved capacity, they only pay a volumetric rate that is different to the firm service’s usage charge. This volumetric rate usually includes a proportion of fixed charges. Rates can be differentiated by mile, zone or follow a “postage stamp” approach.
9.5.3 Australia

The physical and demographical characteristics of Australia are different to those in the UK and the US. The country has a vast territory and a relatively small population, mostly concentrated in the southwest coast. The natural gas industry developed in the form of isolated networks connecting producing basins and consumptions areas, with little or no competition or interconnection. When the industry began, in the 1960s, it was characterized by monopolistic structures along the production chain (ACCC, 2000).

Competition in Australia is still limited in production, transport and distribution. Natural gas is extracted from 11 basins, although around 95% comes from only three of them: Canarvon, Gippsland and Cooper-Eromanga. Production fields are linked to consumption centers through 20,000 km of transport pipelines, as seen in figure 8.5.

Figure 9.5: The Australian Natural Gas System (ACCC, 2000)

During the 1990s, the Australian government took steps to liberalize the industry and promote competition. Reform started in 1991, when the government drafted a national strategy for the industry. Later, in 1994, the Council of Australian Governments (the Council) agreed to remove administrative barriers to interstate gas trade and implementing the separation of transport and distribution from other activities. A second set of reforms involved the restructuring of the Australian natural gas industry. States agreed to separate transport and distribution activities and create separate
corporations. Integration of transport and distribution is allowed but these activities should be legally separated from gas production and retailing. They also agreed to establish “ring-fencing” arrangements for private retailing companies, which require separate accounts, different personnel and limits on information exchange.

The third set of reforms consisted on the adoption of an open access regime. Although the National Access Regime implemented in 1995 provided a general benchmark applied to all essential facilities, the Council considered that the substantial degree of market power enjoyed by the pipelines required an industry-specific regime. As a consequence, the Gas Pipelines Access Code (the Gas Code) was implemented in 1997. The Gas Code is applicable only to those pipelines that were listed when the code was passed or those later included by the NCC. The criteria under which the NCC must analyze whether to declare a pipeline covered by the gas Code are the following (Productivity Commission, 2004):

a. Access to the service would promote competition in at least one market other than the market for the services provided by means of the pipeline;

b. It would be uneconomical for anyone to develop another pipeline to provide the service;

c. Access can be provided without undue risk to human health and safety; and,

d. Access to the services provided by means of the pipeline would not be contrary to the public interest

Once a pipeline has been covered, its owner must submit an access undertaking for approval from to the regulator. An undertaking describes price and non-price terms for access seekers. Even if the access undertaking is approved by the regulator, the pipeline owner and any access seeker are free to negotiate other price and non-price terms of access (except queuing policy). Reference transport charges are set by the pipelines using one of the cost-based methodologies stipulated in the Gas Code. Regulators have some discretion to determine which costs can be included and can use benchmarking to assess if they are excessively priced. In case of dispute, the reference charges and terms stipulated in the access arrangement apply.

The Australian gas access regime also provides access seekers the right to interconnect with a part of an existing transport or distribution pipeline. In this case, operators are required to unbundle their services to allow third parties access only to the components of the network they require. Regulators may also mandate the expansion of a pipeline if it is justified by the requirements of a prospective user.

Since the launch of the reforms, the structure of gas markets in Australia has changed significantly. A number of major gas facilities have been privatized, such as three distribution companies and one transport pipeline in the state of Victoria. The introduction of retail competition has also encouraged competition between gas basins.
and transport pipelines, such as those that serve the Sidney market (Gippsland and Cooper basins and Moomba-Sydney and Longford-Sidney pipelines, respectively). In response to this changing context, the Australian government initiated a review of the Gas Access Regime. The review, carried out by the Productivity Commission (the federal government’s advisor on microeconomic policy), identified several problems. During the inquiry, the Commission found that the fact that reference tariffs would prevail in case of a dispute has discouraged negotiation and thus led to de facto cost-based regulation. This and other findings led the Commission to formulate two main recommendations. First, to make the Gas Code more consistent with the National Access Regime by establishing more specific pricing principles and allowing the use of methodologies different to cost-based ones (incentive regulation, for example). Second, the Commission considered that market power abuse can be prevented by a credible threat of more intrusive regulation, for which recommended to adopt a monitoring approach when the benefits of regulation are not clear. This “light-handed” option would allow pipelines to set their own access terms and conditions, including pricing, queuing, capacity trading and expansion; but they would be subject to the anticompetitive provisions contemplated in the Gas Code (Productivity Commission, 2004).

9.5 Lessons for a port access regime

An increasing number of countries in Europe, Asia and Latin America are undertaking reforms to make their gas markets more competitive. In these countries, open access seems to provide the framework to promote competition.

Since gas basins are unevenly distributed, location plays an important role in this industry. If an area is located close to more than one production field, competition can arise despite the economies of scale or density present exhibited in pipeline networks. This situation is similar to the one where the hinterlands of two or more ports overlap, as in the Hamburg – La Havre range.

Table 9.1 presents the characteristics of the access regimes implemented in the natural gas industries of the UK, US and Australia.
Table 9.1: Main characteristics of natural gas access regimes

<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>US</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vertical structure</strong></td>
<td>Legal separation. Integration is allowed between transport, storage and distribution; and between production and retailing.</td>
<td>Operational separation. Integration is allowed between transport and storage.</td>
<td>Legal separation. Integration is allowed between transport, storage and distribution; and between production and retailing</td>
</tr>
<tr>
<td><strong>Access pricing</strong></td>
<td>Regulated. Charges capped using RPI-X.</td>
<td>Set by the pipeline owner using a regulated methodology.</td>
<td>For transport, negotiable reference tariffs estimated using a regulated methodology.</td>
</tr>
<tr>
<td><strong>Non-price terms and conditions</strong></td>
<td>Regulated. Set in the Network Code</td>
<td>Set by the pipeline owner but consistent with regulated guidelines</td>
<td>Negotiable reference terms established using regulated guidelines</td>
</tr>
<tr>
<td><strong>Mechanism to expand the infrastructure</strong></td>
<td>Regulated. Ofgem set expansion targets.</td>
<td>Market incentives</td>
<td>Market incentives, but the regulator can order expansions if requested by access seekers</td>
</tr>
</tbody>
</table>

Several lessons can be learned from the efforts to adequately regulate the natural gas industry. The first lesson can be drawn from the UK experience: that the market power of the incumbent cannot be underestimated. Indeed, the integrated structure under which BG was privatized reinforced the information asymmetry between regulator and regulated firm, allowing the company to discriminate among users and to deter competition even if an open access regime was put in place.

The second lesson for port regulators is that establishing an open access regime might not be sufficient to encourage entry and competition. As the UK case shows, as long as the incumbent has incentives and the ability to discourage entry, the latter will not occur. In this country, acceptable competition levels were only achieved when the monopolistic activity (transport) was completely separated from the competitive ones (production and retailing). In the US, competition did not develop until pipelines were obliged to separate transport from gas trading activities.

It is worth noting that in the natural gas industry as in the ESI, the introduction of competition required the creation of centralized markets (for natural gas and electricity) that were inexistent previous to the industry reforms. Vertical separation (either legal or operational) seems to be a necessary requirement for the creation of these markets.

The third main lesson port regulators can learn comes from the US. It can be seen that bad regulation can be as effective in preventing the development of markets as leaving the industry unregulated. In this country, regulation establishing different prices for inter- and intra-state trade and forbidding its use for certain purposes led to a disconnected network of pipelines and an artificially-reduced demand that threatened pipeline owners with bankruptcy.
The fourth main lesson to be learned from port regulators regards the relationship between negotiation and the pricing methodology. The natural gas industry seems to confirm what was observed in the previous chapter, that the use of price caps seems to be better suited for regimes where negotiation is not expected (as the UK). In the US and Australia, where negotiation is encouraged, access charges are set using cost-based methodologies.

A further lesson for port regulators regarding access pricing is that the establishment of reference charges seems to discourage negotiation. Indeed, as an inquiry by the Australian Government found, the incentives set by these arrangements seem to have led to a de facto cost-based regulation (Productivity Commission, 2004).

Regarding non-price access terms, one can see that the methodology to establish them in the analyzed countries is the same used for pricing: regulation in UK, and negotiable reference terms established using regulated guidelines in US and Australia. As for the mechanism to expand the infrastructure, each country uses a different approach. In the UK, the regulator uses a mix of administrative-set targets and economic incentives (the pipeline owner is exposed to buy-back costs) to encourage the pipeline owner to invest adequately. In the US and Australia, it is presumed that demand would encourage incumbents to expand capacity. In the former, however, the regulator requires them to hold open seasons to assess demand. In the latter, the regulator may order the expansion of a pipeline if it is justified by the requirements of a prospective user.

Regarding the issue of capacity expansions, the UK experience also provides an interesting lesson for port regulators. It can be seen that the use of auctions constitutes a useful tool to allocate installed capacity, but it does not provide incentives to expand the infrastructure. Moreover, allowing the incumbent to keep at least part of the proceed would be perverse, since a more acute scarcity would produce higher revenues. For these reasons, auctions should be used together with other economic incentives, such as expose the incumbent to buy-back costs.
10. Access regulation in the rail industry

Due to its natural monopoly characteristics and the strategic importance that railways play in the development of the economy, the rail industry has been heavily regulated in most parts of the world. Prices, entry, operations and vertical structure have been typically subject to some form of government control. However, public dissatisfaction with traditional forms of regulation led to the restructuring of the industry in countries such as Japan, Sweden the US, and the UK. Since then, many other countries have followed through.

The following sections describe the economic characteristics of the rail industry, and the main approaches taken to deal with the access problem in several countries. As in previous chapters, the cases of the UK, US and Australia are discussed in detail.

10.1 Economic characteristics of the rail industry

Rail activities can be broken down into i), infrastructure provision, and ii); train services. Infrastructure provision involves the supply of tracks, stations, terminals and other ancillary facilities\(^{59}\); and related services, such as signaling and traffic management. Train services involve the transport of passengers and freight among different origins and destinations (see figure 10.1).

\(^{59}\) Storage yards, marshalling yards, sidings, switches, fueling and maintenance stations, among others.
Rail activities possess the features that characterize network industries, such as sunk costs, economies of scale and scope, and externalities. Sunk costs, for example, arise because rail tracks, rights-of-way and other assets that represent a large part of the total investment costs have little value for activities other than rail transport. Once made, these investments are unrecoverable.

Economies of scale arise because the large share of fixed costs associated with rail operations prevent total costs to rise proportionately with traffic volume. Indeed, costs of track, rights-of-way, locomotive power, crew and related facilities change little if the train carries 10 or 500 passengers. And most ancillary infrastructure (yards, sidings, switches, etc.) can be used for a large number of shippers without further investment requirements (Kessides and Willig, 1995). This characteristic is similar to that of liner shipping, where running costs are relatively stable regardless the number of containers transported.

Rail operations also exhibit a particular type of economies of scale known as “economies of density”. In train services, economies of density arise because is cheaper to add additional cars to scheduled services than operating new trains. Ivaldi and McCullough (2001) measured economies of density for Class I freight railroads in the US. They found that returns to density amount to 1.65; i.e., doubling the number of services would only increase costs by 65%.

Economies of scope derive from the multi-product nature of the industry, which implies that the same labor, facilities and equipment used for producing services for one market can be to produce services in a different one. For example, tracks and
ancillary infrastructure used for passenger services, can also be employed to provide freight services at proportionally lower additional cost.

A distinctive feature of this industry is that rail companies are typically multi-product firms producing a multi-dimensional output. In fact, rail companies typically produce different types of transport services for different users at different origins and destinations, at different quality levels. The combination of these dimensions has different implications in the cost structure of rail companies. For example, long-distance passenger carriers have a different cost structure than local or regional ones, although they can coexist in certain routes (Campos, 2002).

It is worth noting that the multi-product nature of the rail industry, combined with the fact that many of the costs incurred are fixed, joint or common, have the effect of making arbitrary any system to allocate costs among particular services. Some costs may be attributed to a particular type of traffic, such as stations to passenger traffic and terminals to freight, but others, such as energy, cannot. This is an important feature for regulating access, since this arbitrariness can be used to deter competition by charging predatory prices in competitive services (Campos and Cantos, 1999).

Railways also exhibit positive externalities, since the network’s value increases for the existing users when a new link is added. Moreover, as suggested by Button (1993), railways may also create positive externalities for the transport system as a whole. Indeed, costs derived from congestion, accidents or environmental impact would be substantially reduced if part of the road traffic were transferred to railways, since road transport does not fully internalize these costs.

An important characteristic of the rail industry, also shared with ports, is that tracks, stations and rolling stock can only be expanded in discrete amounts, while demand may fluctuate in small amounts. As a consequence, periods of excess capacity may be followed by periods of great congestion. This raises the question of how rail services should be priced. They can be estimated according to short, medium or long-term costs, and may or may not include congestion charges. Moreover, since railways are part of an overall transport system, efficient pricing for rail also depends on factors which cannot be controlled by rail managers, such as policies regarding other transport modes. Furthermore, pricing decisions may require complementary measures to achieve economic efficiency, which are also beyond the reach of rail managers, such the granting of subsidies to certain types of users.

A last characteristic of the industry, although a non-technical one, is that railway transport has been typically considered as a social service. The low rolling resistance of steel wheels on steel rails make rail transport fuel efficient and relatively inexpensive, and the industry is regarded as a mechanism to integrate distant areas, to

---

60 Joint costs are those incurred in production processes involving several products that can be separated or distinguished. A common cost cannot be traced to each product individually (Fernandez-Baca, 2006).
promote economic development and as a guarantee for minimum transport services for particular segments of the population (Campos and Cantos, 1999).

10.2 Markets and services

As in ports, the rail logistics chain is composed by inputs produced in both monopolistic and competitive markets. Indeed, while infrastructure provision is considered a monopolistic activity whose duplication generates additional costs to users, train services can be provided in competitive markets. Inputs from both activities are simultaneously required to transport passengers and freight.

Rail markets can be distinguished by whether the service involves the transport of passengers or freight and the origin and destination of the service. Competition for a particular market can come from other train operators or from other transport modes. The extent of competition, however, depends on geographic, demographic and economic factors (De Rus, Campos and Nombela, 2003). In the US, Canada and Australia, for example, long distances and low population density limit passenger services. In Europe, shorter distances and higher population densities make rail passenger transport feasible. For passenger traffic, flexibility confers advantages to road transport over rail for short distances. For long distances, air transport has the advantage presumably because time savings overcome the higher costs involved. But rail has advantages over other modes for medium-range transport, especially if traffic flows in relatively consistent patterns.

In freight, the ratio of value-added to transport costs confers advantages to one transport mode over other. Rail, for example, has advantages for transporting bulk commodities such as coal, forest and oil because their low value-added per unit of weight requires transport in large shipments, and because they are less sensitive to the long transit times involved. In the US, for example, 43.5% of the rail freight carried in 2006 was coal and other bulks such as chemicals, farm products, metallic and non-metallic minerals, and oil accounted for almost 30% (Association of American Railroads, 2006). Rail also has cost advantages in long-distance shipments. Indeed, according to Eurostat (2007), 83% of the freight carried by rail in the EU is transported 150 km or more. However, where navigable waterways exist, waterborne transport represents an important source of competition for commodities of the same characteristics. Trucking, on the other hand, has advantages in small, short-distance, time-sensitive shipments of products with high value added. Pipelines also compete with rail for oil products and other liquid bulk. Although relatively inflexible, pipelines are considered as energy-efficient, cost-effective transport modes, that may reduce the environmental impacts involved in the transport of fuels.

It is worth noting that the exertion of monopoly power by rail operators may be limited in a more indirect way. For example, since rail-transported coal competes with pipeline-transported oil and natural gas for the generation of electricity, inter-fuel
competition may limit the price rail companies can charge for transporting coal (Kessides, 2004). Geographic proximity may also facilitate competition, especially for large customers. Indeed, proximity may allow captive users to ship their freight to competing railways by truck, or to lay a dedicated link connecting their facilities with the tracks of a competing railway. Another option may be for large users to induce competition by locating their plants close to the railway company that offers better conditions. In reality, however, it is very difficult for the shippers of some commodities to avoid operators’ market power unless alternative transport options exist or an adequate regulatory framework has been implemented.

10.3 Reform and regulation in the rail industry

The rail industry has shown a steady decline in both developed and developing countries since the 1970s. For example, the industry’s share in both passenger and freight traffic more than halved in the 43 countries belonging to the European Conference of Ministers of Transport. As seen in table 10.1, the share of rail in passenger traffic fell from 22.6% in 1970 to 10.4% in 2002.

<table>
<thead>
<tr>
<th>Table 10.1: Share of rail in total passenger transport, ECMT countries</th>
<th>1970</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Europe</td>
<td>10.3%</td>
<td>6.8%</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>48.6%</td>
<td>11.1%</td>
</tr>
<tr>
<td>Commonwealth of Independent States</td>
<td>70.9%</td>
<td>49.0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22.6%</strong></td>
<td><strong>10.4%</strong></td>
</tr>
</tbody>
</table>

Source: ECMT (2004b)

Analogously, the percentage of freight traffic transported by rail fell from 73.6% to 34.4% during the same period (see table 10.2).

