Toward an End-to-End Smart Grid: Overcoming Bottlenecks to Facilitate Competition and Innovation in Smart Grids

Johann J. Kranz
Ludwig-Maximilians-University Munich, Munich School of Management, Institute for Information, Organization, and Management. Ludwigstr. 28 VG/II, 80539 Munich, Germany, kranz@lmu.de

Arnold Picot
Ludwig-Maximilians-University Munich, Munich School of Management, Institute for Information, Organization, and Management. Ludwigstr. 28 VG/II, 80539 Munich, Germany, picot@lmu.de

June 2011
11-12

© 2011 National Regulatory Research Institute
Abstract

Policy makers, practitioners, and researchers are focusing more than ever on smart grid infrastructures due to energy systems’ impact on society and the economy. By integrating a communications and control system with the existing power grid, smart grids provide end-to-end connectivity which enables near to real-time data exchange among all actors and components in the electricity system’s value chain. Dependent on the smart grid communications network’s availability, the way electricity is generated, delivered, and consumed can be improved and optimized. Also, new services, applications, and technologies can emerge that will also aid to improve and optimize the use of electricity. End-to-end communication requires initially developing the missing communications link between consumers’ premises and the rest of the energy network (the “last mile”) by deploying an Advanced Metering Infrastructure (AMI), along with smart meters. Given the German metering and electricity markets’ characteristics—which is comparable to many liberalized markets—incumbent distribution system operators (DSOs) are likely to control the smart grid’s last mile. The last-mile infrastructure cannot be substituted or replicated within a reasonable time and cost frame. Moreover, together with the meter data, the infrastructure provides an essential input allowing efficient downstream markets, i.e., complementary services, products, and applications, to emerge. Such developments give rise to concerns about anti-competitiveness. This paper’s goal is to analyze whether such concerns are justified, since anticompetitive behavior would impede interoperability’s emergence, distort competition, and harm innovation and social welfare. The analysis shows that, contrary to the Chicago School’s rationale regarding vertical integration, DSOs have incentives to discriminate against new market entrants by leveraging entry barriers. We discuss possible regulatory remedies by building upon insights gained in telecommunications regulation. We also consider the implications for theory and regulation, and make recommendations for further research.

Online Access

This paper can be accessed online at http://www.nrri.org/pubs/telecommunications/NRRI_End_to_End_Smart_Grid_june11-12.pdf.
# Table of Contents

1. **Introduction** .................................................................................................................. 1

2. **Conceptual background** .................................................................................................. 3
   
   2.1 Bottleneck regulation ................................................................................................. 3
   
   2.2 Liberalization of electricity markets .......................................................................... 4
       
       2.2.1 Design and operation principle ......................................................................... 4
       
       2.2.2 Status of liberalization in Germany .................................................................. 6

3. **Critical bottleneck areas** ................................................................................................ 10
   
   3.1 Smart grid architecture ............................................................................................ 10
   
   3.2 Potential bottlenecks ............................................................................................... 12

4. **Threats of discrimination** ................................................................................................ 15
   
   4.1 The rationale of the “internalizing complementary efficiencies” theory ............... 15
   
   4.2 Exceptions to the internalizing complementary efficiencies theory ....................... 16

5. **Regulatory instruments** ................................................................................................ 21
   
   5.1 Interoperability and data .......................................................................................... 21
   
   5.2 Last mile .................................................................................................................. 22

6. **Discussion and conclusions** ............................................................................................ 25

**References** ......................................................................................................................... 29
1. Introduction

A national Smart Grid policy should encourage tens of thousands of entrepreneurs to innovate—using new technologies and business models—to create a wide variety of in-building energy management and information services.

Federal Communications Commission – National Broadband Plan

Alfred Kahn famously (1970, pp. xxxvii) said that the “central, continuing responsibility of commissions and legislatures” is to “find the best possible mix of inevitably imperfect competition and inevitably imperfect regulation.” Accordingly, regulation’s central goal is to establish a solid and appropriate framework for balancing public interest and entrepreneurial freedom (Picot 2009). This paradigm is the public-interest theory’s underlying principle. Utilities’ liberalization in Europe and elsewhere was based largely on this theory. The leading objective was to establish competitive markets that increase social welfare (Wernick 2007, pp. 23). In many economic sectors, the transition from monopoly to competition, based on public interest considerations, has been successful. In terms of deregulation, telecommunications is regarded as the leading sector concerning the “form, process and outcome of regulation” (Pollitt 2010).

The energy markets’ reform to a competitive market, which has proven difficult, has been the exception to the successful transition rule (Glachant and Finon 2003, Joskow 2003, Jamasb and Pollitt 2005). In particular, countries that deviated from liberalization’s “textbook model” (see Joskow and Schmalensee 1983), such as the U.S., Japan, and much of continental Europe, failed in developing efficient competition in the potentially competitive electricity value chain segments (Joskow 2006, Joskow 2008).

A major future challenge for electricity grids is the growing addition of intermittent—often distributed—renewable energy sources (RESs) and the low degree of automation, monitoring, and communication within distribution networks. Therefore, instantaneously balancing supply and demand puts a strain on all grids. Without fundamentally modernizing the grid’s infrastructure, RESs’ increasing penetration will result in an increased need for expensive (and higher polluting) balancing power. It will also result in a decline in the grid’s reliability, resilience, efficiency, and environmental sustainability.

Owing to recognizing the need for improved communication and coordination on a global scale, the “smart grid” concept emerged. A smart grid can best be understood as a communications layer’s virtual overlay on the existing power grid. This overlay allows all actors and components within the electricity value chain to exchange information, thereby facilitating supply and demand’s coordination (NIST 2009). This overlay closes the communication gap between consumers’ premises and the rest of the network, but requires the deployment of an AMI infrastructure. Therefore, legacy meters have to be replaced with smart meters. In analogy to the telecommunications industry, the AMI infrastructure and the smart meters can jointly be viewed as the “last mile” of smart grids (Leeds 2009).

In the telecommunications sector, the last mile is represented by the “local loop.” Physically, the local loop is a single, twisted, pair cable that connects consumers’ premises to the backhaul telecommunications network. International regulators treated the local loop as a monopolistic bottleneck, since no alternative infrastructure was available and potential replication was not viable. Telecommunication’s last mile was therefore an essential input that allowed competitors to offer downstream services, such as long-distance calls and internet services. Consequently, incumbents were mandated to grant unbundled access, which allowed a competitive downstream telecommunication and internet services market to be established (Cave 2010).
Similarly, smart meters and meter data’s non-discriminatory access and control rights allow competitors to offer downstream services in a smart grid. Most of these innovative downstream services, applications, and products, which will help improve energy efficiency, depend on seamless and reliable data exchange. Given the configuration of the metering and electricity market, incumbent distribution grid operators are likely to control the smart grid’s last mile. However, this last mile is an essential input for firms seeking entry to the complementary service and application market.

Smart grids will greatly affect traditional business models in the energy industry. Moreover, in many markets liberalization is still insufficient. Therefore, it is essential that regulators identify potential bottleneck facilities, as well as regulatory barriers to new smart grid concepts at an early stage, and find remedies to overcome them (Pérez-Arriaga 2009, ERGEG 2010, Hempling 2011). Thus, the aim of this paper is to identify these barriers and offer potential solutions.

The study therefore draws on the normative theory of regulation and applies insights from diverse literature streams. We investigate whether incumbent distribution network operators have opportunities and incentives to engage in exclusionary behavior and, if so, which regulatory instruments are adequate to relieve these bottlenecks. In sum, the following research questions guided this study:

RQ1: Are there bottlenecks within a smart grid’s communication layer?
RQ2: Do incumbent distribution system operators have incentives to use these bottlenecks to discriminate against independent third parties?
RQ3: If so, which regulatory instruments can remove these bottlenecks?

The remainder of the paper is divided into five sections. Section 2 provides the contextual background by describing the bottleneck regulation’s rationale and delineating the liberalized power market’s regulatory and economic framework, focusing on Germany. Section 3 describes smart grids’ architecture and identifies potential bottlenecks in these grids. In section 4, we examine the existence of incentives to discriminate by building upon the internalizing complementary efficiencies theory. In section 5, we discuss possible regulatory remedies to remove the bottlenecks. In the final section, we discuss the study’s findings and implications, and provide suggestions for future research.
2. Conceptual background

Starting with a monopolistic bottleneck’s definition, we present the prevailing legal and economic reasoning regarding bottleneck regulation, in section 2.1. In section 2.2, we lay the foundation for examining possible bottlenecks by outlining liberalized electricity market’s operating principle, power structures, and failures.

2.1 Bottleneck regulation

Business models in network economies substantially depend on particular networks’ availability and functioning. The irreversible costs and economies of bundling make duplicating such networks unfeasible (Joskow 2005, Viscusi et al. 2005, Picot 2009). Hence, a core element in the liberalization of any network industry is the network access’s regulation for independent market entrants (Schmidtchen and Bier 2005). Without access regulation, potential entrants to these markets would face substantial entry barriers, such as long-term cost asymmetries, that discriminate in favor of the incumbent (Stigler 1968, p. 67). An incumbent might own a facility that cannot realistically be economically and technically substituted. This facility might be essential for reaching customers, and/or for competition to emerge in downstream markets. If the facility has these characteristics, it is regarded as a “monopolistic bottleneck” or an “essential facility” (Knieps 1997, European Commission 1998, Blankart et al. 2007). A facility is always labeled as such whenever there is a natural monopoly. If this is the case, a firm can provide a facility more cost-effectively than several firms can (subadditivity), and the costs for the facility are irreversible (Lipsky and Sidak 1999). As competition in these markets is not feasible, they are regarded as incontestable (Baumol et al. 1982). Consequently, an essential facility’s owner has stable market power (Blankart et al. 2007).