<table>
<thead>
<tr>
<th>Table 10.2: Share of rail in total freight transport, ECMT countries</th>
<th>1970</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Europe</td>
<td>29.5%</td>
<td>13.8%</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>76.2%</td>
<td>35.8%</td>
</tr>
<tr>
<td>Commonwealth of Independent States</td>
<td>85.9%</td>
<td>43.9%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73.6%</strong></td>
<td><strong>34.4%</strong></td>
</tr>
</tbody>
</table>

Source: ECMT (2004b)

Several factors contributed to the industry’s decline. On the one hand, the changes occurred in transport markets. In many countries, rail networks were determined by technologies, market and consumer locations of the 19th century. In passenger markets, technological changes in air and road transport made these modes more competitive than rail. In freight, the development of more sophisticated logistics systems has given preference to other modes, especially road and air transport. The fact that high value-added products are less sensitive to transport costs also confers

---

61 Measured in thousand million passenger-kilometers
62 Measured in thousand million ton-kilometers
advantages to other transport modes. On the other hand, political interference in investment and operational policies often forced rail companies to provide unprofitable services without receiving the corresponding subsidies, thus reducing the amount of resources required to adequately maintain and expand the network. Political interference also reduced rail companies’ ability to withstand pressures from trade unions, which artificially increased the cost of providing train services. Moreover, rigid management structures and lack of competitive incentives made many state-owned companies unresponsive to customer needs (OECD, 2005). According to the same source, the poor performance of rail companies can be related to a problem known as “soft budget constraint”. As explained before, railway transport has been typically considered as a social service, for which governments find it very hard to limit the size of the subsidy offered to rail companies, regardless of their efficiency. Governments are also reluctant to reduce subsidies because railways face competition from other transport modes, such as roads, which are not efficiently priced. Indeed, even though various forms of road pricing are becoming common, road users are not fully charged for some of the costs of using the road, such as the cost of providing the infrastructure or the cost of congestion. Reducing subsidies would incentive traffic switch from rail to road. If the government is not willing to commit to a reduction of subsidies, managers face modest incentives to improve rail companies’ performance.

The industry’s first serious problems appeared during the 1970s, where tight regulation and increasing road competition caused the bankruptcy of several major rail companies in the US. The Congress was then forced to intervene to avoid the disappearance of passenger services. During the 1980s, rail companies in other parts of the world experienced similar problems. In Japan, accumulated debts worth more than US$ 300 billion led Japan National Railways to a severe financial crisis. In Sweden, a series of problems led the state-owned Statens Järnvägar to increasing deficits that reduced investment and service levels. In UK, a series of recurring crises since the late 1950s had made British Railways economically weak. At the time, rail companies in Africa and Latin America faced similar problems, which were worsened by the lack of public funds.

In these circumstances, governments were forced to intervene. In Japan, the state-owned railways were split into nine different infrastructure, passenger and rail traffic companies in 1986. In Sweden and the UK, the reforms initiated also separated infrastructure provision from train operations. In the 1990s, other countries such as Argentina and New Zealand followed through. Table 10.3 shows the characteristics of the reform processes undertaken in these countries. It is worth noting that even though most reform programs sought to introduce competition in rail markets, this was not their main driver. Reform was primarily aimed at reducing the heavy financial burden that railways imposed over public budgets (Guriev, Pittman and Shevyakhova, 2003).
| Country | Before restructuring | After restructuring | | Before restructuring | After restructuring | | Before restructuring | After restructuring | | Before restructuring | After restructuring |
|---------|----------------------|---------------------| |----------------------|---------------------| |----------------------|---------------------| |----------------------|---------------------|
| Market structure | | | | | | | | | | | | | |
| | Public Monopoly (JNR) | Six passenger companies, one for freight | | Public monopoly | Monopoly on infrastructure and quasi-monopoly in services | | Monopoly in hands of New Zealand Rail Ltd. (NZRL) | | Integrated company facing inter-modal competition | | Public monopoly | Franchises for six freight and seven passenger companies |
| Ownership of rolling stock | | | | | | | | | | | | | |
| | Public | Nine infrastructure, passenger and rail companies | | Statens Järnvägar (SJ), government department | SJ and of small companies | | Public agency | Private | Ferrocarriles Argentinos (FA), public enterprise | | Six regional companies |
| Ownership of infrastructure | | | | | | | | | | | | | |
| | State owned | Owned by the new companies | | State owned | Managed by the National Rail Admin. | | State owned | Private ownership until 2004 | | State owned | State network open to third parties |
| Separation between infrastructure and services | | | | | | | | | | | | | |
| | Unified management | Unified management (passenger) and trackage rights (freight) | | Unified management | Separation. Services run by SJ and small firms. | | Unified management | Unified management | Unified management under FA | | Management of companies. Trackage rights exist |
| Regulatory framework | | | | | | | | | | | | | |
| | Regulated prices | Unregulated pricing | | Controlled prices | Control over tariffs—not on access charges | | Prices and service level regulated | Unregulated pricing | Regulated prices | | Price caps, quality service minimum frequencies |
| Reasons for deregulation | | | | | | | | | | | | | |
| | Reduction of state subsidies. Productivity improvement | Reduction of subsidies, improvement of productivity and traffic levels. | | Reduction of subsidies | Reduction of subsidies | | Reduction of subsidies | Improvement of productivity and traffic levels. |

Source: Campos and Cantos (1999)
Given the sizable economies of scale, scope and density that arise from having a national-wide rail network operated by a single rail company, one would expect reform programs aimed at reducing this financial burden while maintaining the integrated structure. In reality, however, few countries opted for doing so. One of the few was New Zealand, and the results were not encouraging.

In 1993, the government privatized the public-owned New Zealand Rail Ltd. as a single, integrated company. In this country, railways face strong competition from road and water transport options, for which previous restructuring was not deemed necessary. More importantly, the government chose the privatization option to make clear its commitment not to continue subsidizing the industry. At the beginning, new technology was introduced and productivity, as well as profits, rose considerably. However, around the seventh year after privatization, the company was put under financial distress because the new owners had not invested enough in maintaining and improving the infrastructure. The result was a series of problems related to lack of maintenance, safety concerns and reduced service quality. In these circumstances, the “soft budget constraint” problem discussed previously appeared. In 2004, the government could not withhold its promise not to continue subsidizing the industry, and took over the network. Since then, it is investing an important amount upgrading the network (Williams, Greig and Wallis, 2005).

In other countries, governments chose to reform their industries without maintaining a single integrated rail company, opting for dividing the system into competing regional networks or implementing open access regimes. In the EU, for example, reforms were oriented at introducing vertical separation, implementing open-access regimes and liberalizing freight and passenger markets. Between 1991 and 1995, the European Commission issued Directives 91/440/EC, 95/18/EC and 95/19/EC; separating infrastructure from train services (operationally, at least), and requiring incumbents to grant access to third parties under non-discriminating conditions. Moreover, since 2001, the Commission has issued three so-called “railway packages”. The first two were aimed at introducing competition in freight markets while the third, issued in late 2007, sought to liberalize international passenger services.

During the 1990s, Latin American countries chose to reform their rail sectors by separating and concessioning their networks to the private sector. In Argentina, the network was separated into six freight and seven passenger networks, and concessioned as vertically-integrated companies. In Brazil and Mexico, whose railways only provided freight services, the network was also separated and concessioned to be run as vertically-integrated companies. According to Sharp (2005), the overall results of the Latin American rail reform programs are positive: governments were relieved of financial burdens, service quality improved, investment increased and the rail market share stopped declining. However, since their main aim was to reduce financial burdens, not enough attention was given to achieving social objectives such as accessibility and affordability to small shippers and passengers, for which the results on these fields are not encouraging.
In Africa, thirteen railways were concessioned between 1993 and 2005. Two of them were cancelled, one was affected by war and another suffered from natural disasters. According to Bullock (2005), despite these problems the results are encouraging. Throughput has generally increased and the railways are arguably performing better than if they had not been concessioned.

Although specific measures are different from country to country, the reforms carried out in the rail industry have four common characteristics (OECD, 2005):

a. Horizontal and vertical restructuring. As discussed in the next section, this may involve the vertical separation of the industry (legal, as in UK and Sweden; or operational, as in Italy and the Netherlands); or the partition of the network into integrated competing rail companies (as in Brazil or Mexico).

b. Rebalance of the roles played by the public and private sectors. Reform has given the private sector a larger role in providing competitive services, especially in freight markets. In several countries a regulator was set up to monitor the industry, such as the Federal Railway Office (Eisenbahn-Bundesamt) in Germany, the Office of the Rail Regulator in UK or the Swedish Rail Agency (Järnvägsstyrelsen). In Europe, the public sector continues to provide rail services directly, while in Latin America they have been almost entirely concessioned to the private sector.

c. More flexibility has been granted to rail companies to set prices and shut down non-profitable services.

d. Subsidies to operators have been granted using open bidding mechanisms, thus improving the transparency of their allocation.

10.4 Access arrangements in the rail industry

After twenty years of reform in the railway industry, three restructuring models have emerged: i) competition between integrated rail companies (without access), ii) vertical integration with competitive access, and iii) vertical separation. All of these options address the relationships between rail companies, markets served and functions performed, including ownership, maintenance and improvement of facilities, operation management, marketing, and financial activities.
10.4.1 Competition between integrated rail companies

As discussed before, nation-wide networks operated by an integrated company maximize the economies of scale, scope and density present in rail operations. However, in absence of effective competition, vertical integration leads to lack of commercial dynamism, low productivity, poor service quality, among other results. (Guriev, Pittman and Shevyakhova, 2003). Therefore, the first restructuring option consists on separating the system into regional networks. In this way, competition from other rail networks or transport modes would incentive the operators of regional networks to achieve efficiency and improve service levels.

This restructuring option has the advantage that in the presence of effective competition, regulatory intervention is only necessary in cases where shippers do not have any alternative transport option. Only in these cases, regulation of prices, quality or access would be warranted.

One country whose government chose this restructuring option is Mexico. The state-owned Mexican railways were divided into three regional networks and concessioned to the private sector during the mid-1990s. No concessionaire has exclusive access to major cities and all of them serve Mexico City (the county’s main economic center). Two railways, located in the north of the country, have access to ports located in both oceans and extend as far as the border with the US. They were concessioned to regional companies which entered the business in order to develop inter-modal transport systems (Campos, 2002). Since its concession, the industry’s productivity has improved notably and railways have recovered market share from roads. Other countries that have restructured their industries under this model are Cameroon, Côte D’Ivoire, Burkina Faso, Argentina, and Brazil. Two further countries also operate under the system of competing integrated rail companies: US (which will be analyzed in detail later) and Canada. In the latter, two rail companies (Canadian National and Canadian Pacific) operate networks covering the southern part of the country, from coast to coast, and serving all major cities. Both companies also have tracks connecting their networks with main US mid-west cities: Detroit, Chicago and Minneapolis.

The main disadvantage of separating the national network is reduced connectivity. Indeed, the more fragmented the rail system, the less likely that a rail company can provide any given end-to-end service. As a result, shippers wanting to send freight to a certain destination may need to send it first to an intermediate point and from there make arrangements with a second rail company to carry it to its final destination (Campos, 2002).

This restructuring option seems to be especially attractive for large, freight dominated railways like those of US, Canada and Mexico. Less dense networks are unlikely to have multiple paths connecting the same origin-destination pairs, for which the resulting level of competition may not be intense enough as to eliminate inefficiencies
and limit market power. And even if it is possible to separate the network in such a way that all origin-destination pairs are connected, it is unlikely that the alternatives will be similarly attractive to passengers. Indeed, since passenger traffic is more time-sensitive than freight, only few users would be willing to switch to alternative options involving longer routes or more train changes. It would be necessary to duplicate expensive infrastructure to achieve competition.

10.4.2 Vertical integration with access

The second restructuring option seeks to preserve economies of scale, scope, and density and to introduce a competitive discipline even in markets where no other transport option exists. Under this option, integrated rail companies operate the infrastructure and provide train services, but they are obliged to grant third parties access to the infrastructure under non-discriminating terms.

In theory, this restructuring option has two main benefits. First, it enhances competition, since train companies compete with each other in the same routes and shippers are even allowed to transport their own freight. Competition ensures that rail services users need are produced efficiently and sold at an efficient price. Second, unlike the previous option where the network was separated, preserving and integrated system provides a greater scope for seamless end-to-end services.

In reality, however, open access has not attracted as much entry as expected. Incumbents still have large market shares or totally dominate the market, especially in the passenger segment. In the Netherlands, for example, a second company called “Lovers Rail” was allowed to compete for passenger traffic in two routes (Amsterdam-Haarlem and Amsterdam-Nijmegen) in 1996, but it had to withdraw after only three years. The company left the market complaining that the incumbent was not neutral in allocating paths (periods in which a company can use the shared infrastructure). In Germany, only one other company competes for passengers with the incumbent DB (and only in two routes), although there are charter companies providing seasonal services in international routes (Van de Velde, 2000).

Competition seems to be more intense in freight markets, although market shares of competing companies are small in comparison with those of the incumbents. In Germany and Switzerland, for example, the competitors of the incumbent operators had only reached a market share of 6.9% and 12%, respectively; after at least four years of open access (OECD, 2005). According to Pittman (2005), this lack of entry can be explained because the economies of scale, scope, and density enjoyed by integrated railway operators are so large that prevent competition from non-integrated companies.

The main problem with this option is similar to those in other industries: the incumbent has incentives to discriminate in favor of its own train services. In fact, it
relies heavily on the regulator’s ability to set adequate access terms, ensure efficient and timely investments, identify incumbents’ discriminatory behavior and manage congestion, among others. As argued by Pittman (2005) and Fagan (2007), infrastructure is more closely related in this industry than in others, making this task a very difficult one; and especially difficult for developing countries without experienced regulators.

10.4.3 Vertical separation

Under this option, infrastructure services are provided by an entity separated from the companies supplying train services, with the infrastructure owned by the government, a consortium of operators or a regulated private entity. This is the model used in Sweden and the UK.

As discussed in chapter 3, the nature of the access problem is different under vertical separation. In theory, separation would allow a more effective competition among providers of train services, since the incumbent would not have incentives to discriminate nor foreclose entry. However, vertical separation can create coordination problems, undermine economies of scope and impose higher transaction costs. In fact, separation implies tasks that were previously coordinated within a single management structure need to be negotiated in processes that are costly: reaching a deal requires time, its implementation requires supervision and differences may require some kind of arbitration. For example, high-speed passenger services can only be provided if the track reaches higher maintenance standards than for regular services. Under vertical separation, the terms of such coordination need to be reflected in contracts whose negotiation is lengthy and costly. Separation also requires rail companies to have personnel that would not be required under integration. In the UK, for example, where a system of financial penalties has been implemented to deal with delays, rail companies employ more than 300 people to operate it (OECD, 2005).

There is some econometric evidence that the costs that vertical separation imposes are considerable. Ivaldi and McCullough (2004) analyzed Class I railways in the US for the period 1978-2001 and concluded that a fully integrated firm would have a 20-40% cost advantage over a vertically separated company. Further problems of separation is that it requires complicated timetabling and path allocation that hinder investment planning, since the interests of the network owner and train services providers do not necessarily coincide.

Vertical separation has not been more successful that the previous option in generating competitive entry. In Sweden, for example, the reform implemented in 1988 separated the functions of the state-owned rail operator Statens Järnvägar (SJ). The National Rail Administration took the responsibility of maintaining and operating the infrastructure, while SJ was required to continue providing train services, but under commercial terms. In passenger markets, SJ operates commercially viable
routes in the national trunk system (regional governments have responsibility over regional systems). Non-commercially viable routes are franchised to the bidder asking the least subsidy. In this market, SJ has a market share of 77%. In freight markets, the incumbent provides services through subsidiaries, which compete with more than 10 other companies. The largest competitor, however, only holds 12% of the market; while SJ controls 85% (Statens Järnvägar, 2005).

The EU first tried to adopt a model of vertical integration with competitive access. Indeed, the directives 91/440/EC, 95/18/EC and 95/19/EC required member states to separate the cost accounting records of infrastructure facilities from those of train services and to allow the use of that infrastructure by third parties. This model, however, was not yet adopted by all member states when the first “railway package” was issued. This package favors the adoption of a model of vertical separation, setting deadlines for the complete separation of infrastructure from train services and the implementation of an open access regime. The second rail package was aimed at introducing competition in freight markets while the third sought to liberalize international passenger services.

It can be seen that no restructuring model to be superior. Conserving a national-wide integrated rail operator leads to inefficiencies. Dividing the system into regional networks reduced connectivity and is difficult to implement in non-dense networks and passenger markets. Implementing access policies does not necessarily lead to effective competition and requires regulatory skills that may be difficult to find in developing countries; and separation imposes sizeable transaction costs that might even discourage entry.

10.5 Reform, regulation and access regimes in selected countries

10.5.1 United Kingdom

Railways in the UK were built by the private sector during the 19th century. Nowadays, the network comprises 21,000 miles of tracks, 2,500 stations, more than 9,000 level crossings and 40,000 bridges and tunnels (Network Rail, 2005). Figure 9.2 shows the current British rail system.
At the time when the rail system was owned and operated privately, industry structure, rates, and terms of service were regulated by the government. Nationalization of the industry occurred in 1948. The four existing rail companies were restructured into six regional groups within a single national railway. Further changes took place in the 1960s, when the increasing financial difficulties of the state-owned British Rail (BR) led to major cuts and reduction of train services. During the 1980s, BR’s core businesses were structured into business units and non-core businesses, such as train manufacturing, were privatized. However, despite these efforts, BR continued to lose traffic and accumulate financial deficits. According to Freeman and Shaw (2002), at the time of privatization BR suffered from low-quality services and was regarded by a large part of the public as “dear, often dirty, overcrowded and unreliable”.

In 1993, the Parliament passed the Railways Act, which provided the framework for the privatization of BR. In 1994, BR was vertically and horizontally separated into approximately 100 companies providing infrastructure, passenger, freight and

---

maintenance services. Infrastructure services were to be provided by Railtrack, a company prohibited to provide train services. These include track, signaling, bridges, tunnels, and certain stations, freight terminals, sidings, yards, warehouses and other facilities. Passenger services were to be provided by train companies operating in franchised routes. Freight services were divided into six corporations: three regionally-separated bulk companies and three additional ones specialized in container traffic, non-bulk international services and postal services, respectively. Likewise, BR’s rolling stock was distributed among three companies which would lease the equipment to train operating companies.

The structure of the freight businesses nonetheless changed after the initial separation. Three companies were sold to a single operator (English Welsh Scottish Railways), which later acquired two more of the privatized companies, and at present carries a large part of the rail freight.