Owing to an essential facility’s owner transferring the market power from the primary (upstream) market to a secondary (downstream) market in which the facility provides an essential input (Salinger 1989), the firm can take unfair advantage of its dominant position. The firm, for example, might refuse to deal with certain consumers or implement predatory pricing practices. The firm can also impede competitors’ access to large markets, and negatively affect the emergence of innovative services and products.

Thus, in order to avoid deadweight losses, to promote maximum efficiency, and allow active competition in complementary markets, non-discriminatory access to essential facilities is subject to ex ante regulation that should be in place before the market power can be abused (Lipsky and Sidak 1999, Blankart et al. 2007). The access problem is closely linked to the essential facilities doctrine (EFD), which was originally a U.S. antitrust law instrument (Renda 2010). Today, the EFD’s reasoning helps identify situations in which regulatory interventions are required (OECD 1996), since “any solution to the problems of economic inefficiency is inherently regulatory” (Lipsky and Sidak 1999). In this respect, competition law is insufficient to neutralize an owner’s network-specific market power. Furthermore, ex post interventions involve significant time lags (Gabelmann 2001).

Since its initial application in Europe (European Commission 1994), the EFD has been largely adopted in the European Commission’s Access Notice (1998, section 68). Based on the EFD in 2003, the European Commission proposed a three-criteria test for the electronic communications sector to define situations that require ex ante regulation (European Commission 2003, recital 9):
1) High and non-transitory entry barriers are present, whether of structural, legal or regulatory nature
2) The market structure does not tend towards effective competition within the relevant time horizon
3) Merely applying competition law will not address the market failure(s) adequately

Hence, if a monopolistic bottleneck in an upstream market threatens an efficient downstream market’s emergence, and the bottleneck owner has significant market power, the EFD’s reasoning is applied in order to substantiate regulatory intervention, which sets out to influence the secondary markets’ structure (Lipsky and Sidak 1999, Blankart et al. 2007, Renda 2010).

Most bottlenecks that were regarded as essential in the past (for examples see Lipsky and Sidak 1999) were “tangible” in nature, such as the local loop’s single twisted pair cable. However, there are also “intangible” bottlenecks based on intellectual property rights, such as proprietary standards, protocols, or interfaces. These could hinder competition in downstream markets, as argued by the U.S. Department of Justice (2002) and the European Commission (2004) in two antitrust cases against Microsoft. In these lawsuits, Microsoft was alleged to abuse the dominance of its Windows platform to discriminate against competitors in complementary markets by means of the non-disclosure of interoperability information (see Renda 2004). Intangible bottlenecks’ prevalence is likely to increase in ever more “digitally renewed” economies (Davis 2000).

In the following section, the EFD’s rationale and its application by the European Commission will serve as the guiding, underlying principle to analyze potential tangible and intangible bottlenecks.

2.2. Liberalization of electricity markets

Power systems’ structures have evolved over several decades. Hence, operation modes and design configurations differ across countries. However, in the course of liberalization, energy systems’ structures are converging. This section provides a brief overview of liberalized electricity markets’ regulatory, organizational, and technical structures with a focus on Germany. Although we focus on Germany, the functional pattern can be applied widely in most liberalized energy supply systems.

2.2.1 Design and operation principle

Within the electricity industry, three major areas of activity can be identified: generation, transport, and consumption. Since electricity markets’ deregulation, the generation and retail markets have been organized competitively. Conversely, the transport functions—transmission and distribution—continue to be treated as natural monopolies because of sunk costs, as well as economies of scale and scope in electricity delivery (Joskow and Schmalensee 1983, Monopolkommission 2002). In order to avoid monopolistic exploitation of these natural monopolies, third-party network access and revenues for network usage are regulated (Wilson 2002, Glachant and Finon 2003, Woo et al. 2003, Shioshani and Paffenberger 2006).

In Germany, four transmission system operators (TSOs) control the transmission network. Each TSO operates a control zone. Within a control zone, each TSO controls the voltages and stabilizes the frequency by contracting balancing energy via a separate market (Verhaegen et al. 2006) at usually high price levels (Rebours et al. 2007).

Dependent on the required response time, the market for balancing energy is divided into primary, secondary, and tertiary reserves, as depicted in Figure 1 (ENTSOE 2009). Owing to high technical requirements, only six providers can supply primary and secondary reserves in Germany (Monopolkommission 2009).
DSOs deliver power to end-consumers. In Germany, distribution grids are operated by 70 regional suppliers and about 870 municipal utilities, which are responsible for power quality and supply security in their areas. Besides planning, operating, and maintaining distribution grids, DSOs are legally obliged to procure the information required for electricity suppliers’ (ESs) energy accounting tasks.

Aside from large industrial consumers, ESs procure power for their consumers, either from the energy exchange or from wholesalers. ESs charge consumers for the electricity that they use as well as for the network usage costs, the costs of balancing power, and the costs for metering services. Traditionally, DSOs operated the metering service market as a regulated monopoly. In many electricity markets, however, the metering market has recently been liberalized to increase competition and to promote innovation. With the exception of the network functions, all the electricity markets’ segments have now been liberalized.

Figure 2 illustrates liberalized electricity markets’ complex structure by showing the relationships between selected market actors.
2.2 Status of liberalization in Germany

The previous section mentioned that electricity markets are—at least in principle—open to competition. However, liberalization does not necessarily imply effective competition. Furthermore, energy markets have structural characteristics that facilitate the exercise of market power (OECD 2004, Jamash and Pollitt 2005). Hence, a few companies, or groups of companies, still dominate electricity markets. They perform at least one function of power transmission and distribution, as well as at least one function of generation and supply. These companies, or groups of companies, are referred to as conglomerates, or vertically integrated utilities (VIUs).

In Germany, the electricity market is dominated by four large VIUs, namely E.ON, RWE, EnBW, and Vattenfall (Gleave 2010). Owing to their high degree of integration under company and obligation law (see Table 1) and their enormous market power, these companies can limit and restrict competition in the electricity industry value chain’s potentially competitive segments (Monopolkommission 2009).
These four conglomerates have a market share of about 90% in net electricity generation, which strikingly illustrates the high degree of concentration (Laird and Stefes 2009). Owing to this generation dominance, the VIUs can exploit their position (e.g., withholding generation capacity) to set higher prices than would be possible in a competitive generation market (Borenstein et al. 2000, Joskow and Tirole 2000, Hirschhausen et al. 2007). Studies analyzing electricity generation’s simulated marginal costs and wholesale prices at the European Energy Exchange (EEX) conclude that the observed price differences are not the result of exogenous factors, but rather that of market dominance and strategic behavior (Müsgens 2006, Hirschhausen et al. 2007).

Moreover, VIUs that simultaneously operate a control zone and own peak power plants have financial incentives to excessively use balancing power. Hence, the Monopoly Commission (2009) assumes that VIUs purposefully withhold generation capacities in order to sell more expensive balancing power, knowing that they will not be sanctioned by other market participants’ idle capacities. The German regulator is, however, increasingly aware of these inefficiencies, as well as of the potential exploitation of market power. Therefore, the regulator has taken the first effective steps (Rammerstorfer and Wagner 2009) towards a more efficient reserve market: Since the end of 2006, all German TSOs have been obliged to procure balancing power in a transparent web auction based on merit orders. In addition, they have been obliged to jointly coordinate balancing energy’s usage in all four control zones since 2010 (BMWI 2010). Recently, due to imminent antitrust suits, political pressure, and substantial future investment needs, two TSOs (E.ON and Vattenfall) sold off their transmission subsidiaries (Bundesnetzagentur 2009). However, the four large VIUs still own several distribution networks or hold shares of DSOs.

At first glance it might seem that the VIUs’ dominance in the retail market is less pronounced as they only have a market share of around 50%. However, only about 5% of the consumers obtain power from truly (ownership unbundled) independent retailers (A.T. Kearney 2009). The remaining market is shared between municipal utilities and regional energy suppliers, many of which are integrated with VIUs in one form or another, as already shown in Table 1. Given the outlined lack of competition and the high degree of concentration in the German energy industry, it is not surprising that electricity prices in Germany are among the highest in Europe (Eurostat 2010). Although country-specific conditions have to be taken into account, the price level is probably the single most important performance assessment indicator of liberalization (Jamasb and Pollitt 2005). Given Germany’s price level, this indicates an insufficient liberalization of the electricity industry. Further, the price level is aligned with the extent to which countries have separated the network functions from the competitive segments, as shown in Table 2. Based upon five indicative criteria—ownership, accounting, regulatory aspects, legal aspects, and physical aspects—the analysis indicates that the extent of German DSOs’ unbundling is very low.