Responsibility for regulating the privatized industry was divided between the Rail Regulator and the Office of Passenger Rail Franchising. The Rail Regulator is an independent agency responsible for supervising access to track and rail markets. The regulator must ensure that the infrastructure operator’s income (access charges and transfers from the government) is sufficient to finance the maintenance and enhancement of the network. The main responsibility of the Office of Passenger Rail Franchising was to auction passenger services and to monitor the conditions offered by train operating companies. In 2001, this office was replaced by the Strategic Rail Authority and entrusted with wider responsibilities. This authority franchises the routes to the highest bidder or to the operator which requires the least subsidy. Most of the routes are, however, franchised through the second method.

Access charges for passenger services are set by the Rail Regulator in the form of price-caps, using the RPI-X methodology, in five-year periods. The regulated revenue of the infrastructure operator is a function of the value of the network assets, the returns the regulator considers the company needs to finance investments and the costs of operating and maintaining the railway to achieve standards set by the regulator. Plans for upgrading and expanding the network are scrutinized by the regulator and only those that are considered necessary are allowed to be remunerated by access charges. Most non-price access terms are set by the regulator.

For the first review period following the privatization of the company (1996-2001), access charges for passenger services were calculated to cover all of Railtrack’s costs and investments. For the second review period (2001-2006), the government decided to subsidize network usage to encourage rail transport, for which one third of the company’s investment plans were to be financed by the Strategic Rail Authority.

In the case of freight, criteria for setting access charges changed from the first to the second review period. During the first one, access charges were negotiated directly with individual operators and the resulting contracts reviewed and approved by the regulator. The criteria to approve freight access charges were the following:
a. The structure of charges should enable the network operator to recover its total freight costs plus any expected contribution to shared common costs;

b. Charges must be set below stand-alone costs; and,

c. Charges should be neither higher nor lower than for other operators taking into account specific factors related to the services provided.

The regulator changed its approach for the second review period. It noticed that access charges represented a large part of rail freight costs, and therefore had a direct impact on the choice between rail and road transport. In line with the government’s policy to encourage freight to switch to rail, the Rail Regulator concluded that a more deterministic procedure would be used to set access charges. Under this new approach, scope for negotiations would be reduced and charges would be set by the Rail Regulator according to published criteria: facilitation of competition, reduction of entry barriers and entry encouragement. Negotiation would only be used to settle issues such as the split between fixed and variable costs and the performance regime. This change in approach resulted in a reduction of access charges of approximately 50%, thus encouraging freight to switch to rail. The Strategic Rail Authority agreed to finance the deficit in Railtrack’s revenue originated by this reduction and to pay for the specific costs of major enhancements to the network required by freight traffic. Other changes were to estimate charges using longer term efficiency gains, differentiated by commodity according to the likelihood of competition from other transport modes. Charges for iron ore and coal, for example, would reflect only half of the efficiency gains, since higher rates would encourage these commodities to switch to road. Overall, freight access charges were reduced by approximately 50% (Drew, 1994).

It is worth noting that the underestimation of the amount of investments needed to upgrade the network, and a series of events triggered by the accident occurred at Hatfield in 2000 led to Railtrack’s bankruptcy in 2001. The accident prompted a large program of repairs and replacements that disrupted timetables and led to long delays across the network. Railtrack was forced to spend vast sums, both on replacing track and on compensating the train operators for the resulting delays to their services. Partly as a result of the Hatfield accident, Railtrack predicted a funding gap of almost £5 billion between 2001 and 2006. In 2002, Railtrack went bankrupt and was restructured into a “not for profit company” and renamed Network Rail. This new company is financed entirely by debt, with the government providing a guarantee (Thompson, 2004)\textsuperscript{64}. In a review of the British restructuring and privatization process, this author arrived to the following conclusions:

\textsuperscript{64} According to Thompson (2004), Network Rail “is a specific form of legal entity under U.K. law that occupies a “not for dividend” status. It was set up for the sole purpose of acquiring Railtrack and enabling the business to come out of administration. It is a company limited by guarantee, owned by members rather than shareholders, which does not pay dividends. The Board of the new company is
a. The British was one of the most contentious of rail reforms. At the time of privatization, there was no real consensus on the problem to solve and the options to be pursued.

b. Criticisms have been directed mainly at the costs of the restructuring process and the fragmented nature of the services that resulted.

c. Separation of infrastructure from operations cause problems of complexity and created transaction costs. It was not, though, the cause of increased accidents. Moreover, the safety record has improved significantly under the new system.

d. The objective of the government of transferring a significant degree of risk (demand, operating cost and investment scheduling) to the private sector turned out to be unrealistic and was not achieved.

e. Passenger and freight traffic have grown rapidly, with passenger demand at levels higher than at any point in the last 60 years. Investment levels in equipment and infrastructure are higher than they have ever been.

10.5.2 United States

The rail industry in the US is one of the largest in the world. The geography and economic structure of the country causes the industry to be dominated by freight traffic. There are about 500 private freight companies, out of which seven mobilize nearly 80% of the cargo. The remaining portion is transported by regional firms that operate branch lines that feed into the major networks. These small rail companies tend to operate on low-density routes that major companies abandoned after the regulatory reforms of the 1970s (Krohn, 1998).

Rail companies in the US are classified in three classes. Class I railroads are those with revenues over US$ 250 million; Class II railroads are those with revenues between US$ 20 million and US$ 250 million, while Class III railroads are those with revenues under US$20 million. Intercity passenger services are provided by the National Passenger Railroad Corporation (Amtrak), a public corporation. Commuter networks are financed locally.

accountable to and may be removed by the members, although the Board is involved in selecting the public members. Network Rail’s members comprise the SRA, license holders (…), and other interested organizations and individuals drawn from the general public. There are between 100 and 120 members, of which representatives of the public interest form a majority but no more than 75 percent of the total. SRA has special membership rights, including the right to appoint a member of the Board of Directors and the right to remove all other members in the event of fundamental financial failure. Despite the unusual structure of Network Rail and its underlying government guarantees, it is still designated as a private company: it has not been formally re-nationalized.” (p. 19)
Since its creation in the 19th century, US rail companies have been vertically integrated and privately owned. In 1887, the Interstate Commerce Commission was created to regulate the industry, which then dominated the transport sector and enjoyed substantial market power. Regulation was further strengthened during much of the 20th century. Since the 1950s, the share of freight transported by the railways declined as competition from other modes increased. Much of this decline was attributed to restrictive regulation from the Interstate Commerce Commission. During the 1970s, returns on investments fell to low or negative levels, and about 20% of the industry went on bankruptcy (OECD, 1998.)

In 1970, the Congress passed an Act creating Amtrak, a federally-owned company who took over unprofitable intercity passenger traffic from private railways. Amtrak is funded by the Congress through federal capital and operating grants, which are determined on an annual basis. The company is operated as a for-profit corporation and managed as a private company. It owns the rolling material and 20% of the infrastructure it uses, in particular the North-East corridor (Washington-Philadelphia-New York-Boston). In 1980, the US Congress passed the Staggers Rail Act, which deregulated the industry and enhanced competition levels in the industry. The Act relaxed the method of regulating freight rates, allowed abandoning unprofitable lines and removed antitrust limitations. The Act severely limited the powers of the Interstate Commerce Commission, particularly in the areas of rate setting and access rights. These reforms led to an increase in the productivity of labor and capital and a
decrease in freight rates. Around 40 Class I railroads were consolidated into nine companies, and some small rail operators emerged (Productivity Commission, 1999).

In 1995, the Interstate Commerce Commission was replaced by the Surface Transportation Board (STB). The STB is an independent body attached to the Department of Transportation. It supervises access and antitrust issues for all rail companies operating in the US, although it is primarily focused on Class I railways. For example, the STB has the right to approve or decline mergers or to impose conditions on the merger to avoid market dominance. The US rail policy largely relies on intra and inter-modal competition. The STB is only allowed to regulate prices when a shipper does not have alternative transport options and the rate asked by the rail company is considered unreasonable. Less than 30% of freight traffic is subject to regulation (ECMT, 2001).

In the US, access to facilities owned by other companies is arranged through the use of voluntary contractual arrangements. The arrangements are negotiated freely between the parties. The STB does not regulate access terms; it only requires that they are not discriminatory. Access agreements can take many forms. The most important are the following (Jahanshahi, 1998):

a. **Trackage rights.** One railway uses the tracks of another railway and pays an agreed charge based on tonnage or a fee for service;

b. **Paired track arrangements.** When railways have parallel tracks and agree to integrate them into a double track line, thus increasing operating efficiency;

c. **Joint track arrangements.** When the track is co-owned by two or more railways. Dispatch and maintenance are shared but each operator is responsible for scheduling its own trains;

d. **Joint subsidiaries.** When a new company is established to undertake track maintenance and some operations;

e. **Joint facilities.** The railway is owned by two or more operators;

f. **Reciprocal operating agreements.** When one operator provides rolling material and other provides infrastructure services. The track owner receives a fee;

g. **Reciprocal switching.** A mutual exchange of wagons from one line to another;

h. **Detours.** The use of tracks of another railway due to unforeseen events, such as natural disasters and derailments;

i. **Pooling.** Aggregation of several rail operators to serve large industries; and
j. **Haulage and car-handling contracts.** One operator pulls the wagons of another company when it is not permitted or economically justified to run a separate train.

It is worth mentioning that the non-regulated nature of access arrangements implies that competition is not introduced in all segments of the market. In some circumstances, for example, although there are many competing railways, there is only one track to the final destination. As a matter of fact, Jahanshahi (1998) reports that there have been a number of disputes between energy companies and railways transporting coal regarding the access charges for the use of these essential facilities.

Amtrak also has the right to operate over tracks owned by freight companies under negotiated access agreements. This allows the company to use 24,000 additional miles of track through the payment of access rights to around 20 freight companies. Access charges are based on a formula based on avoidable costs that comprises a cost for gross tonnage and speed, linked to incentive and penalty payments. It is worth mentioning that the future of Amtrak is uncertain. The company requires around US$ 20 billion to repair its network and the current administration has eliminated its operational subsidy of US$1.2 billion for 2006 without restructuring plans been approved (The Economist, 2002; 2005a).

### 10.5.3 Australia

The structure of Australian railways follows the economic and demographic patterns of the country. Long distances, scarce population concentrated in the east, west and south coasts; and a significant mineral production, favor the development of a freight network. In fact, over 50% of the freight transported in the country is via rail. Coal represents almost 70% of the weight transported by the system. Unlike Europe and US, inland waterways are not important for the Australian transport system. Most inter-modal competition comes from road (Productivity Commission, 1999).

The interstate network links the capital cities of all states and is administered by the Australian Rail Track Corporation (ARTC), a public corporation owned by the commonwealth government. There are two main corridors in this network: (i) north-south, connecting the states of Victoria, New South Wales and Queensland along the east coast; and (ii), east-west, from New South Wales to Western Australia along the southern coast of the country (see figure 10.4).
Railways began operating in Australia in the 1850s, linking ports, cities and rural areas. They were privately owned and operated until the early 1900s, when the bankruptcy of several of these companies prompted state governments to take over the industry. Within state jurisdictions, vertically-integrated, public-owned monopolies operated the systems. Differences in standards and regulations led to a lack of integration of the state networks. In the 1970s, the commonwealth government tried to consolidate the diverse state systems, but only two state governments agreed to sell their railways.

Besides network fragmentation, legal restrictions to compete with road freight and inconsistent cost recovery policies between rail and road resulted in the decline of the industry. This poor performance led to commonwealth and state governments to promote a large-scale reform of the industry in the 1990s. One if the major difficulties in addressing the industry’s problems, however, was that railways are primarily a state responsibility. Decisions about the public or private provision of services and licensing of operators are made at state level. Hence, regulatory reform was aimed primarily at (i) harmonizing the different regulatory regimes; and (ii), standardizing the infrastructure. In 1995, the gauge of all interstate tracks was standardized through a substantial investment made by the commonwealth government (ECMT, 2001).

65 Source: Australasian Railway Association (2005)
A wide range of structural, ownership and access policies were introduced by the states during the 1990s. Some states vertically separated the provision of infrastructure and train services and others privatized parts of the network. In Queensland, train services are still provided by an integrated state-owned company.

Table 10.4 shows the structure and ownership of Australian railways after the reform.

<table>
<thead>
<tr>
<th>Jurisdiction</th>
<th>Train Services Provider</th>
<th>Infrastructure Services Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commonwealth</td>
<td>Private</td>
<td>Government</td>
</tr>
<tr>
<td>New South Wales</td>
<td>Private</td>
<td>Government</td>
</tr>
<tr>
<td>Victoria</td>
<td>Private</td>
<td>Government</td>
</tr>
<tr>
<td>Queensland</td>
<td>Government</td>
<td>Government</td>
</tr>
<tr>
<td>Western Australia</td>
<td>Private</td>
<td>Private</td>
</tr>
<tr>
<td>South Australia</td>
<td>Private</td>
<td>Private</td>
</tr>
<tr>
<td>Tasmania</td>
<td>Private</td>
<td>Private</td>
</tr>
</tbody>
</table>


Access to rail infrastructure is governed by the provisions of the National Access Regime. The basis for access to interstate rail infrastructure in Australia is the ARTC’s access undertaking, accepted by the ACCC in 2002. According to this undertaking, access charges are based on published reference tariffs which can be negotiated according to the particular needs of the service providers. Charges are structured as a two-part tariff. One component is levied on a fee per kilometer and is specific to each train service type and segment. The second component is related to distance and weight, and levied on a fee per gross tonne s per kilometer. The published reference access charges are adjusted annually using the RPI-X methodology. Access prices are also subject to floor-ceiling revenue limits. The floor is given by the incremental cost of ARTC providing a service, while the ceiling is the full economic cost of providing access to a certain segment of track, including the costs specific to a service, depreciation and an allocation of indirect costs, and a return on assets employed. Access to intrastate rail networks is nonetheless regulated through state-based regimes. These regimes vary in the independence of the arbitrator, the transparency of the arbitrator’s decision, the scope to appeal against decisions, pricing principles and the mechanisms used to resolve disputes. The regimes are similar in that non-price access terms are in all cases determined through a process of negotiation and arbitration (Productivity Commission, 2001).

As in other countries, entry has been modest despite the implementation of the access regime. Two companies have entered the national market (both providers of inter-modal services); although there has been additional entry for niche markets and also
from non-rail companies providing their own transport services. According to some estimates, these companies account for 15% of the total market (Fagan, 2007).

10.6 Lessons for a port access regime

Decades after the nationalization of the industry, changes in markets and technology that favor other transport modes and the difficulties faced by organizations structured as monopolies to adapt to these circumstances, caused the crisis that questioned the way rail services were provided. Different approaches were taken to deal with the problem. In the UK, the network is mostly used for passenger traffic, which makes it operationally different from those in US and Australia. The approach was to privatize the infrastructure separated from the provision of services and to grant Railtrack a monopoly to provide infrastructure. The provision of passenger services is franchised through auctions to non-related train operators which receive subsidies from the Strategic Rail Authority. Freight services were to be supplied in competition with other transport modes. It seems that the investment needed to upgrade the network and maintain reasonable safety standards was severely underestimated. This might have led the government to acknowledge that under these circumstances, the provision of infrastructure services also requires financial support from the public treasury. In fact, higher access charges and thus, higher train services’ prices, would have encouraged passengers to use road transport, which does not fully internalize the negative externalities it causes. In Australia, where the problem was the lack of integration, the approach taken by the commonwealth and state governments was to harmonize regulatory frameworks and to create a single corporation to operate the interstate infrastructure. In the US, the approach taken by the government was to deregulate the industry to facilitate intra- and inter-modal competition. To avoid passenger services to disappear, a public corporation was created to take over the market.

Table 10.5 presents the characteristics of the rail access regimes implemented in the UK, US and Australia.
Governments and port regulators can draw several lessons from the reform of the rail industry. First, that the decisions regarding the port sector’s vertical structure take into account the country’s geographical and economic particularities, since they have a large influence in the degree of competition that will arise: its geographical features, its production/consumption patterns, the characteristics of the existing infrastructure, etc. As the history of the rail industry shows, there cannot be a “one-size fits all” approach to reform, since a model that works well under certain circumstances may have negative consequences under different ones. In fact, the vertical structure is different in each of the analyzed countries.

The second main lesson comes from the US experience, which indicates that even in industries where competition is strong, regulation may still be needed in some segments. As mentioned before, in this country the regulator still sets prices for 30% of the freight (EMCT, 2001). In the port industry, where terminals face different degrees of competition, this implies that services for some cargoes will need to be regulated while others are liberalized.

The third main lesson is that access may not occur despite the characteristics of the implemented regime. As Pittman (2005) argues, the lack of entry is an indicator that the economies of scale, scope and density enjoyed by integrated operators are so large that prevent competition from non-integrated companies. This is relevant for ports because they also present economies of scale and scope (Heaver, 1975; Tovar, Jara-Diaz and Trujillo, 2004 and 2004). In these circumstances, the absence of an effective competition to discipline the incumbent might require regulators to intervene regulating not only the monopolistic segments of the logistics chain (such as infrastructure provision) but even potentially-competitive markets (shipping and nautical services). In these cases, the conditions offered by the few entrants to their customers may serve as a benchmark for regulating service conditions.

Two valuable lessons for regulators in the port industry can be drawn from the analysis of access pricing regimes in the rail industry. The first is the confirmation of the trend consisting in the use of price caps when the scope for negotiation is small.

<table>
<thead>
<tr>
<th>Vertical structure</th>
<th>UK</th>
<th>US</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separation for interstate; mixed for intrastate networks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access pricing</td>
<td>Regulated, with some scope for negotiation in freight</td>
<td>Freely negotiated</td>
<td>Negotiable reference charges</td>
</tr>
<tr>
<td>Non-price terms and conditions</td>
<td>Regulated, with some scope for negotiation in freight</td>
<td>Freely negotiated</td>
<td>Negotiable reference terms</td>
</tr>
<tr>
<td>Mechanism to expand the infrastructure</td>
<td>Regulated. Only expansions that are considered necessary are remunerated.</td>
<td>Market incentives</td>
<td>Regulated. Only expansions that are considered necessary are remunerated.</td>
</tr>
</tbody>
</table>
The second is that policies aimed at encouraging the use of a particular transport mode may use access charges as a tool, but it will require a close coordination between the government and the regulator. One policy that would facilitate this coordination is having a single regulator in charge of all transport modes, such as in Peru and Australia.
Part III: An access regime for the port industry
11. Lessons for formulating access policies in the port industry

More than two decades of reforms in network industries allow drawing useful lessons on how to promote competition by implementing access policies. As seen in previous chapters, even though the economics and reform drivers differ from industry to industry and from country to country, the access problem is common to all of them.