<table>
<thead>
<tr>
<th>Minority holding (&lt;25%)</th>
<th>E.ON</th>
<th>RWE</th>
<th>EnBW</th>
<th>Vattenfall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>22</td>
<td>22</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Qualified minority holding (&gt;25%—&lt;50%)</td>
<td>96</td>
<td>32</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>Majority holding (&gt;50%)</td>
<td>35</td>
<td>29</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td><strong>153</strong></td>
<td><strong>83</strong></td>
<td><strong>41</strong></td>
<td><strong>9</strong></td>
</tr>
</tbody>
</table>

Table 1. Number of holdings in regional or local power utilities (based on data of Monopolkommission 2007)

Although country-specific conditions have to be taken into account, the price level is probably the single most important performance assessment indicator of liberalization (Jamasb and Pollitt 2005). Given Germany’s price level, this indicates an insufficient liberalization of the electricity industry. Further, the price level is aligned with the extent to which countries have separated the network functions from the competitive segments, as shown in Table 2. Based upon five indicative criteria—ownership, accounting, regulatory aspects, legal aspects, and physical aspects—the analysis indicates that the extent of German DSOs’ unbundling is very low.
<table>
<thead>
<tr>
<th>Country</th>
<th>Transmission System Operator (Max. Score=5)</th>
<th>Distribution System Operator (Max. Score=5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Belgium</td>
<td>4</td>
<td>3.5</td>
</tr>
<tr>
<td>Denmark</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Finland</td>
<td>5</td>
<td>1.5</td>
</tr>
<tr>
<td>France</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Germany</td>
<td><strong>4</strong></td>
<td><strong>1.5</strong></td>
</tr>
<tr>
<td>Greece</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Ireland</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Italy</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Netherlands</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Portugal</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Spain</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Sweden</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>UK</td>
<td>5</td>
<td>4.5</td>
</tr>
<tr>
<td>Norway</td>
<td>5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

• TSO: Ownership unbundling, Yes=1, No=0; DSO: Legal unbundling, Yes=1, No=0
• Published accounts, Yes=1, No=0
• Compliance officer, Yes=1, No=0
• Separate corporate identity, Yes=1, No=0, Often=0.5
• Separate locations, Yes=1, No=0, Partly=0.5

Table 2. Extent of network unbundling (based on European Commission 2005)

An equally complex and difficult liberalization process looms in the market regarding metering services’ future. The market was liberalized in 2005 and opened to third parties in 2008. The aim was to establish a competitive market for metering services and, additionally, to rapidly deploy electronic meters that could measure time-differentiated energy usage (Bundesregierung 2007). In liberalized metering markets, consumers can freely choose whom they authorize to fulfill the metering-related functions of measuring point operation and measurement service provision. They can even choose to authorize a single economic actor, or several actors, to fulfill these functions. We simplify this issue for our analysis, however, by assuming that these functions are provided by a single economic entity, to which we refer as the “metering provider” (MP). The MP’s services involve various tasks. The most essential tasks are “purchase, installment and maintenance of the meter, meter data collection, management and provision of meter data to other market players” (ERGEG 2007).

For a variety of reasons, virtually no competition has emerged since the opening of the German metering market. The roll-out of the new metering infrastructure is therefore
proceeding at a snail's pace. Potential entrants and network operators criticize the lack of investment security. The latter is due to the absence of standard business processes and minimum technical requirements regarding the new metering devices, as well as a lack of clarity with regard to financing. Furthermore, non-integrated entrants to the MP market are obliged to install meters that comply with technical and data provision requirements that DSOs partly specify for each of their distribution networks. Consequently, very few market actors offer consumers smart meters who in turn also cannot be considered as a dynamic competitive element (Trend Research 2010, pp. 1070).

The Netherlands and the UK employed a similar roll-out approach driven by competition policy (Wissner 2009). In a response to the slow diffusion, the Netherlands, however, recently switched from a competitive market design to a regulated model (Wissner and Growitsch 2010). As in the UK, the German regulator—at least for the present—seems to favor a competitively organized metering market. However, in the federal government’s new Energy Concept, it clarified that it anticipates a nationwide roll-out of smart meters. This intention is a breeding ground for a potential changeover to a regulated market design like in the Netherlands if the deployment continues to be slow (Bundesregierung 2010). However, regardless of the actual market design, DSOs are likely to dominate the metering service market for several reasons that we outline in the following.

In regulated markets, DSOs will probably be in charge of the smart meter roll-out. For example, this is what is happening in the Netherlands and Sweden. In these countries DSOs continue to earn regulated returns. In competitive markets, however, DSOs will almost certainly play a dominant role in the metering market because otherwise they would face various disadvantages: First, DSOs are legally obliged to take over metering services immediately in situations in which a competing MP fails to carry out its responsibilities. This entails DSOs having to have knowledge of the new metering technology although they do not operate in the metering market themselves. Second, losing customers to a competing MP means that DSOs would not only lose dependable revenue sources but also long-established customer relationships, which are valuable business assets. Third, as long as consumers do not proactively choose another MP, DSOs remain responsible for providing metering services. However, there is very little consumer demand for metering services. Therefore, new competitors from outside the industry are reluctant to enter the market, which results in DSOs continuing to act as MPs under a revenue cap regulatory regime. In addition, new competitors’ market entry is complicated by DSOs’ right to partly specify the technical requirements for their distribution areas. This increases competitors’ transaction costs and limits economies of scale.

Thus, given the current regulatory and legal provisions, as well as the electricity and metering markets’ characteristics, DSOs will continue to act as regulated monopolists in the metering market by carrying out metering services themselves. The following analysis is based on this hypothesis.
### 3. Critical bottleneck areas

Before identifying potential monopolistic bottlenecks, it is important to understand the structure of future grids from a technological point of view. In the following section, we therefore outline smart grids’ architecture.

#### 3.1 Smart grid architecture

There is, as yet, no precise definition of the smart grid concept (Pérez-Arriaga 2009). The concept’s novelty and complexity, as well as the extensive variety in associated technologies make it hard to define it concisely (OECD 2009, Orlamünder 2009). Consequently, we find that existing studies and literature define the smart grid in line with each of their respective focal areas (see Table 3).

<table>
<thead>
<tr>
<th>Source</th>
<th>Smart Grid Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERGEG (2010)</td>
<td>An electricity network that, cost efficiently, can integrate the behaviour and actions of all users connected to it—generators, consumers and those that do both—in order to ensure a sustainable power system with low losses and high levels of quality, security of supply and safety.</td>
</tr>
<tr>
<td>EPRI (2005)</td>
<td>A smart grid uses digital technology to improve reliability, security, and efficiency (both economic and energy) of the electric system from large generation, through the delivery systems to electricity consumers and a growing number of distributed-generation and storage resources.</td>
</tr>
<tr>
<td>DOE (2008)</td>
<td>A smart grid as part of an electricity power system can intelligently integrate the actions of all users connected to it—generators, consumers and those that do both—in order to efficiently deliver sustainable, economic and secure electricity supplies.</td>
</tr>
<tr>
<td>ENSG (2009)</td>
<td>A Smart Grid as part of an electricity power system can intelligently integrate the actions of all users connected to it—generators, consumers and those that do both—in order to efficiently deliver sustainable, economic and secure electricity supplies.</td>
</tr>
<tr>
<td>OECD (2009)</td>
<td>The smart grid is an innovation that has the potential to revolutionise the transmission, distribution and conservation of energy. It employs digital technology to improve transparency and to increase reliability as well as efficiency. ICTs and especially sensors and sensor networks play a major role in turning traditional grids into smart grids.</td>
</tr>
<tr>
<td>DECC (2009)</td>
<td>Building a ‘smarter grid’ is an incremental process of applying information and communication technologies (ICTs) to the electricity system, enabling more dynamic real-time flows of information on the network and greater interactivity between suppliers and consumers. These technologies can help deliver electricity more efficiently and reliably, from a more complex range of generation sources, than the system does today.</td>
</tr>
<tr>
<td>FERC (2009)</td>
<td>Smart Grid advancements will apply digital technologies to the grid, and enable realtime coordination of information from generation supply resources, demand resources, and distributed energy resources (DER). This will bring new efficiencies to the electric system through improved communication and coordination between utilities and with the grid, which will translate into savings in the provision of electric service. Ultimately, the smart grid will facilitate consumer transactions and allow consumers to better manage their electric energy costs.</td>
</tr>
<tr>
<td>European Technology Platform Smart Grids (2006)</td>
<td>Electricity networks that can intelligently integrate the behaviour and actions of all users connected to it—generators, consumers and those that do both—in order to efficiently deliver sustainable, economic and secure electricity supplies.</td>
</tr>
<tr>
<td>Adam and Wintersteller (2008)</td>
<td>A smart grid would employ digital technology to optimise energy usage, better incorporate intermittent “green” sources of energy, and involve customers through smart metering.</td>
</tr>
<tr>
<td>Climate Group (2008)</td>
<td>A “smart grid” is a set of software and hardware tools that enable generators to route power more efficiently, reducing the need for excess capacity and allowing two-way, real time information exchange with their customers for real time demand side management (DSM). It improves efficiency, energy monitoring and data capture across the power generation and T&amp;D network.</td>
</tr>
<tr>
<td>CISCO Systems (2009)</td>
<td>The smart grid is a data communications network integrated with the electrical grid that collects and analyzes data captured in near real-time about power transmission, distribution and consumption. Based on these data, smart grid technology then provides predictive information and recommendations to utilities, their suppliers, and their customers on how best to manage power.</td>
</tr>
</tbody>
</table>

**Table 3. Smart grid definitions**
The selected definitions suggest that there are two different approaches to defining a smart grid (OECD 2009): One perspective highlights the technical components, the other focuses on the smart grid’s capabilities.