Section 11.1 discusses the similarities between the economic characteristics of port and network industries. This comparison is helpful for determining which of the access policies implemented in network industries are more suitable for port terminals. Section 11.2 summarizes the lessons learned from implementing access policies in network industries. Section 11.3 uses these lessons to determine the characteristics that an access regime for naturally monopolistic port terminals should have. Finally, the appendix at the end of the chapter analyzes the characteristics of access regimes in each of the studied countries. This analysis, even though complementary, is useful to determine if the characteristics of some of the discussed access policies can be attributed to particular conditions to a broader regulatory philosophy.

11.1 Similarities among ports and network industries

Table 11.1 shows the economic characteristics of the analyzed industries.

As discussed in chapter 6, network industries share several features. One of the most notorious characteristic is that final services (a call, electricity or gas delivery, the completion of the transport logistics chain) are composed by inputs produced in both competitive and monopolistic markets. In these circumstances, a regulated company providing integrated services has incentives to discriminate against competitors in the competitive markets, thus producing the access problem.
Table 11.1: Economic characteristics of ports and network industries

<table>
<thead>
<tr>
<th>Ports</th>
<th>Telecommunications</th>
<th>ESI</th>
<th>Natural Gas</th>
<th>Railways</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inputs produced in competitive and monopolistic markets are needed to complete the port's transport chain</td>
<td>Inputs produced in competitive and monopolistic markets are needed to complete one call</td>
<td>Inputs produced in competitive markets and monopolistic activities are needed to deliver electricity</td>
<td>Inputs produced in competitive markets and monopolistic activities are needed to deliver natural gas</td>
<td>Rail services are composed by inputs produced in both monopolistic and competitive markets</td>
</tr>
<tr>
<td>Ports are multi-service entities: different services are produced for different markets</td>
<td>Telecommunication networks support different markets: local, long-distance, mobile, etc.</td>
<td>Electricity is the only final good produced in the industry</td>
<td>Natural gas is the only final good produced in the industry</td>
<td>Rail companies produce different services for different users and markets</td>
</tr>
<tr>
<td>Asset indivisibility: ports can only be expanded in discrete amounts of relative large minimum sizes. Yards, berths or breakwaters are considered sunk investments.</td>
<td>Some transmission and distribution technologies require large capital investments. Fixed-wire networks involve sunk costs.</td>
<td>Some generation technologies require large capital investments. Transmission and distribution involve large sunk costs.</td>
<td>Activities require large capital investments (much of them sunk)</td>
<td>Asset indivisibility: infrastructure can only be expanded in discrete amounts. A large part of total investments are sunk.</td>
</tr>
<tr>
<td>Economies of scale and scope</td>
<td>Economies of scale, scope and density</td>
<td>Economies of scale and density</td>
<td>Economies of scale and density</td>
<td>Economies of scale, scope and density</td>
</tr>
<tr>
<td>Positive and negative externalities</td>
<td>Positive externalities: the network’s value increases for the existing users when a new subscriber is added</td>
<td>Positive and negative externalities</td>
<td>Positive and negative externalities</td>
<td>Positive externalities: the network’s value increases for the existing users when a new link is added</td>
</tr>
<tr>
<td>Competition largely depends on factors beyond port’s management control: existence of suitable shores, port’s location, size of the regional economy, among others</td>
<td></td>
<td></td>
<td>Production and competition depends on the location of geological basins in relation with consumption centers</td>
<td></td>
</tr>
</tbody>
</table>
Unlike network industries this is not necessarily the case in the port industry. As explained in chapter 5, naturally monopolistic terminals arise only in ports whose traffic is too small to allow building a second terminal (due to sub-additivity of costs). But when they arise, the resulting access problem is similar to that of network industries: an integrated terminal operator has incentives to discriminate against service providers in related markets (stevedoring, storage, towage) to control the whole transport chain and recover profits foregone by regulation.

It can be seen in table 11.1 that all the analyzed industries exhibit economies of scale. Ports, telecommunications and railways also exhibit economies of scope, which allows providing several services more cheaply than if they were produced by independent producing units. Moreover, network industries also exhibit economies of density, arising from the fact that it is cheaper to add a new link than building a new network. In the case of railways, economies of density arise from the fact that it is cheaper to add a new wagon to an existing train than operating a new one.

Likewise, ports produce positive and negative externalities. Positive, when the assets considered public goods are used by other users; and negative, when vessel traffic produces congestion or environmental pollution. The externalities produced by the other industries are of a different nature. In telecommunications and railways, positive ones are produced over existing networks when a new subscriber or rail link is added to the network. In the ESI, since electricity systems need to be balanced all the time, actions from one generator may have consequences on the remaining generating plants or the transmission network.

It can also be seen in table 11.1 that an important characteristic shared among port terminals, telecommunications and railway networks is that their infrastructure supports several markets. For example, long-distance and mobile services in telecommunications; shipping, cargo handling, pilotage and others in port terminals, and freight and passenger services in railways. The ESI and the natural gas industry, on the contrary, deliver just only one good.

Port terminals also share with railways the characteristic that their infrastructure can only be expanded in discrete amounts of relatively large minimum sizes. Moreover, a large part of the investments is considered sunk, since they have few alternative uses. Something similar occurs in the natural gas industry, were production and transport requires large capital investments. In telecommunications and the ESI, only some technologies and fixed-wire networks require large capital investments and involve sunk costs.

An interesting feature shared between the port industry and natural gas is that exogenous factors may have a large impact on competition. Indeed, as explained in chapter 9, the location of consumption centers relative to the producing wells may allow competition between pipelines even if economies of scale or density are not exhausted. As discussed in section 5.1.3, the degree of competition in the port industry largely depends on factors such as the location of competing terminals, the
existence of suitable shores (to build new terminals), the size of the regional economy, among others.

11.2 Lessons learned from implementing access policies

This section presents the lessons port regulators and policy makers can learn from the formulation and implementation of access policies in network industries. As explained in chapter 5, formulating access policies involves making decisions or setting rules regarding four main issues: vertical structure, access pricing, non-price terms and the incentives used to expand the infrastructure.

Regulators have at least two theoretical options for each of them, as shown in table 11.2.

<table>
<thead>
<tr>
<th>Option 1</th>
<th>Option 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical structure</td>
<td>Separation</td>
</tr>
<tr>
<td>Access pricing</td>
<td>Regulation</td>
</tr>
<tr>
<td>Non-price terms and</td>
<td>Regulation</td>
</tr>
<tr>
<td>conditions</td>
<td>Negotiation</td>
</tr>
<tr>
<td>Mechanism to expand the</td>
<td>Regulation</td>
</tr>
<tr>
<td>infrastructure</td>
<td>Incentives</td>
</tr>
</tbody>
</table>

In reality, however, regulators may opt for intermediate or mixed approaches, as allowing integration for freight but enforcing separation for passengers (as in US railways) or setting charges jointly with the incumbent in consultation with other operators (as in UK telecommunications).

11.2.1 Vertical structure

As explained in chapter 3, the importance of vertical structure lays on the fact that integration or separation gives incumbents totally different incentives whether to grant access to third parties or not. Indeed, integration gives firms controlling monopolistic markets incentives to seek deterring competition in non-monopolistic ones. This behavior is facilitated by the information asymmetry that characterizes the interaction between regulator and regulated firm (as discussed in chapter 2) and the large information costs involved in defining the terms of non-discriminatory access to infrastructure facilities (as discussed in chapter 5). As seen in previous sections, many access terms and conditions involve subtle issues that make discrimination and the creation of entry barriers difficult to identify, such as service timetabling, technical
specifications for equipments, minimal amounts of insurance cover, personnel qualifications requirements, among others. Therefore, implementing access policies under vertical integration requires significant regulatory capabilities in terms of experience and financial resources\(^\text{66}\).

On the other hand, under vertical separation the incumbent has no incentives to discriminate, since it cannot provide the service in the competitive market directly or through a related company. Under this structure, the more profitable scenario is to charge for access to as many companies as possible. Nonetheless, even though this behavior reduces the regulatory burden, it also increases transaction costs. And as railways illustrate, in the presence of economies of scope these transaction costs can be so large that impede entry.

The vertical structure of network industries in the analyzed countries is shown in table 11.3. It can be seen that policies regarding the vertical structure of telecommunications, ESI and natural gas seem to be consistent across countries: integration is allowed for the first, but separation is enforced between monopolistic and competitive activities in the remaining two. In railways, UK opted for separation, US for integration and Australia for an intermediate approach.

| Table 11.3: Vertical structure of network industries |
|-----------------------------------|-----------------------------------|-----------------------------------|
|                                   | UK                                | US                                | Australia                                       |
| Telecomm.                         | Integration allowed for long-distance; legal separation required for mobile; LLU for internet services | Integration allowed for long-distance to incumbents opening their networks; legal separation required for mobile; LLU for internet | Integration allowed for long-distance and mobile; LLU for internet services; sharing for data services |
| ESI                               | Legal separation between transmission, distribution and other activities. Integration allowed for generation and retailing. | Operational separation; incentives given for divestiture and the formation of ISOs and RTOs. | Separation. Integration subsists among some public-owned companies. |
| Natural Gas                       | Legal separation. Integration is allowed between transport, storage and distribution; and between production and retailing. | Operational separation. Integration is allowed between transport and storage. | Legal separation. Integration is allowed between transport, storage and distribution; and between production and retailing |
| Railways                          | Vertical separation               | Integration                        | Separation for interstate; mixed for intrastate networks |

The first lesson that port regulators can learn from the experiences of network industries is that the decision regarding the vertical structure of an industry has to be the consequence of a realistic estimation of the potential for competition in a

\(^{66}\) Defilippi (2004) estimates that the costs of regulating monopolistic ports in Peru is equivalent to allowing a concessionaire to obtain monopolistic rents in excess of 6% above a normal return.
particular market. As UK experience in regulating the natural gas industry shows, the cost for consumers of allowing integration when separation was needed, is high. A related lesson is that the market power of incumbents cannot be underestimated. They will take advantage of any flaws in regulation to discourage competitors to enter related markets.

The second lesson is that despite the many criticisms, integration constitutes a workable option to organize an industry characterized by having both monopolistic and non-monopolistic markets. Indeed, as seen in chapter 7, integration is the preferred form of vertical organization in the telecommunications industry, despite the fact that incumbents enjoy large amounts of market power (much of it inherited from the time when they were publicly owned). A related lesson for port regulators is that even in cases where a single operator is allowed to provide several services, some form of separation might be needed. Moreover, the fact that either operational or legal separation is implemented in a context of vertical integration indicates that the cost of implementing these options are not large; or, if large, that these costs are overcome by a more transparent operation or an improved regulation.

The third main lesson concerning vertical structure is that separation seems to be necessary when there is the need to create a new market. As explained in chapters 8 and 9, the main goal of the reforms carried out in the industries of electricity supply and natural gas was to create a market for these commodities, regulating only the monopolistic segments. But in both industries, the creation of markets necessarily requires the separation of the provision of monopolistic from non-monopolistic services, as the experience of the UK natural gas industry indicates. In this country, the incumbent was able to enjoy monopolistic rents until it was forced to separate by the Gas Act of 1995.

The fourth main lesson regarding vertical structure is that, as the experiences of the telecommunications and the railways industries show, the creation of an access regime is not enough to ensure entry to non-monopolistic markets, regardless of the vertical structure of the industry. This is more evident in the railways industry, where entry has been very limited under both vertical integration (The Netherlands, Germany, Switzerland) and separation (UK, Sweden). Therefore, regulators should be prepared to regulate prices and conditions or closely monitor markets that in theory, are competitive. This is the case of internet (where LLU has not encouraged entry as expected) or rail freight services.

11.2.2 Access pricing

As in many other public policy matters, selecting the mechanism to set access prices involves a trade-off between the advantages and disadvantages of each option. Indeed, setting access prices or the methodology to estimate them through a regulatory decision may avoid the delays characteristic of negotiation, but it has a higher risk of
regulatory failures (due to the information asymmetry between regulator and regulated firm and other problems discussed in chapter 2). On the other hand, allowing the parties to negotiate may reduce the risk of regulatory failure (since the information asymmetry between incumbent and access seeker is smaller than the one between regulator and incumbent), but negotiations take time and there is the risk that the terms of the agreement may result favoring the incumbent. Negotiation has also limits. Indeed, as seen in previous chapters access arrangements include a large number of issues that need to be determined, and this complexity could induce incumbents to strategically delay their completion.

Table 11.4 shows how access prices are set in the studied industries and countries.

<table>
<thead>
<tr>
<th>Industry</th>
<th>UK</th>
<th>US</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telecomm.</td>
<td>Set jointly by the regulator and the incumbent in consultation with other operators</td>
<td>Set at state level, although parties are free to negotiate other rates</td>
<td>Negotiated</td>
</tr>
<tr>
<td>ESI</td>
<td>Transmission charges set annually using LRMC. Total revenues set using RPI-X. Connection charges are negotiated.</td>
<td>Set by transmission owners according to guidelines set by the Ferc.</td>
<td>Regulated using price-caps (RPI-X methodology)</td>
</tr>
<tr>
<td>Railways</td>
<td>Regulated, with some scope for negotiation in freight</td>
<td>Freely negotiated</td>
<td>Negotiable reference charges</td>
</tr>
</tbody>
</table>

The main lesson regarding access pricing is that there is a strong relationship between the option to set access charges (regulation or negotiation) and the methodology to determine them (cost-based or price caps). In countries and industries where negotiation is encouraged (telecommunications in US and Australia, for example), the tendency is to use a cost-based methodology. This occurs even in situations when the room for negotiation is small, such as in the UK’s electricity supply and railways industries. On the other hand, in industries where vertical separation is mandated, such as in the natural gas industry and the ESI in US and Australia, the tendency is toward the regulator setting access charges using price caps.

The main reason behind this finding seems to be the incentives each pricing methodology gives the regulated company. When RPI-X is used, the company is given an incentive to reduce costs (and achieve productive efficiency) because it is allowed to retain the difference between the regulated price and their real costs. In this sense, price caps are superior to cost-based methodologies, under which the company does not have incentives to reduce costs, since all incurred ones are passed on to the users (Bernstein and Sappington, 1998). For this mechanism to work, however, long horizons are needed: price reviews are usually performed every four or five years, for
which it does not seem to be a useful methodology for a negotiation-arbitration framework designed to provide incentives to potential entrants to negotiate their access terms.

It is worth commenting the option used by the UK telecommunications regulator to set access prices and the problems it would pose for port regulators, especially in the developing world. In this case, negotiation does not occur between incumbent and access seeker but between incumbent and regulator, in a process where the regulator receives advise from other operators. Even though this may seem and interesting alternative, it seems to be a very difficult one to adopt in developing countries. Indeed, to achieve superior results than a negotiation between incumbent and access seekers, this option requires the regulator to know better than the access seekers, for example, the characteristics of the demand and the costs involved for all the technologies used. This implies, in fact, that the regulator is in a position to overcome the information asymmetry that characterizes its interaction with regulated firms. This option also requires having a regulator that is totally immune from the influence of interest groups. As explained in chapter 2, one of the main causes of regulatory failure is the regulator making biased decisions because it has been “captured” by interest groups (Guasch and Spiller, 1999). Having the regulator involved in direct negotiations with one of the parties not only facilitates the occurrence of this failure but also changes its role as the industry’s arbitrator.

The second main lesson regarding access prices concerns the use of negotiable reference tariffs. As explained in chapter 9, the Productivity Commission of Australia found during an inquiry that the fact that reference tariffs would prevail in case of a dispute discouraged negotiation and led to de facto regulation. To remedy this, the Commission recommended to establish more specific pricing principles and allowing the use of methodologies different to cost-based ones and to adopt a monitoring rather than a regulatory approach when the benefits of regulation are not clear (Productivity Commission, 2004).

### 11.2.3 Non-price access terms

As seen in chapter 5, the purpose of creating a regulatory framework to set non-price access terms is to avoid an integrated incumbent creating entry barriers; but also to prevent excess entry. As in the previous section, non-price access terms can be either set by the regulator or allow parties to negotiate to determine them.

Table 11.8 shows how non-price access terms are regulated in the network industries of UK, US and Australia.
<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>US</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telecomm.</td>
<td>Regulated via “standard interconnect agreement”</td>
<td>Set at state level, but parties are free to re-negotiate them.</td>
<td>Main ones regulated via “standard access obligations”. Others are negotiated.</td>
</tr>
<tr>
<td>ESI</td>
<td>Main terms are regulated by a Code. Others are negotiated.</td>
<td>For integrated utilities, regulated minimum terms. In organized markets, set by the system operator.</td>
<td>Regulated by the National Electricity Code.</td>
</tr>
<tr>
<td>Natural Gas</td>
<td>Regulated. Set in the Network Code</td>
<td>Set by the pipeline owner but consistent with regulated guidelines</td>
<td>Negotiable reference terms established using regulated guidelines</td>
</tr>
<tr>
<td>Railways</td>
<td>Regulated, with some scope for negotiation in freight.</td>
<td>Freely negotiated</td>
<td>Negotiable reference terms</td>
</tr>
</tbody>
</table>

There are three main lessons port regulators can learn from the analyzed experiences. The first is that the mechanism to set non-price access terms tends to be the same as for setting access charges, but not necessarily. In the Australian telecommunications industry, for example, even though access charges are set through negotiation, the main non-price access terms are regulated. Only the remaining ones are negotiated.