From a capabilities perspective, a smart grid is characterized by enabling a two-way flow of electricity and information, thereby ensuring a high degree of interconnectivity between all actors and components of the electricity system. By collecting, processing, and analyzing data on power generation, transmission, distribution, and consumption in (near) real-time, a smart grid is expected to provide a wide range of benefits across the entire electricity value chain. These anticipated benefits are summarized in Figure 3.

From a technical perspective, a smart grid is comprised of three layers. Each of these layers integrates a multitude of digital and non-digital technologies and systems from the realms of telecommunication, information, and energy technology (see Figure 4). From an architectural point of view, a smart grid can be best understood as an additional communication layer that is virtually overlaid on to the existing power grid and on which an application layer is built.

<table>
<thead>
<tr>
<th>Supply Side</th>
<th>Electricity Network</th>
<th>Demand Side</th>
</tr>
</thead>
</table>
| - Optimizing the facility’s utilization and reducing the need for exceed capacity that peak load power plants provide  
- Improving the connection between and operations of generators of all sizes and technologies  
- Reducing the entire electricity supply system’s environmental impact | - Preventive maintenance and remote grid management through better monitoring and control features  
- Minimizing energy losses through efficient energy routing  
- Increasing the degree of automation and “self-healing” responses to system disturbances  
- Incorporating DERs and PHEVs effectively | - Providing consumers with better information  
- Increased responsiveness and demand flexibility  
- Enhanced efficiency through better management options and greater awareness of energy consumption  
- Giving power consumers a more participative and active role  
- Enabling innovative services and applications |

**Figure 3. Anticipated smart grid benefits.**

**Figure 4. Smart grid architecture.**
By employing a layered approach of this kind, the design problem’s complexity is reduced, because the functionality is modularized in components and subcomponents (van Schewick 2010, pp. 50). By interconnecting formerly isolated components, actors, networks, and technologies, a smart grid facilitates the creation of a system of systems (NIST 2009). Hence, a smart grid can be conceived as a system product. By definition, it requires the components to be compatible. The different systems must function seamlessly with each other to produce the desired outputs (Langlois 1999). Each layer’s components perform specific functions, and have well-defined interfaces with the upper layer in order to make their services available. Simultaneously, they make use of the layer beneath’s services. A smart grid therefore emulates the internet’s original design principle by employing an “end-to-end” architectural approach. Within this architecture application specific functionalities are implemented at higher layers at the network’s end hosts or end points, while lower layers are kept as general and application-independent as possible (Saltzer et al. 1981). In an end-to-end network, components and actors can send and receive data without knowing the network’s structure (Economides and Tåg 2009). The network itself therefore remains neutral. This encourages innovations at the network’s end (Cerf 2006a, Cerf 2006b), and is widely regarded as the key driver for the internet’s rapid development. This development is also characterized by low entry barriers and non-discriminatory access for innovators (van Schewick 2007). Similarly, in a smart grid, the innovation is also expected to come from the network’s end (FCC 2010). While there might be some innovation at the network’s core, the innovative applications and services at higher layers will provide the literal “smartness”.

Hence, our work focuses on identifying bottlenecks that require regulatory interventions within the communications layer. The interventions will ensure a “neutral” smart grid that promotes entrepreneurship and grants non-discriminatory access and low entry barriers for new market entrants. In the next section, we will analyze which facilities can act as bottlenecks.

3.2 Potential bottlenecks

Since a communication layer is overlaid on top of the power layer, the communications gap between customers’ premises and the remaining actors and components in the energy value chain will be bridged for the first time in power systems’ existence. Utilities have already deployed a communications network (LAN, MAN, or WAN) that connect parts of their infrastructure (especially transmission systems) with supervisory control and data acquisition systems (SCADA) to manage grid operations. However, the missing link to consumers has to be built from scratch for which various narrow- (e.g., GPRS and GSM) and broadband (e.g., fiber-optics, BPL, LTE, and satellite) technologies can be used. Depending on regional geographical conditions, and wired or wireless access platforms’ penetration, utilities can build upon the commercially available communication infrastructure. This will allow economically inefficient investments to be avoided if the existing infrastructure is able to cope with the smart grid’s quality-of-service, security, reliability, and resiliency requirements (FCC 2010). If there is no other option, utilities may deploy the necessary infrastructure themselves, or they can do so in joint ventures. In the U.S., for example, both alternatives are in use: Utilities piggy-back on commercial network infrastructure, or they build their own infrastructure, using wireless mesh networks or power-lines, to connect smart meters (Heidell and Ware 2010).