The second lesson regards the difference between the theory and the practice of negotiation when setting non-price access terms. As explained in chapter 5, negotiation is used as a mean to overcome the information asymmetry existing between the regulator and the regulated firm. In theory, access seekers have better information than the regulator regarding the real costs of providing services, for which, to avoid regulatory failures, the regulator should only intervene if negotiations between the regulated company and the access seeker fail. However, the analysis carried out in previous chapters indicates that the information asymmetry does not prevail in all non-price aspects. Indeed, as shown by the access regimes in the telecommunications industry, at least some of the non-price access terms can be set outright (arguably, the most obvious ones, or those that have been set in previous access agreements), thus reducing the scope for negotiation and shortening the process.

The third lesson is analogous to the previous one. Under an access regime that favors setting non-price access terms through regulation, the practice indicates that at least some terms can be left to be determined though negotiation, such as in the UK’s railways and electricity supply industries. Terms left for negotiation are those where the information asymmetry between regulator and regulated firm is larger. For example, in the UK’s ESI, entrants have to negotiate with the incumbent which works will have to be undertaken for connecting to the transmission grid. Trying to regulate this in advance is of little help, since the nature and costs of these works will depend on the particular conditions of the entrant’s infrastructure. In this case, negotiation would lead to better results.
11.2.4 Mechanism to expand the infrastructure

In competitive markets, companies have incentives to expand their infrastructure to meet demand; otherwise, they might lose clients to competitors offering better services. Monopolists, however, do not have these incentives. In fact, as explained in chapter 2, it is more profitable for them to restrict supply until the point where marginal costs equal marginal revenue. Regulated monopolists face similar incentives than non-regulated ones. They will not expand their infrastructure unless they face incentives of some sort; especially if they are vertically integrated and the infrastructure will also be used by other operators in related competitive markets. For these reasons, regulators need to decide what mechanism will be used to ensure that the infrastructure will be expanded. They have two options: they can either order the incumbent to expand the infrastructure or give him economic incentives to do it when necessary.

It is worth recalling what was discussed in chapter 5 regarding the form under which the private sector participates in the port industry. Unlike telecommunications, electricity supply and natural gas industries, where privatization through the sale of public assets is the rule, port terminals are transferred to the private sector through several arrangements which may require or not the expansion of the infrastructure. Concessions, Greenfield Projects and Divestitures generally require private companies to expand the infrastructure, whereas Management Contracts and Leases do not (World Bank, 2007). More importantly, out of the 299 port projects with private participation carried out in developing countries between 1990 and 2006, the large majority (276) belong to the first group; i.e., require private companies to expand the terminal’s infrastructure when needed (PPIAF, 2007).

The main difficulty with opting for ordering the monopolist to expand the infrastructure is the already discussed information asymmetry regarding demand and the incumbent’s costs. Because of that, the regulators may lack information to decide whether the expansion (or what kind of expansion) is needed. The second difficulty with this option arises from the regulatory capture’s problem: the regulator’s decision may be biased because of influence from an interest party, such as the government, consumers or the regulated company itself. It also incentives incumbents to spend resources in unproductive activities (such as lobbying) trying to convince the regulator that their planned investments are, in effective, necessary (Lasheras, 1999).

The option of giving the incumbent economic incentives to expand the infrastructure has also disadvantages, as discussed in chapter 5. Financial rewards allow incumbents to obtain monopolistic rents; market incentives are only suited for incumbents facing competition, and subsidies only work for unprofitable services. As we will see, regulators tend to use mixed approaches.
Table 11.6 shows the incentives given to incumbents to expand the infrastructure in the analyzed network industries.

| Table 11.6: Mechanism to expand the infrastructure in network industries |
|-----------------------------|-----------------------------|-----------------------------|
|                            | UK                          | US                          | Australia                    |
| Telecomm.                   | Universal service obligations | Universal service obligations | Universal service obligations |
| ESI                         | Regulated. Only expansions that are considered necessary are remunerated | In organized markets, the system operator decides over expansions. Other operators face market incentives. | Regulated. The regulator will only remunerate investments passing the "Regulatory Test". |
| Natural Gas                 | Regulated. Ofgem set expansion targets. | Market incentives | Market incentives, but the regulator can order expansions if requested by access seekers |
| Railways                    | Regulated. Only expansions that are considered necessary are remunerated. | Market incentives | Regulated. Only expansions that are considered necessary are remunerated. |

The first lesson for port regulators is that the decision regarding the expansion of the infrastructure in monopolistic markets cannot be left to the monopolist. As one can see in Table 11.6, only in industries where infrastructure owners compete against each other (natural gas and railways in the US) the decision is left to the incumbent, who can decide whether to expand the infrastructure according to market conditions. In all the remaining cases, the decision is made or authorized by the regulator, with the sole exception of the US ESI (in Organized Markets, the decision is made by the System Operator).

The second lesson is that regulators in network industries have implemented two policies to deal with the problem of expanding the infrastructure, each with important shortcomings.

a. The first is to allow an independent planner to decide when infrastructure should be expanded, as in the US Organized Markets. In theory, independent system operators have enough information to make opportune and demand-adjusted decisions. However, as discussed in section 8.4, i) these bodies may lack incentives to make the right decisions, since they do not bear the cost of their mistakes; and, ii) they tend to have complex governance structures (Arizu, Dunn and Tenenbaum, 2001).

b. The second option is to allow the asset owner to propose expansions but only approve those that are considered necessary, which is the method used in the electricity supply and railways industries in UK and Australia. However, as argued by Parker (2002), this option is prone to regulatory failures due to the existing information asymmetries between the regulator and the regulated firm.
The third lesson regarding the mechanisms used to expand the infrastructure is that the use of auctions may constitute a useful tool to allocate installed capacity, but it does not provide adequate incentives to expand the infrastructure. Moreover, as the UK’s natural gas industry experience indicates, allowing the incumbent to keep even a small percentage of the proceeds introduces incentives that are perverse, since the more acute the infrastructure scarcity is, the higher the revenues it receives. For these reasons, auctions should be used together with other incentives that penalize the incumbent for not expanding the infrastructure. The UK natural gas regulator combines long-term options with exposing the incumbent to buy-back costs by forcing it to repurchase the required capacity at spot prices\footnote{When capacity is scarce, spot prices are much higher than long-term ones (Chan, Laplagne, and Appels, 2003).}.

The following box summarizes the lessons learned from analyzing the access policies implemented in network industries.
Box 1: Access Policies Implemented in Network Industries: Lessons for the Port Industry

**Vertical structure**

1. The decision regarding the vertical structure of an industry has to be the consequence of a realistic estimation of the potential for competition in a particular market. As experience shows, the cost for consumers of allowing integration when separation is needed, is high.

2. Despite the many criticisms, integration constitutes a workable option to organize an industry where monopolistic and non-monopolistic segments coexist. However, experience shows that even in cases where a single operator is allowed to provide several services some form of separation might be needed.

3. Vertical separation seems to be necessary when there is the need to create a new market. Indeed, the main goal of the reforms carried out in the industries of electricity supply and natural gas was to create a market for these commodities, regulating only the monopolistic segments. But in both industries, the creation of markets necessarily requires separating the provision of monopolistic from non-monopolistic services.

4. As the experiences of the telecommunications and the railways industries show, the creation of an access regime is not enough to ensure entry to non-monopolistic markets, regardless the vertical structure of the industry.

**Access pricing**

5. There is a strong relationship between the option to set access charges (regulation or negotiation) and the methodology to determine them (cost-based or price caps). In countries and industries where negotiation is encouraged (telecommunications in US and Australia, for example), the tendency is to use a cost-based methodology. In industries where vertical separation is mandated, the tendency is toward the regulator using price caps to set access charges.

6. The use of negotiable reference tariffs may discourage negotiation, as the experience of the Australian natural gas industry indicates.

**Non-price access terms**

7. The mechanism to set non-price access terms tends to be the same as for setting access charges, but not necessarily.

8. Under an access regime that favors setting non-price access terms through negotiation, at least some of the non-price access terms can be set outright, thus reducing the scope for negotiation and shortening the process.

9. Under an access regime that favors setting non-price access terms through regulation, at least some terms can be left to be determined though negotiation.

**Mechanism to expand the infrastructure**

10. The decision regarding the expansion of the infrastructure in monopolistic markets cannot not be left to the monopolist.

11. Regulators in network industries have implemented two policies to deal with the problem of expanding the infrastructure:
   a. To allow an independent planner to decide when infrastructure should be expanded (as in US Organized Markets).
   b. To allow the asset owner to propose expansions but only approve those that are considered necessary (as in the electricity supply and railways industries in UK and Australia).

12. The use of auctions may constitute a useful tool to allocate installed capacity, but it does not provide adequate incentives to expand the infrastructure. For this reason, auctions should be used together with other incentives that penalize the incumbent for not expanding the infrastructure.
11.3 Elements of an access regime for the port industry

The purpose of analyzing the access policies implemented in network industries is to use these experiences to propose an access regime for the port industry. According to the analysis carried out in previous sections, such a regime should have the following characteristics.

<table>
<thead>
<tr>
<th>Table 11.7: Characteristics of an Access Regime for the Port Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Element</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Vertical structure</td>
</tr>
<tr>
<td>Access pricing</td>
</tr>
<tr>
<td>Non-price terms and conditions</td>
</tr>
<tr>
<td>Mechanism to expand the infrastructure</td>
</tr>
</tbody>
</table>

11.3.1 Vertical structure

The preceding sections have shown that the trend in the telecommunications industry is to allow vertical integration with some form of operational or legal separation; whilst in the energy industries (ESI and natural gas), there is a clear tendency toward enforcing vertical separation. In railways, the tendency is not clear. Access regimes with both separation (UK, Sweden) and integration (Germany, The Netherlands) have been implemented and entry has been limited under both options. Which of these options is the more suitable for the port industry?

We have seen that the port terminals are similar to telecommunications networks in the sense that both support several markets. For example, pilotage, towage and others in port terminals; and long-distance and mobile services in telecommunications. The monopolistic segments of both the ESI and the natural gas industry are used to deliver just one commodity. Moreover, in these industries separation was the consequence of a necessity, since the newly created markets for these commodities could not work without separating the monopolistic segments from the competitive ones.

Another argument to be taken into account is that one of the main trends in the transport industry (not only in ports) is toward competition among integrated logistics chains (ADB, 2000). Ports (and terminals) constitute the most important link (node) in competing supply chains (Haralambides, 2007).

Under these considerations, integration is the most recommendable option of vertical structure for naturally monopolistic port terminals. However, this is only a
preliminary conclusion. As the experience of the UK natural gas industry shows, the consequences of erring in this aspect are dear for consumers. Indeed, as one can recall for chapter 9, regulators underestimated the incumbent’s potential for exploiting market power, for which integration allowed the incumbent to deter competitors despite the implementation of an open access regime. The market for natural gas could perform competitively only after separation was enforced.

For this reason, even though allowing integration seem to be a logical decision, port regulators still need to carry out adequate estimations of economies of scale and scope at particular ports. In this sense, the experience of the railways industry is exemplifying. As Ivaldi and McCullough (2004) found, estimation of economies of scope indicate that a fully integrated firm would have a 20-40% cost advantage over a vertically separated company. In these circumstances, very little entry would be expected regardless of the characteristics of the access regime.

11.3.2 Price and non-price access terms

The analysis carried out through this study shows that in industries where integration is allowed, there is a tendency toward the use of negotiation to determine price and non-price terms (telecommunications in US and Australia). The analysis also shows that in industries where integration is allowed, cost-based methodologies are used to determine access charges, even in situations when the room for negotiation is small (as in the UK’s electricity supply and railways industries).

Since integration is the most recommendable option of vertical structure for naturally monopolistic port terminals, negotiation constitutes the most recommendable option to determine price and non-price terms. In these circumstances, the regulator should only intervene when parties do not reach an agreement.

For the same reasons, a cost-based methodology should be used to determine access charges. However, as the experience of the Australian natural gas industry indicates, port regulators should refrain from establishing reference tariffs that would prevail in case of a dispute. Experience shows that this arrangement has discouraged negotiation and led to de facto regulation (Productivity Commission, 2004).

One of the main lessons of access policies implemented in network industries indicates that even in industries where non-price terms are set through regulation, some of these terms can be set outright by the regulator, especially those that are considered industry standards or where the information asymmetry between regulator and regulated firm is small. For example, the amount of insurance cover that the incumbent might require to third parties operating inside the terminal, the characteristics of the equipments to be used, the qualifications of the laborers, or the terminal’s queuing policy (as in the Australian natural gas industry).
11.3.3 Mechanism to expand the infrastructure

As the statistics of public-private partnerships in the port industry show, the large majority of projects carried out in developing countries require the operator to expand the terminal’s infrastructure when needed (PPIAF, 2007). Only leases and management contracts do not impose this requirement. The regulator’s problem is then to devise a mechanism to assure that the infrastructure will be expanded when needed. One of the main lesson from the experiences of network industries is that this decision should not be left to the incumbent.

The analysis shows that regulators in network industries have implemented two policies to deal with the problem of expanding the infrastructure. The first is to allow an independent planner to decide when infrastructure should be expanded (as in the US Organized Markets). The second option is to allow the asset owner to propose expansions but only approve those that are considered necessary (as in the electricity supply and railways industries in UK and Australia). Both options have important shortcomings that were discussed in chapter 11.

However, none of these options seem better than simply indicating the maximum congestion level a terminal can stand, which is the method used in the concession of Matarani port terminal, in Peru (Alcazar and Lovaton, 2005). In this case, congestion acts as “trigger”: once the maximum level is reached, the operator has to expand the terminal. Two considerations that make the asymmetry between regulator and regulated firm almost inexistent support this conclusion. First, unlike network industries there are few technical options to expanding a port terminal when occupancy reaches congestion levels. Second, the criterion that a terminal reaches congestion levels when it reaches between 60% and 70% of berth utilization is considered the norm in the industry (Nathan Associates, 2004). For these reasons, the most recommendable mechanism to determine when to require the incumbent to expand the terminal’s infrastructure is the use of “triggers”.

232
Appendix 4: National access regimes in the studied countries

The analysis of access regimes by country is a useful exercise because: (i) it allows one to identify national regulatory patterns; and (ii), to understand the extent to which the choice of options made to regulate access in a particular industry is influenced by a broader regulatory philosophy or national policy.

A4.1 United Kingdom

The scope and extent of the regulatory reforms undertaken in the UK’s allows extracting valuable lessons. Indeed, this was the first country to restructure its network industries at large scale. Reforms were based on the guiding principle that competitive markets yield more efficient results than government supply. As a result, state-owned monopolies were privatized and markets traditionally reserved to be served by public companies were opened to competition. The role of the government in network industries shifted from service provider to regulator and arbitrator. This approach was later followed by a number of countries around the world.

However, as discussed in previous chapters, some of the restructuring policies initially implemented in the UK were flawed, inconsistent or implicitly made assumptions that were later proven wrong. For example, the design of the Electricity Pool allowed the owners of mid-priced plants to collude or game the system, and the instrument to trade futures distorted the price of electricity. These and other flaws forced the government to replace the pool with NETA in 2001 (Newbery, 1999). Likewise, assessments about the contestability of certain markets were inconsistent. In the telecommunications industry, only one competitor was initially licensed to challenge the incumbent in each market; while in the natural gas industry the government trusted that competition could develop even if BG owned the transmission system and was the only purchaser of the gas produced in the North Sea.

In addition, some of the implemented policies were based on assumptions that were later proven wrong. For example, the division of BR’s freight services into six companies implicitly assumed that on-rail competition could be introduced even in the presence of sizeable economies of scope. At the time of privatization, nonetheless, three of these companies were sold to a single operator which later acquired two more of the privatized companies. Similarly, it was assumed that the amount of investments
needed to maintain and upgrade the rail system could be solely recovered through access charges. This assumption also ignored that higher rail prices would encourage passengers to shift to road traffic, thus creating more congestion and other related externalities.

Table A4.1 presents the characteristics of the access regimes currently in effect in the UK.

<table>
<thead>
<tr>
<th>Table A4.1: Access regimes in UK network industries</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Telecom</strong></td>
</tr>
<tr>
<td>Vertical structure</td>
</tr>
<tr>
<td>Access pricing</td>
</tr>
<tr>
<td>Non-price terms and conditions</td>
</tr>
<tr>
<td>Mechanism to expand the infrastructure</td>
</tr>
</tbody>
</table>

It seems that there is a trend toward vertically separated activities among the access regimes implemented in the UK. Although the approach is more flexible for telecommunications, transmission and other services should be provided by legally separated companies in both energy industries. In the rail industry, the operator of the network is also banned from providing train services.

A tendency can also be observed toward the regulation of access prices. Indeed, parties are allowed to negotiate only connection charges to the transmission grid and certain aspects of rail access charges for freight services. In the telecommunications industry, access charges were initially set after a negotiation involving BT, the regulator and representatives of the consumers, but since then are adjusted by the regulator using the RPI-X methodology.
The UK’s trend toward regulation is more evident regarding access terms and conditions. As a matter of fact, there is no scope for negotiation in any of the analyzed industries. Only in the ESI, parties can negotiate complementary agreements which establish the works required to connect to the transmission system.

As for the incentives used to expand the network, the tendency toward regulation is less clear. In telecommunications, incumbents with significant market power are obliged to meet universal service obligations; in the ESI and railways, expansions to be remunerated using access charges are subject to regulator’s approval; while in the natural gas industry, Ofgem complements regulated targets with economic incentives.

A4.2 United States

The US economic system has a large history of reliance on the private sector to deliver services traditionally supplied by state-owned companies in other countries. Indeed, while railways and telephone companies were nationalized in Australia and the UK, they remained private in the US.

Private provision of public services was nonetheless complemented by government regulation as early as 1887, when the Interstate Commerce Commission was created to counteract the market power of railway owners. Regulation was later extended to other network industries with the creation of the FCC and the Ferc. The responsibility for regulation of network industries in the US is shared by federal and state regulators. In industries such as railways and natural gas, tight or flawed regulation has been regarded as one of the main causes of their gradual decline between the 1950s and 1980s (OECD, 1998).

In telecommunications and electricity supply, on the contrary, it is the lack of regulation and authorities’ regulatory powers that was considered as the cause for markets not working properly. For example, the public status of many participants in the ESI and the legal constraints faced by federal and state regulators to mandate the unbundling of transmission and generation has limited the implementation of ISOs and RTOs (Brown and Sedano, 2003). As a result, some state PUCs have been obliged to provide financial incentives to integrated utilities to divest their assets. Likewise, the implementation of the Telecommunications Act by the FCC was legally challenged by the incumbents in a process that lasted six years.

Two aspects of the US access regimes are worth noting. The first is that the access framework in the natural gas industry encouraged the creation of a secondary market for capacity rights, which led to the development of a spot gas market. The second is the non-mandatory nature of the access arrangements in the rail industry, which implies that competition has not been introduced in all segments of the market.

Table A4.2 presents the characteristics of the access regimes in US network industries.
The characteristics of US access regimes show a tendency toward non-intervention. Only the ESI seems to be closely regulated, presumably because of its unique features discussed in previous chapters.

It can be seen that although regulatory criteria may differ among federal and state regulators, vertical integration is common. However, unbundling of electricity activities is encouraged and ring-fencing arrangements are imposed to enforce operational separation in both energy industries. Prices, terms and conditions for access are negotiated in all studied industries with the exception of the ESI. In the natural gas industry, reference tariffs for transmission are estimated by the pipeline owners themselves following Ferc’s methodologies and guidelines.

In gas and railways, networks expansion is guided by market incentives rather than by regulatory targets. In the ESI, this is the mechanism used by ISOs and RTOS, although systems not using these organized markets may follow different guidelines. In telecommunications, the coverage of high cost areas is assured through the imposition of universal service obligations. Further network expansions are guided by commercial incentives.

<table>
<thead>
<tr>
<th>Table A4.2: Access regimes in US Network industries</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Telecom</strong></td>
</tr>
<tr>
<td>Vertical structure</td>
</tr>
<tr>
<td>Access pricing</td>
</tr>
<tr>
<td>Non-price terms and conditions</td>
</tr>
<tr>
<td>Mechanism to expand the infrastructure</td>
</tr>
</tbody>
</table>

**A4.3 Australia**
As a late comer, Australia could reap from the reforming experiences of other countries, in particular the UK. The approach taken was to institute a general access regime that applies to all facilities of national importance, regardless of their ownership or industry to which they belong.

As seen in chapter 3, the National Access Regime establishes the criteria to bring facilities and services under access regulation, and subjects the determination of their prices, terms and conditions to a negotiation-arbitration framework. Although specific access regimes have been developed for the telecommunications, electricity and natural gas industries, these are consistent with the regulations of the National Access Regime.

Since Australia is politically organized as a federal government, regulatory responsibilities are shared among the commonwealth and state regulators, akin to the US system. For this reason, industry reforms required close coordination between national and state governments.

Non-economic factors such as the size of the population relative to the extension of the country and its distribution within the territory play a relevant role in Australian network industries. In the ESI, for example, although the interstate electricity market is called the National Electricity Market, it only comprises eastern and southern Australia. The natural gas industry developed in the form of isolated networks with little or no interconnection, and differences in standards and regulations led to a lack of integration of the state rail networks.

Table A4.3 presents the characteristics of the access regimes in Australian network industries. It can be seen that the Australian scope to regulate access in network industries is intermediate among those implemented in the US and the UK. Although integration is allowed in telecommunications, even for mobile services; separation is the preferred structure in the remaining industries, at least for private companies and interstate networks.

As in other countries, price and non-price conditions for the ESI are regulated. These elements are subject to negotiation (and arbitration) in the remaining industries with the exception of telecommunications, where non-price terms are regulated via standard access obligations.

The incentives used to encourage network expansion are not uniform across industries. In both energy industries, market incentives guide network expansions. But in the ESI, new assets will secure a return only if they pass a “regulatory test”; while in the natural gas industry regulators may mandate the expansion of a pipeline if it is justified by the requirements of a prospective user. In telecommunications, carriers are required to meet universal service obligations, while in the rail industry, regulators estimate access charges including only justifiable investments.
### Table A4.3: Access regimes in Australian network industries

<table>
<thead>
<tr>
<th>Vertical structure</th>
<th>Telecom</th>
<th>ESI</th>
<th>Natural gas</th>
<th>Railways</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration allowed for long-distance and mobile; LLU for internet services; sharing for data services</td>
<td>Separation. Integration subsists among some public-owned companies.</td>
<td>Legal separation. Integration is allowed between transport, storage and distribution; and between production and retailing</td>
<td>Separation for interstate; mixed for intrastate networks</td>
<td></td>
</tr>
</tbody>
</table>

| Non-price terms and conditions | Main ones regulated via “standard access obligations”. Others are negotiated. | Regulated by the National Electricity Code. | Negotiable reference terms established using regulated guidelines | Negotiable reference terms |

| Mechanism to expand the infrastructure | Universal service obligations | Regulated. The regulator will only remunerate investments passing the "Regulatory Test". | Market incentives, but the regulator can order expansions if requested by access seekers | Regulated. Only expansions that are considered necessary are remunerated. |
12. A model of access regime for the port industry

The aim of this chapter is to propose an access regime to regulate naturally monopolistic port terminals using the lessons learned from network industries. Section 12.1 presents the basic elements of the model (objectives, actors and elements). Section 12.2 describes its operation using the example of Matarani port terminal. Section 12.3 uses game theory to analyze the expected consequences of the proposed model on the contestability and efficiency of port-related markets.

12.1 Basic elements

12.1.1 Objectives

As explained in chapter 2, the aim of regulation is to eliminate the barriers that impede the functioning of competitive markets; or, if this is not possible, to replicate the discipline that market forces would impose on the regulated firm if they were present (Guasch and Spiller, 1999). To achieve these goals, regulation is based on two main policies: (i) **price regulation** (regulation of prices and quality of service), and; (ii) **access regulation** (regulation of how firms access the facilities they need to compete) (Defilippi and Flor, 2008).

The main objective of the proposed access model is to allow competition to occur in markets that otherwise would need to be price-regulated. By limiting regulatory intervention only to situations when it is effectively needed, this policy is less costly and less likely to spawn market distortions.
12.1.2 Parties

There are three parties:

a. **Regulator**: It can be a government agency specialized in monopoly regulation, such as Ositran (Peru), Superintendencia de Puertos y Transporte (Colombia), or the state regulators of Victoria\(^{68}\) and South Australia\(^{69}\); as well as empowered ministries such as Mexico’s Secretaría de Transporte, or competition agencies such as Australia’s ACCC, or the ICCC in Papua New Guinea.

b. **Incumbent**: A private firm that operates a terminal under any contractual or ownership arrangement. It can be a concessionaire, lessee or terminal owner.

c. **Entrant**: A private firm that would like to enter the market of any service necessary to complete the logistics chain. It can provide one or many services.

12.1.3 Assumptions

The model has the following assumptions:

a. The port has a single terminal, handling liquid bulk; dry bulk; break-bulk; and containerized cargo.

b. The terminal constitutes a natural monopoly for all cargo originated in or destined to its hinterland.

c. The terminal operator:

   - Is a private firm.
   - Is in charge of maintaining and expanding the terminal’s infrastructure.
   - Has the right to collect wharfage (to cargo) and berthage (to vessels) to pay for maintaining and expanding the infrastructure.

d. The port authority has normative and regulatory duties regarding traffic, safety and security issues.

e. Port services (stevedoring, warehousing, pilotage, towage, mooring, etc.) can be provided by the terminal operator, a related company or competing firms. There are no sizeable entry barriers.

\(^{68}\) The Essential Services Commission.

\(^{69}\) The Essential Services Commission of South Australia.
f. Cargo is handled using equipment that is not fixed to the infrastructure: mobile cranes and conveyor belts, forklifts, tipper trucks, yard trucks, loading shovels, and the ship’s own equipment. Containers are shipped in geared vessels.

12.1.4 Characteristics of the model

Consistently with the analysis and recommendations of section 11.3, the proposed access regime should have the following characteristics:

a. **Vertical structure**: Integration with legal or operational separation (similar to access regimes implemented in the telecommunications industry). The terminal operator is allowed to provide all the services needed to complete the logistics chain by itself (operational separation) or through related companies (legal separation).

b. **Access pricing**: Access charges result from a negotiation between the incumbent and entrant using a cost-based methodology. Similar to the arbitration-negotiation devised in the Australian National Access Regime, the regulator only intervenes if the parties do not reach an agreement.

c. **Non-price terms**: Non-price terms are also subject to negotiation between incumbent and entrant. The regulator sets the rules of negotiation and is expected to set some basic terms outright. As for access pricing, the regulator only intervenes if the parties do not reach an agreement.

d. **Mechanism to expand the infrastructure**: Use of “triggers”. Similarly to the system used in the UK’s natural gas industry, the regulator pre-announces the maximum congestion level and penalizes the operator if this level is surpassed. Under this arrangement, the operator has incentives to expand the infrastructure when needed.

Unlike network industries, in ports the number of service providers can be limited by safety, security or operational reasons. In this case, auctions would be used as a mechanism to select the operator. As discussed in chapter 2, Demsetz (1968) claims that it is possible to achieve optimal results by generating *ex ante* competition (the “Demsetz Approach”). In the absence of collusion, equal access to essential inputs and symmetric information among the bidders, the auction would make prices approach the average cost of the most efficient firm, thus minimizing simultaneously productive and allocative inefficiencies.

12.2 Operation of the model

Once the basic characteristics of the regime are decided, principles are needed to determine which services will be covered (stevedoring, warehousing, pilotage,
towage, mooring, etc.) and which facilities the operators should be granted access to (piers, wharves, docks, aprons, yards, transit and parking areas, etc.) To decide this, the model borrows two important characteristics of the Australian National Access Regime: the use of the Essential Facilities Doctrine (EFD) and the declaration of “Essential Services”.

As discussed in chapter 3, the National Access Regime applies to all infrastructures considered of “national importance”, regardless of the industry. The regime works by declaring these infrastructures “Essential Facilities” and granting access to parts of them according to the criteria of the EFD. So far, the regime has been applied to the network industries analyzed in this study as well as to port and airport terminals (Productivity Commission, 2001). It is important to take note that not all services provided with essential facilities are covered by the Australian regime. Only those that are needed to deliver the good or service provided by the industry in question are, and are listed by regulators as “Essential Services”.

In the port industry, “Essential Services” are only those that are needed to complete the logistics chain: pilotage, towage, mooring, stevedoring, etc. These are the services that a regulated monopolist would try to monopolize to recover profits foregone by regulation. But a service such as bunkering, for example, may not be declared an Essential Service if the regulator considers that ships can refuel in another location. In this case, competition alone would limit the incumbents’ ability to charge disproportionate prices and thus, to recover monopolistic rents. Regulation would not be warranted and thus should be avoided.

In this model, “Essential Facilities” are those the providers of “Essential Services” require access to. For example, to provide towage services, towboats require access to the basin and mooring areas. Their crews also require access to transit areas within the terminal. Since an integrated terminal operator might impede non-related towage companies provide services by restricting access to these facilities, the basin, as well as mooring and transit areas constitute Essential Facilities for the provision of towage services. Once a service is declared as Essential Service, the conditions to access Essential Facilities are regulated.

Table 12.1 shows the three types of market that exist in a port with a naturally monopolistic terminal: markets that are effectively monopolistic (and require price regulation); those that would be competitive if the incumbents’ ability to deter competition is limited (require access regulation); and those that do not require regulation (regular antitrust policies could be used to deal with anti-competitive behavior).
Table 12.1: Types of Markets in Ports with Monopolistic Terminals

<table>
<thead>
<tr>
<th>Type of market</th>
<th>Contestability</th>
<th>Regulation required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monopolistic</td>
<td>No</td>
<td>Price regulation</td>
</tr>
<tr>
<td>Competitive (but the incumbent has the ability to deter competition)</td>
<td>Yes, but access policies are needed to sustain competition</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Competitive (the incumbent does not have the ability to deter competition)</td>
<td>Yes. No regulatory safeguards are needed to sustain competition</td>
<td>No regulation</td>
</tr>
</tbody>
</table>

To distinguish among these three types of markets, the proposed model uses a sequential rationale that is depicted in figure 12.1 (Defilippi and Flor, 2008).

Figure 12.1: Rationale to Determine Regulatory Option

At first, the regulator asks itself if the market of a certain service is competitive (or potentially competitive) or not. If it is not competitive, price regulation is needed. If the market is indeed competitive (or potentially competitive), the regulator should ask itself if it corresponds to an Essential Service or not. If it does not, the service should not be regulated. But, if the market corresponds to an Essential Service, access can be granted through negotiation or auction. If the number of providers is not limited, negotiation may proceed. Otherwise, an auction must be called. This process has to be repeated for each service.

The operation of the model can be exemplified using the Matarani port as an example. As previously mentioned, this is a common-user port located in southern Peru. Its only terminal handles most types of cargo: dry bulk (grains and minerals), liquid bulk (sulfuric acid), break bulk and containerized cargo. Moreover, this terminal constitutes a natural monopoly for users located in southern Peru (Tamayo, Paredes and Flor, 1999).

Matarani port terminal was concessioned in 1999 to Romero group, one of the largest business conglomerates in Peru, with interests in several sectors of this economy\(^\text{70}\). The terminal’s concession contract allows the operator (a special-purpose company...

\(^\text{70}\) Besides logistics, the Romero group has interests in finance, manufacturing, agribusiness, fishing and retailing, among others.
named Tisur) to charge wharfage and berthage tariffs for the use of the infrastructure, the exclusive right to provide mooring services and the non-exclusive right to provide storage services. Tariffs for wharfage, berthage and mooring are regulated, as well as for some storage services: liquid bulk (sulfuric acid) and dry bulk (grains and mineral ore). Besides operating the terminal, Romero group provides other logistics services through related companies with various degrees of integration: shipping, shipping agency services, pilotage, towage, stevedoring, and cargo agency services, among others. This structure corresponds to vertical integration with legal separation. Other non-related firms provide these services as well (Alcazar and Lovaton, 2005). It is important to mention that the Matarani port terminal operates in an institutional and regulatory framework that complies with the assumptions made in section 12.1.3.

Table 12.2 presents the results of applying the rationale of the proposed model to the Matarani port terminal. One can see that the tariffs for wharfage, berthage and storage for liquid and dry bulk (grains and mineral ore) would still need to be regulated, since the terminal faces no competition in these markets (painted in green). On the other hand, the terminal faces competition for the provision of ancillary services such as bunkering, ship-store and waste collection, for which they do not require neither price nor access regulation (painted in yellow). One can also see that all of the remaining services have two common characteristics: (i) their provision is strictly necessary to complete the logistics chain, and; (ii) their providers require using the terminal facilities. These characteristics give the monopolist operating the terminal the incentives and the ability to try deterring competition in these markets. Therefore, they require to be covered by an access regime.
<table>
<thead>
<tr>
<th>Services provided at the terminal</th>
<th>Is their market competitive?</th>
<th>Essential Service?</th>
<th>Regulatory policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shipping</td>
<td>Yes</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Pilotage</td>
<td>Yes</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Towage</td>
<td>Yes</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Mooring</td>
<td>Yes</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Wharfage</td>
<td>No</td>
<td>Yes</td>
<td>Price regulation</td>
</tr>
<tr>
<td>Berthage</td>
<td>No</td>
<td>Yes</td>
<td>Price regulation</td>
</tr>
<tr>
<td>Stevedoring</td>
<td>Yes</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Storage for:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Containerized cargo(*)</td>
<td>Yes</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Dry bulk (grains)</td>
<td>No</td>
<td>Yes</td>
<td>Price regulation</td>
</tr>
<tr>
<td>Dry bulk (mineral ore)</td>
<td>No</td>
<td>Yes</td>
<td>Price regulation</td>
</tr>
<tr>
<td>Break bulk(*)</td>
<td>Yes</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Liquid bulk(*)</td>
<td>No</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
<tr>
<td>Ancillary services:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bunkering</td>
<td>Yes</td>
<td>No</td>
<td>No regulation</td>
</tr>
<tr>
<td>Ship-store</td>
<td>Yes</td>
<td>No</td>
<td>No regulation</td>
</tr>
<tr>
<td>Waste collection</td>
<td>Yes</td>
<td>No</td>
<td>No regulation</td>
</tr>
<tr>
<td>Ship repair</td>
<td>Yes</td>
<td>Yes</td>
<td>Access regulation</td>
</tr>
</tbody>
</table>

(*) Storage for containerized cargo and break bulk can be provided by warehouses located outside the port. Access is required for trucks carrying the cargo to and from the terminal.

12.3 Analysis\(^{71}\)

In this section, game theory will be used to analyze the access strategies of the terminal operator (the incumbent) and a potential entrant. This analysis will serve to assess the consequences of implementing the proposed access model on the contestability and efficiency of Essential Services’ markets.

Two further assumptions need to be made:

a. Both price and non-price access terms will be determined through negotiation.

b. If the parties do not reach an agreement, the regulator can enact a mandate setting both price and non-price access terms. Neither party can foresee the terms of the mandate, thus creating incentives for them to reach a “Nash equilibrium”\(^{72}\).

---

\(^{71}\) This section is based on Flor and Defilippi (2003)

\(^{72}\) A “Nash equilibrium” is a situation in which neither party has incentives to change their strategies (Vega Redondo, 2000).
Figure 12.2 shows the possible outcomes of the interaction between incumbent and a potential entrant in the market of any Essential Service.

Table 12.3: Interactions between Incumbent and Potential Entrant

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Situation</th>
<th>Result</th>
<th>Profits (incumbent, challenger)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Competitive market</td>
<td>Market returns</td>
<td>(r, r)</td>
</tr>
<tr>
<td>b</td>
<td>Challenger monopolizes the market</td>
<td>Monopolistic rents for challenger</td>
<td>(0, r+)</td>
</tr>
<tr>
<td>c</td>
<td>Contestable market</td>
<td>Market returns</td>
<td>(r, 0)</td>
</tr>
<tr>
<td>d</td>
<td>Incumbent monopolizes the market</td>
<td>Monopolistic rents for incumbent</td>
<td>(r+, 0)</td>
</tr>
<tr>
<td>e</td>
<td>Competitive market</td>
<td>Market returns</td>
<td>(r, r)</td>
</tr>
<tr>
<td>f</td>
<td>Price war</td>
<td>Returns below market levels</td>
<td>(r-, r-)</td>
</tr>
<tr>
<td>g</td>
<td>Contestable market</td>
<td>Market returns</td>
<td>(r, 0)</td>
</tr>
<tr>
<td>h</td>
<td>Incumbent monopolizes the market</td>
<td>Monopolistic rents for incumbent</td>
<td>(r+, 0)</td>
</tr>
</tbody>
</table>

Where:
- r  market returns
- r+ Returns above market
- r- Returns below market

Before the proposed access model is implemented, the incumbent is the exclusive supplier of the essential service. The incumbent can assume an active or passive attitude. Once the regime has been implemented, potential entrants may decide whether to enter the market or not, regardless of the attitude of the incumbent.
Since the incumbent is vertically integrated, it can grant access to third parties on fair terms or try to discriminate in favor of its related firm. In the first case, the parties may easily reach an agreement. In the second case, the incumbent risks the regulator mandating access in terms that are less convenient than those that could be obtained through negotiation. In addition, it faces the possibility of being fined for denying or delaying access unrightfully.

It can be seen that the incentives created by the access regime make the markets for essential services potentially more competitive than before. Indeed, as a result of the interaction between incumbent and entrant, eight different outcomes can be reached, out of which five result in a competitive market (a, c, e, f and g, shaded in yellow). However, only four of them constitute a “Nash equilibrium” (Vega Redondo, 2000). But as a result of a price war (outcome f), both parties would be obtaining below-market returns, for which it is likely that they will continue playing until one of them is driven out of the market or a competitive equilibrium (r, r) is reached.

It is important to mention that in the absence of entry barriers, it is unlikely that a price war would result in a predatory situation (Tarzijan and Paredes, 2001). Even if the incumbent could temporarily reduce the price of a service below its cost, it would probably be unable to recover its losses after expelling competitors because a later increase in prices would encourage the entry of new ones (or reentry of the firms that had left the market).

An interesting result of the implementation of the proposed model is that the increase in contestability may generate a situation in which a service is provided by just one supplier, but at competitive prices (c and g). This outcome is produced by the difficulty of increasing margins caused by the threat of new entrants, which disciplines the provider and produces competitive profits.

The outcomes b, d and h will lead to either party obtaining returns above market levels. In the absence of entry barriers, the most likely outcome is that these returns will attract more competitors to the market, which will interact until a competitive equilibrium is reached. Only then neither player would have incentives to change their strategies.

It is also worth mentioning that the possibility of collusion between the incumbent and competitor is also present. But such arrangements are very hard to enforce in absence of entry barriers, for which such outcome is unlikely to occur (Ivaldi et al., 2003).

From this analysis, the following can be concluded:

a. The implementation of the proposed access regime is likely to reduce entry barriers, increase contestability and introduce competition in markets for port services. Five out of eight possible outcomes are clearly competitive, and the

247
above-the-market returns obtained in the remaining three are expected to attract competitors until the rents disappear.

b. The strengths of the proposed regime are that it makes markets more contestable, has low transaction costs and reduces the cost of regulation. By implementing the rule that the regulator intervenes only when a market solution is not attainable, the regime avoids unwarranted regulation. In countries with history of discretionary regulatory policies, this rule may encourage private investment in port infrastructure.

c. The model has three important weaknesses:

- The presence of multiple operators may increase coordination costs, especially in small ports.
- The procedures required to implement the proposed access regime may be lengthy, which may increase transaction costs and thus, reduce the potential benefits of implementing the regime.
- It requires the regulator to be credible. In developing countries this requirement is related to the prevailing political environment and may make the proposed regime difficult to implement.
13. Conclusions

Although usually located far from main maritime routes and poorly endowed with transport infrastructure, developing countries urgently need to increase trade to provide jobs and improve the living conditions of their populations. Despite problems caused by mismanaged privatization processes and learning costs, an increasing number of governments are concessioning their infrastructure to be operated by the private sector, so they can concentrate their efforts and scarce funds in other fields, such as education and health.

However, due to the economic characteristics of port operations (economies of scale, scope and density, and lumpy investments) some port terminals constitute natural monopolies for carriers and shippers located in their hinterlands (Flor and Defilippi, 2003). Unfortunately, regulation of monopolies is still an unknown discipline in the port sector. Initially developed to deal with monopolies in network industries such as telecommunications and electricity supply, its adaptation to transport has been limited. Indeed, the only one known to have established port regulators are Australia, Colombia and Peru. The further development of the discipline would allow governments to involve the private sector in the operation and construction of monopolistic terminals without the fear of harming carriers and shippers.

The objective of this thesis is to analyze the characteristics of access policies implemented in the telecommunications, electricity supply, natural gas and railways industries, and to use the lessons learned from these experiences to propose a model suitable for the port industry.

The problem of access arises in industries where inputs from monopolistic and competitive markets are complementarily needed to provide a service, or in the case of ports, to complete the logistics chain. In these circumstances, the firm controlling the monopolistic segment has incentives to deter competition in the competitive segments (markets) to recover profits foregone by regulation (Paredes, 1997). To avoid such situations from occurring, governments have two options. They can either (i) forbid integration between terminal operators and providers of port services or, (ii) establish a framework under which all service providers are allowed to access and use
the terminal under similar conditions. As suggested by Vickers (1995), the first option (vertical separation) may create non-trivial transaction costs that result in higher prices for the consumers, for which the second option (the implementation of access policies) constitutes a more desirable policy.

This research analyzes access policies implemented in the network industries of three countries: UK, US and Australia. These countries were chosen because they have approached the access problem from different perspectives, and because they are considered best practice cases among regulation practitioners (ADB, 2000). The UK was the first developed country that carried out a comprehensive privatization program, which required formulating access policies without the benefit of previous experiences. As seen in the previous chapters, this lack of experience led UK regulators to underestimate the monopolists’ ability to deter competition despite the implementation of open access policies. The US, on the other hand, has a long tradition of encouraging competition and allowing the private supply of public services. In this country, the reform of network industries focused more in restructuring than in changing the nature of their ownership. The Australian strategy, on the other hand, consisted on implementing the National Access Regime whose provisions cover all relevant infrastructures (regardless of the nature of their ownership and the industry they belong).

Part I

Chapter 2 presented the basic elements of regulation theory: monopoly pricing, regulation rationale and regulatory failures. It also introduced the five basic dilemmas a regulator faces when addressing the problems caused by the existence of monopolies. The first one is that in the presence of sub-additive costs, productive efficiency is achieved by having only one firm supplying the market; but the consequent lack of competition will cause allocative inefficiency instead. The second dilemma is that although the use of Ramsey prices may assure economic efficiency, these are not necessarily subsidy-free, sustainable or socially equitable. The third dilemma a regulator faces is between using auctions that generate ex ante competition but introduce contractual rigidities, or a more flexible regulation that may better address unexpected problems but also generates higher social costs. The fourth dilemma is created by the information asymmetry between the regulator and the regulated firm. The regulator may encourage the regulated firm to reveal its costs, but to do so it will need to offer an economic incentive that generates an informational rent. Finally, the regulator must choose between using rate-of-return and RPI-X as pricing methodologies. Both work well under certain circumstances but may cause undesired effects and social costs if the choice is wrong.

Chapter 3 discussed various topics related to the problem of introducing competition in naturally monopolistic industries: the convenience of vertical separation, access pricing and the Essential Facilities Doctrine (EFD). This chapter showed that although vertical separation facilitates regulation, it may cause diseconomies of scope or
coordination. It also discussed the diverse methodologies of access pricing and the problems found when implementing them. Finally, this chapter analyzed the EFD from an economic point of view and explained its role in the Australian National Access Regime. It showed that the application of the EFD can be seen as a trade-off between static and dynamic economic efficiency. On one hand, is desirable to limit the incumbent’s property rights to avoid the loss of welfare caused by the reduction of competition in at least one market. On the other hand, this limitation on property rights also reduces the incumbent’s incentives to invest, innovate and improve the coverage of infrastructure.

Chapter 4 analyzed the evidence on the effects of privatization processes around the world. It started by discussing the reasons argued by governments to embark on privatization programs and the extension of these programs in the world. It also reviewed diverse assessments on efficiency, tariffs and welfare, concluding that the majority support the idea that privatization had brought net positive effects to the implementing countries. Finally, the chapter reviews privatization assessments in the port industry. Chapter 4 concluded that that successful processes are those complemented by sound regulatory policies, thus supporting the idea that better regulatory policies are needed to deal with monopolies in the port industry.

Chapter 5 provided an overview of the basics of port economics, discussed the main drivers for port reform around the world, and illustrated the processes undertaken in three developing countries. Its aim was to demonstrate that privatization could be counterproductive in countries where ports are naturally monopolistic unless an access regime is implemented. It also discusses the main characteristics that such access regime should have and that will serve as guidelines to analyze access regimes in network industries.

**Part II**

Part I showed the complexities involved in formulating access policies. Fortunately, regulators in network industries have already decided over these issues. The main features of these access regimes and the lessons learned from their implementation provide valuable inputs for the design of access policies for the port industry. The aim of part II is to summarize these lessons and use them to propose an access regime for naturally monopolistic port terminals.

This part starts with chapter 6, which described the economic characteristics of network industries and their relationship with diverse market structures. The discussion focused on the main economic characteristics of network industries (externalities and complementarities; economies of scale, scope and density; compatibility and standards) and how they may lead to market structures that do not necessarily result from lack of competition or anti-competitive practices.

Chapters 7 to 10 analyzed the access regimes implemented in telecommunications, electricity, natural gas and rail industries. To facilitate comparisons, they followed the
same structure. With the aim of contextualizing the analysis, the economic characteristics of the industry and its reform trends were first presented. This was followed by a discussion on the access arrangements typically implemented in the industry; and a study of the UK, US and Australian cases. Each chapter concluded with a section summarizing the lessons that can be drawn for designing an access regime for naturally monopolistic port terminals.

**Part III**

The aim of part III was to propose an access model for naturally monopolistic port terminals using as inputs the lessons drawn from analyzing access regimes implemented in network industries. Chapter 11 aimed at drawing lessons from network industries in order to propose an access regime for naturally monopolistic port terminals. It started by discussing the similarities between the economic characteristics of ports and the analyzed industries. The second section presented the lessons learned from implementing access policies in network industries, while the last section used these lessons to propose a model of access regime for the port industry.

The lessons learned from implementing access policies in network industries are shown in the following box:
Lessons Learned from Implementing Access Policies in Network Industries

Vertical structure
1. The decision regarding the vertical structure of an industry has to be the consequence of a realistic estimation of the potential for competition in a particular market. As experience shows, the cost for consumers of allowing integration when separation is needed, is high.
2. Despite the many criticisms, integration constitutes a workable option to organize an industry where monopolistic and non-monopolistic segments coexist. However, experience shows that even in cases where a single operator is allowed to provide several services some form of separation might be needed.
3. Vertical separation seems to be necessary when there is the need to create a new market. Indeed, the main goal of the reforms carried out in the industries of electricity supply and natural gas was to create a market for these commodities, regulating only the monopolistic segments. But in both industries, the creation of markets necessarily requires separating the provision of monopolistic from non-monopolistic services.
4. As the experiences of the telecommunications and the railways industries show, the creation of an access regime is not enough to ensure entry to non-monopolistic markets, regardless the vertical structure of the industry.

Access pricing
5. There is a strong relationship between the option to set access charges (regulation or negotiation) and the methodology to determine them (cost-based or price caps). In countries and industries where negotiation is encouraged (telecommunications in US and Australia, for example), the tendency is to use a cost-based methodology. In industries where vertical separation is mandated, the tendency is toward the regulator using price caps to set access charges.
6. The use of negotiable reference tariffs may discourage negotiation, as the experience of the Australian natural gas industry indicates.

Non-price access terms
7. The mechanism to set non-price access terms tends to be the same as for setting access charges, but not necessarily.
8. Under an access regime that favors setting non-price access terms through negotiation, at least some of the non-price access terms can be set outright, thus reducing the scope for negotiation and shortening the process.
9. Under an access regime that favors setting non-price access terms through regulation, at least some terms can be left to be determined through negotiation.

Mechanism to expand the infrastructure
10. The decision regarding the expansion of the infrastructure in monopolistic markets cannot not be left to the monopolist
11. Regulators in network industries have implemented two policies to deal with the problem of expanding the infrastructure:
   a. To allow an independent planner to decide when infrastructure should be expanded (as in US Organized Markets).
   b. To allow the asset owner to propose expansions but only approve those that are considered necessary (as in the electricity supply and railways industries in UK and Australia).
12. The use of auctions may constitute a useful tool to allocate installed capacity, but it does not provide adequate incentives to expand the infrastructure. For this reason, auctions should be used together with other incentives that penalize the incumbent for not expanding the infrastructure.

According to these lessons an ideal access regime for naturally monopolistic terminal should have the following characteristics:
1. Port terminals are similar to telecommunications networks in the sense that both support several markets. Moreover, in these industries separation was the consequence of a necessity, since the newly created markets for these commodities could not work without separating the monopolistic segments from the competitive ones.

Another argument to be taken into account is that one of the main trends in the transport industry (not only in ports) is toward competition among integrated logistics chains (ADB, 2000). Ports (and terminals) constitute the most important link (node) in competing supply chains (Haralambides, 2007). Under these considerations, integration is the most recommendable option of vertical structure for naturally monopolistic port terminals.

2. Since integration is the most recommendable option of vertical structure for naturally monopolistic port terminals, negotiation constitutes the most recommendable option to determine price and non-price terms. In these circumstances, the regulator should only intervene when parties do not reach an agreement.

For the same reasons, a cost-based methodology should be used to determine access charges. However, as the experience of the Australian natural gas industry indicates, port regulators should refrain from establishing reference tariffs that would prevail in case of a dispute. Experience shows that this arrangement has discouraged negotiation and led to de facto regulation (Productivity Commission, 2004).

3. The analysis shows that regulators in network industries have implemented two policies to deal with the problem of expanding the infrastructure. However, none of these options seem better than simply indicating the maximum congestion level a terminal can stand. In this case, congestion acts as “trigger”: once the maximum level is reached, the operator has to expand the terminal. For these reasons, the most recommendable mechanism to determine when to require the incumbent to expand the terminal’s infrastructure is the use of “triggers”.

Chapter 12 proposed an access regime to regulate naturally monopolistic port terminals using the lessons learned from network industries. The main objective of the proposed access model is to allow competition to occur in markets that otherwise would need to be price-regulated.

The model has the following assumptions:

a. The port has a single terminal handling liquid bulk, dry bulk, break-bulk and containerized cargo.
b. The terminal constitutes a natural monopoly for all cargo originated in or destined to its hinterland.

c. The terminal operator:
   • Is a private firm.
   • As the only user of the port’s infrastructure, it is in charge of maintaining and expanding it.
   • Has the right to collect wharfage (to cargo) and berthage (to vessels) to pay for maintaining and expanding the infrastructure.

d. The port authority has normative and regulatory duties regarding traffic, safety and security issues.

e. Port services (stevedoring, warehousing, pilotage, towage, mooring, etc.) can be provided by the terminal operator, a related company or competing firms. There are no sizeable entry barriers.

f. Cargo is handled using equipment that is not fixed to the infrastructure: mobile cranes and conveyor belts, forklifts, tipper trucks, yard trucks, loading shovels, and the ship’s own equipment. Containers are shipped in geared vessels.

The model borrows two important characteristics of the Australian National Access Regime: the use of the Essential Facilities Doctrine (EFD) and the declaration of “Essential Services”. In the port industry, “Essential Services” are only those that are needed to complete the logistics chain: pilotage, towage, mooring, stevedoring, etc. These are the services that a regulated monopolist would try to monopolize to recover profits foregone by regulation.

Game theory was used to analyze the access strategies of the terminal operator (the incumbent) and a potential entrant. From this analysis, the following can be concluded:

a. The implementation of the proposed access regime is likely to reduce entry barriers, increase contestability and introduce competition in markets for port services. Five out of eight possible outcomes are clearly competitive, and the above-the-market returns obtained in the remaining three are expected to attract competitors until the rents disappear.

b. The strengths of the proposed regime are that it makes markets more contestable, has low transaction costs and reduces the cost of regulation. By implementing the rule that the regulator intervenes only when a market solution is not attainable, the regime avoids unwarranted regulation. In countries with history of discretionary regulatory policies, this rule may encourage private investment in port infrastructure.
c. The model has three important weaknesses:

- The presence of multiple operators may increase coordination costs, especially in small ports.

- The procedures required to implement the proposed access regime may be lengthy, which may increase transaction costs and thus, reduce the potential benefits of implementing the regime.

- It requires the regulator to be credible. In developing countries this requirement is related to the prevailing political environment and may make the proposed regime difficult to implement.
Alcazar, Lorena y Rodrigo Lavation (2005): “Evaluación de la Concesión del Puerto de Matarani: ¿Quién ganó y quién perdió?” Documento de Trabajo 47. Grade, Lima


Australasian Railway Association Inc (2005): "Australian Rail Network". www.ara.net.au


Boyer, Marcel (1997): “Competition and Access in Telecoms: ECPR, Global Price Cap, and Auctions”. Center for Interuniversity Research and Analysis on Organizations Université de Montréal


Charles River Assoc Ltd. (2002): “Port Companies and Market Power - A Qualitative Analysis”. Wellington


Crandall, Robert and Gregory Sidak (2002): "Is Structural Separation of Incumbent Local Exchange Carriers Necessary for Competition?" Yale Journal on Regulation, 19:335-411


De la Cruz, Ricardo and Raúl García (2003): "La Problemática de la Actividad de Transmisión de Energía en el Perú: Algunas Opciones de Política". CIES, Lima


Economides, Nicholas (1996): "The Economics of Networks". International Journal of Industrial Organization, 14: 673-699


Economides, Nicholas and Lawrence White (1994): "Networks and Compatibility: Implications for Antitrust". European Economic Review, 38: 651-662


Foxley, Juan and José Luis Mordones (2000): "Port Concessions in Chile". Public Policy for the Private Sector Note # 223, October. The World Bank, Washington DC.


*Review of Industrial Organization*, 16: 251-266


Harding, Alan and Jan Hoffmann (2002): “Trade between the CARICOM and CACM countries: The role to play for ports and shipping services” Working paper presented in a conference at Ciudad del Saber, 12 November 2002, Panama


Ivaldi, Marc; Bruno Jullien; Patrick Rey; Paul Seabright y Jean Tirole (2003): “The Economics of Tacit Collusion”. Final report for DG Competition, European Commission


Lapuerta, Carlos, Juan Benavides and Sonia Jorge (2003): "Regulation and Competition in Mobile Telephony in Latin America". Paper presented at the First
Meeting of the Latin American Competition Forum. OECD/Inter-American Development Bank


Nash, Chris and Cesar Rivera-Trujillo (2004): "Regulatory Reform in Europe - Principles and Practice". Paper presented at the STELLA Focus Group 5 synthesis meeting, Athens, June 2004


Network Rail (2005): "Company Information". www.networkrail.co.uk


The Economist (2005b): "Argentina. Full Steam Ahead into the Dark" 16th June 2005


Tovar, Jara-Diaz and Trujillo (2004): "Econometric Estimation of Scale and Scope Economies Within the Port Sector: A Review". City University Economics Discussion Papers No. 07/04


Yap, Wei Yim and Jasmine Lam (2006): “Competition dynamics between container ports in East Asia”. Transportation Research Part A 40:35–51

Summary

The problem of access arises in industries where inputs from monopolistic and competitive markets are complementarily needed to provide a service. In these circumstances, the firm controlling the monopolistic segment has incentives to deter competition in the competitive segments (markets) to recover profits foregone by regulation (Paredes, 1997). In the port industry, for example, a number of services need to be provided to complete the logistics chain: pilotage, towage, stevedoring, storage, etc. Without any of these, cargo cannot be delivered. In ports where a terminal constitute a natural monopoly, an integrated terminal operator\(^\text{73}\) has incentives to deter competition in the markets of services that are necessary to complete the logistics chain, since this would allow him to charge disproportionate prices and recover monopolistic rents. This strategy can be implemented by preferential treatment to itself or sister companies, or by restricting competitors access to the terminal.

To avoid the situations from occurring, regulators have two options. They can either (i) forbid integration between terminal operators and carriers or, (ii) establish a framework under which all service providers are allowed to access and use the terminal under reasonable conditions. As suggested by Vickers (1995), the first option (vertical separation) may create non-trivial transaction costs that result in higher prices for the consumers, for which the second option (the implementation of access policies) constitutes a more desirable policy. However, formulating access policies is not an easy task. If access conditions are too high, a limited number of entrants will use the terminal, allowing providers to obtain economic rents. If conditions are too relaxed, an excess of entry may occur, thus reducing the terminal operator’s incentives to adequately maintain and expand the infrastructure (Laffont and Tirole 1994).

The objective of this thesis is to analyze the characteristics of access policies implemented in the telecommunications, electricity supply, natural gas and railways

\(^{73}\) A terminal operator that also provides shipping or logistics services (by itself or through related companies).
This research analyzes access policies implemented in the network industries of three countries: UK, US and Australia. These countries were chosen because they have approached the access problem from different perspectives, and because they are considered best practice cases among regulation practitioners (ADB, 2000).

The lessons learned from implementing access policies in network industries are the following:

**Vertical structure**

1. The decision regarding the vertical structure of an industry has to be the consequence of a realistic estimation of the potential for competition in a particular market. As experience shows, the cost for consumers of allowing integration when separation is needed, is high.

2. Despite the many criticisms, integration constitutes a workable option to organize an industry where monopolistic and non-monopolistic segments coexist. However, experience shows that even in cases where a single operator is allowed to provide several services some form of separation might be needed.

3. Vertical separation seems to be necessary when there is the need to create a new market. Indeed, the main goal of the reforms carried out in the industries of electricity supply and natural gas was to create a market for these commodities, regulating only the monopolistic segments. But in both industries, the creation of markets necessarily requires separating the provision of monopolistic from non-monopolistic services.

4. As the experiences of the telecommunications and the railways industries show, the creation of an access regime is not enough to ensure entry to non-monopolistic markets, regardless the vertical structure of the industry.

**Access pricing**

5. There is a strong relationship between the option to set access charges (regulation or negotiation) and the methodology to determine them (cost-based or price caps). In countries and industries where negotiation is encouraged (telecommunications in US and Australia, for example), the tendency is to use a cost-based methodology. In industries where vertical separation is mandated, the tendency is toward the regulator using price caps to set access charges.

6. The use of negotiable reference tariffs may discourage negotiation, as the experience of the Australian natural gas industry indicates.

**Non-price access terms**

7. The mechanism to set non-price access terms tends to be the same as for setting access charges, but not necessarily.
8. Under an access regime that favors setting non-price access terms through negotiation, at least some of the non-price access terms can be set outright, thus reducing the scope for negotiation and shortening the process.

9. Under an access regime that favors setting non-price access terms through regulation, at least some terms can be left to be determined through negotiation.

**Mechanism to expand the infrastructure**

10. The decision regarding the expansion of the infrastructure in monopolistic markets cannot be left to the monopolist.

11. Regulators in network industries have implemented two policies to deal with the problem of expanding the infrastructure:
   a. To allow an independent planner to decide when infrastructure should be expanded (as in US Organized Markets).
   b. To allow the asset owner to propose expansions but only approve those that are considered necessary (as in the electricity supply and railways industries in UK and Australia).

12. The use of auctions may constitute a useful tool to allocate installed capacity, but it does not provide adequate incentives to expand the infrastructure. For this reason, auctions should be used together with other incentives that penalize the incumbent for not expanding the infrastructure.

According to these lessons, an ideal access regime for naturally monopolistic terminal should have the following characteristics:

1. Integration is the most recommendable option of vertical structure for naturally monopolistic port terminals.
2. Negotiation constitutes the most recommendable option to determine price and non-price terms. In these circumstances, the regulator should only intervene when parties do not reach an agreement. For the same reasons, a cost-based methodology should be used to determine access charges.
3. The most recommendable mechanism to determine when to require the incumbent to expand the terminal’s infrastructure is the use of “triggers”.
Samenvatting

In industrieën waar de inbreng van monopolistische en concurrerende markten nodig is om een dienst aan te kunnen bieden, speelt het probleem van *markttoegang*. In dergelijke omstandigheden heeft de monopolist een duidelijk motief om concurrentie in de segmenten (markten) met meerdere tegenspelers te ontmoedigen teneinde de nadelen van regulering te beperken (Paredes, 1997). Zo zijn in de havenindustrie verschillende diensten tezamen nodig om een volledige logistieke keten te realiseren: loodswezen, sleepdiensten, stuwadoors, opslag, enzovoorts. Als één van deze diensten ontbreekt, dan kunnen goederen niet worden afgeleverd. In havens met een natuurlijk monopolie voor de terminal heeft een geïntegreerde terminalbeheerder een duidelijk motief om concurrentie op het gebied van deze voor de logistieke keten noodzakelijke diensten te ontmoedigen. Op deze manier kan de beheerder disproportionele prijzen rekenen en de economische marge naar zich toetrekken. Een dergelijke strategie kan worden gerealiseerd in de vorm van een voorkeursbehandeling voor de eigen of zusterondernemingen, of door de toegang van concurrenten tot de terminal te beperken.

Regelgevers hebben twee mogelijkheden om dergelijke situaties te voorkomen. Ze kunnen de integratie tussen terminalbeheerders en rederijen verbieden of een raamwerk vaststellen voor de toegang tot en het gebruik van de terminal onder redelijke voorwaarden. Zoals voorgesteld door Vickers (1995), kan de eerste optie (verticale scheiding) leiden tot niet-triviale transactiekosten die resulteren in een hogere prijs voor de consument. De tweede optie, het invoeren van een toegangsregime, is daarom een aantrekkelijker beleid. Het is echter niet eenvoudig om een dergelijk toegangsregime te formuleren. Indien de toetredingsbarrières te hoog zijn, zal slechts een beperkt aantal toetreders de terminal gebruiken en kan de aanbieder economische marges naar zich toetrekken. Zijn de voorwaarden daarentegen te gemakkelijk gesteld, dan kan er te groot aantal gebruikers van de terminal gebruik willen gaan maken. In dat geval is er voor de beheerder van de terminal onvoldoende stimulans om de infrastructuur te onderhouden en uit te breiden (Laffont en Tirole, 1994).
Het doel van dit proefschrift is om de eigenschappen van toegangsregimes, zoals deze in de telecommunicatie-, electriciteits-, aardgas- en spoorwegindustrieën zijn ingevoerd, te analyseren en om de lessen uit deze industrieën te gebruiken teneinde een theoretisch model te ontwikkelen dat geschikt is voor de havenindustrie.

In dit onderzoek worden toegangsregimes die zijn ingevoerd in de netwerkindustrieën van drie landen (Verenigd Koninkrijk, Verenigde Staten van Amerika en Australië) onderzocht. Deze landen zijn geselecteerd omdat deze de toegangsregimes vanuit verschillende perspectieven hebben benaderd en omdat zij door regelgevingexperts als de beste praktijkvoorbeelden worden gezien (ADB, 2000).

Op basis van deze ervaringen komen we tot de volgende aanbevelingen voor een ideaal toegangsregime voor een haventerminal met een natuurlijk monopolie:

1. *Integratie* is de meest aan te bevelen keuze voor de verticale structuur voor een haventerminal met een natuurlijk monopolie.
2. Prijzen en andere condities kunnen het beste door middel van *onderhandelingen* tot stand komen. In dergelijke omstandigheden dient de regelgever alleen tussenbeide te komen indien de partijen niet tot overeenstemming kunnen komen. Om dezelfde reden dient een op kosten gebaseerde methodologie voor het bepalen van de toegangsprijzen te worden gebruikt.
3. Als mechanisme om de gevestigde partij aan te zetten tot het uitbreiden van de terminalinfrastructuur stellen wij het gebruik van zogenaamde *triggers* voor.
Resumen

El problema del acceso surge en industrias en las que la producción de un bien o servicios requiere complementariamente insumos producidos en mercados competitivos y monopólicos. En estas circunstancias, la firma que controla el segmento monopólico tiene incentivos para limitar la competencia en los mercados conexos con el fin de recuperar las rentas perdidas como consecuencia de la regulación. En la industria portuaria, diversos servicios son necesarios para completar la cadena logística: practicaje, remolcaje, estiba, almacenamiento, etc. En puertos en los que un terminal constituye un monopolio natural, un operador integrado tiene incentivos para limitar la competencia en los mercados mencionados, ya que ello le permitiría cobrar precios desproporcionadamente altos y recuperar las rentas monopólicas que la regulación le impide obtener.

Para impedir la aparición de situaciones como las descritas, los reguladores tienen dos opciones: (i) impedir la integración entre operadores de terminales y prestadores de servicios logísticos o, (ii) establecer un marco en el cual todos los prestadores de servicios logísticos pueden acceder a las instalaciones del terminal bajo condiciones razonables. Como sugiere Vickers (1995), la primera opción (separación vertical) tiene el potencial de crear costos de transacción que pueden resultar en mayores costos para los usuarios, por lo cual la segunda opción (la implementación de políticas de acceso) constituye una política deseable. Sin embargo, la formulación de políticas de acceso no es una tarea fácil. Si las condiciones de acceso son demasiado exigentes, sólo ingresará al mercado un número limitado de firmas, permitiendo a las empresas existentes obtener rentas económicas. Si las condiciones son muy laxas, ocurrirá un exceso de entrada que reducirá los incentivos del operador del terminal a mantener y expandir la infraestructura (Laffont y Tirole, 1994).

El objetivo de esta tesis es el de analizar las características de los regímenes de acceso implementados en las industrias de telecomunicaciones, electricidad, gas natural, y ferrocarriles, y utilizar las lecciones aprendidas de estas experiencias para proponer un régimen de acceso para la industria portuaria.

Las lecciones aprendidas son las siguientes:

**Estructura vertical**

1. Las decisiones sobre la estructura vertical de la industria deben ser consecuencia de una estimación realista del potencial de surgimiento de
competencia. Como lo demuestra la experiencia de la industria de gas natural del Reino Unido, el costo para los consumidores de permitir la integración cuando la separación es necesaria, es muy alto.

2. La segunda lección es que a pesar de las múltiples críticas, la integración vertical constituye una opción válida para organizar industrias caracterizadas por requerir insumos de mercados competitivos y monopólicos. La experiencia indica que el costo de estas políticas no es muy alto, o que éste es compensado por una operación más transparente o por una mejor regulación.

3. La separación parece ser necesaria cuando existe la necesidad de crear un mercado, como los de gas y electricidad, que requiere necesariamente la separación de las actividades monopolísticas de las competitivas.

4. Tal como lo demuestran las experiencias de las industrias de telecomunicaciones y ferrocarriles, la implementación de un régimen de acceso no garantiza la entrada a los segmentos competitivos.

Cargos de acceso
5. Existe una fuerte relación entre la opción para fijar cargos de acceso (regulación o negociación) y la metodología para determinarlos (basados en costos o precios tope). En países e industrias en donde se incentiva la negociación, la tendencia es a utilizar metodologías basadas en costos. En industrias en donde la separación vertical es obligatoria, la tendencia es que el regulador fije los cargos de acceso usando precios tope.

6. El uso de tarifas referenciales negociables que prevalezcan en caso de disputa desincentiva la negociación y lleva a una regulación de facto basada en costos.

Condicioness de acceso
7. Si bien las condiciones de acceso pueden ser determinadas usando el mismo mecanismo que para fijar el cargo, ello no tiene que ser necesariamente así. En la industria australiana de telecomunicaciones, por ejemplo, aunque los cargos de acceso son negociados, las principales condiciones del mismo están reguladas.

8. En regímenes que favorecen el uso de negociación para determinar las condiciones de acceso, al menos algunas de ellas pueden ser fijadas desde un principio, lo que reduce el ámbito de la negociación y acorta el proceso.

9. En regímenes que favorecen el uso de regulación para determinar las condiciones de acceso, la práctica indica que algunos términos pueden ser sujetos de negociación entre las partes.

Mecanismo para ampliar la infraestructura
10. La decisión de expandir la infraestructura no puede ser dejada en manos del monopolista. Sólo en las industrias en las que los operadores de la infraestructura enfrentan una competencia efectiva, éstos tienen la libertad de expandirla de acuerdo con los incentivos proporcionados por el mercado.

11. Los reguladores de las industrias de red implementan dos políticas para enfrentar el problema de la expansión de la infraestructura, ambos con importantes desventajas:
a. Permitir que la decisión sea tomada por un planificador independiente, como en “mercados organizados” de la industria eléctrica estadounidense.
b. Permitir que el operador de la infraestructura proponga las expansiones, pero sólo aprobar las que el regulador considera necesarias, como en las industrias de electricidad y ferrocarriles del Reino Unido y Australia.

12. El uso de subastas constituye una herramienta útil para asignar capacidad existente, pero no provee incentivos adecuados para expandir la infraestructura. Más aún, puede constituir un incentivo perverso, ya que el monto a recibir es mayor cuanto más aguda sea la escasez de infraestructura. Por ello, las subastas deben ser usadas conjuntamente con otras medidas que penalicen al monopolista por no llevar a cabo las expansiones necesarias.

De acuerdo con estas lecciones, un régimen de acceso ideal para terminales portuarios naturalmente monopólicos debería poseer las siguientes características:

1. La integración es la opción de estructura vertical más recomendable para terminales portuarios con características de monopolio natural.
2. La negociación constituye el mecanismo más recomendable para determinar cargos y condiciones de acceso. En estas circunstancias, el regulador sólo debe intervenir cuando las partes no alcanzan un acuerdo. Por las mismas razones, deberían usarse metodologías basadas en costos para determinar cargos de acceso.
3. El mecanismo más recomendable para determinar cuándo debe ser obligatorio para un operador expandir un terminal es el uso de triggers
About the Author

Enzo Defilippi (Lima, 1969) studied Economics at Pontificia Universidad Católica del Perú, where he graduated in 1993. In 2000, he received Master’s degrees in both Public Policy and Business Administration from IESA (Caracas, Venezuela).

After working several years as a financial analyst and public officer, he started his academic career in 2002. In that year, he co-authored a paper that obtained the “International Journal of Maritime Economics and Logistics Prize”. Since then, his work has been published in journals such as *Transport Research Part A: Policy and Practice* and *Maritime Economics and Logistics*, as well as in various research series at Peruvian universities. He has also developed a successful career as an economic consultant, having provided services to the European Commission, the International Finance Corporation, the Andean Community, USAID, GTZ, diverse Peruvian ministries and government agencies, and a large number of private companies.

He currently works in Lima as a researcher and consultant in regulation, antitrust, and public-private partnerships.


Access Regulation for Naturally Monopolistic Port Terminals

The access problem arises in industries where services are produced using inputs from both monopolistic and competitive markets. In the port industry, a number of services need to be jointly provided to complete the logistics chain: pilotage, towage, stevedoring, storage, etc. In ports where a terminal constitutes a natural monopoly, an integrated operator has incentives to deter competition from providers of the competitive services. To avoid situations from occurring, regulators have two options: forbidding integration or implementing access policies under which all service providers are allowed to use the terminal under reasonable conditions. The first option may create non-trivial transaction costs that result in higher prices for the consumers, for which the second option constitutes a more desirable policy.

Unfortunately, formulating access policies is not an easy task. If conditions are too high, a limited number of firms will enter the market, allowing existing ones to obtain economic rents. If conditions are too relaxed, an excess of entry may occur, thus reducing the operator’s incentives to adequately maintain and expand the terminal. This may produce dear consequences in developing countries, where there is urgency for updating and expanding port infrastructure.

The objective of this thesis is to propose a model to deal with the access problem in the port industry using as inputs the lessons learned in the telecommunications, electricity, natural gas and railways industries.