By linking the existing utilities’ communication networks with smart meters, the AMI (also referred to as Field Area Network (FAN)) facilitates an end-to-end network. The AMI allows data to be transported back and forth between consumers and other market actors (see Figure 5). In buildings, the smart meter serves as the central gateway to in-house devices such as home appliances, consumer electronics, water heaters, lighting systems, and programmable thermostats connected via Home Area Networks (HAN). Thus, to enable innovative
applications, such as demand response and virtual power plants, authorized market actors like independent energy service providers (ESP), need access to the smart meter. This access would be via the AMI, which will allow direct communication with the smart meter, and enable authorized market actors to send price signals, control appliances, and change tariffs (ERGEG 2007). Thus, smart meters, together with AMI, serve as an essential gateway. This gateway can be deemed synonymous with the last mile in telecommunications, as it acts “as the final leg delivering connectivity from a utility to a consumer” (Leeds 2009, pp. 11). On the communication layer’s one end, the AMI connects smart meters, while on the other end, it interfaces with the backhaul network that aggregates and transports the data to the WAN, as illustrated in Figure 5 (NIST 2009).

Similarly to telecommunications, the last-mile infrastructure in a smart grid is an essential input. The last-mile infrastructure provides access, which is necessary to enter the downstream market. The infrastructure cannot be substituted or replicated within a reasonable time and/or cost frame, due to substantial sunk costs and economies of bundling.

The data retrieved from smart meters can also be regarded as essential inputs for authorized actors. The data aids them in improving grid management and monitoring, streamlining business processes, and enabling innovative energy efficiency measures and value-added services (ERGEG 2007, FCC 2010, OFGEM 2010). Hence, it is crucial that MPs who are in charge of collecting and administrating the meter data, provide authorized parties with non-discriminatory and efficient access to the meter data, in compliance with national security and privacy requirements. In order to ensure an efficient data provision also standardized data formats are necessary. Table 4 provides an overview of the market actors and their respective data needs.

<table>
<thead>
<tr>
<th>Actors</th>
<th>Use of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSO</td>
<td>Grid operation, ES billing, forecasting, loss detection, and customer service process automation, customer switching, power quality monitoring</td>
</tr>
<tr>
<td>Supplier</td>
<td>Billing, tendering, forecasting, and trading</td>
</tr>
<tr>
<td>Generation (distributed)</td>
<td>Plant operation, fulfillment of supply contracts</td>
</tr>
<tr>
<td>Customer</td>
<td>Information, usage control, decision making</td>
</tr>
<tr>
<td>ESPs and other third parties</td>
<td>Using energy efficiency measures, input to home and building automation, aggregation of supply and demand data for electronic electricity markets</td>
</tr>
<tr>
<td>Government Body or Regulators</td>
<td>Monitoring power quality, statistics, and disaster management</td>
</tr>
</tbody>
</table>

Table 4. Actors and their data needs (based on ERGEG 2007)
Ultimately, smart grids’ goal is to enable actors and components’ end-to-end communication. Currently, only limited information exchange is possible in power systems, due to specialized rules for data exchange. For example, the core utilities’ information systems (SCADA) typically use their own specialized communications protocol. These protocols only enable communication within a subsystem, and impede communication between subsystems (CISCO Systems 2009). Therefore, in order to achieve end-to-end interoperability, it is crucial to build a smart grid’s communication network on a basic set of open and non-proprietary communication protocols and standards (DKE 2010, ERGEG 2010, NIST 2010).

In sum, an interoperable end-to-end smart grid communication layer’s development is essential for competitive downstream markets’ emergence. With regard to the communication layer, we identified three critical bottlenecks areas: the last mile, meter data, and interoperability. DSOs control access to the last mile and the meter data, and influence interoperability considerably. They would therefore have manifold opportunities to discriminate against independent third parties in the complementary market.

With regard to the European Commission’s three-criteria test for electronic communication markets, recital 11 of the 2003 Recommendation states that a “structural barrier can also exist where the provision of service requires a network component that cannot be technically duplicated or only duplicated at a cost that makes it uneconomic for competitors”. Hence, the three-criteria test’s first condition applies to the smart grid’s last mile. On the one hand, once DSOs deployed the infrastructure, sunk costs create a long-term cost asymmetry between DSOs “inside” the market and potential entrants “outside” the market. On the other hand, the replication is practically and economically “not easy” for competitors (European Commission 2002). Thus, in the next section, we discuss whether DSOs have reasons to engage in discriminatory conduct.
4. Threats of discrimination

Owing to the current regulatory and legislative provisions, as well as the metering and electricity market’s characteristics, there are potential bottlenecks in a smart grid’s communications layer. DSOs can discriminate against independent producers of complementary applications, services, and products (summarized as applications in the following) by impeding access to the data and/or the last mile. We refer to the data and the last mile as the “product” or “platform” in this section. Consequently, our further analysis is based on the hypothesis that DSOs will roll out the new metering infrastructure and keep on acting as metering providers. Hence, they will continue to act as regulated monopolists in their distribution areas. We refer to a “monopolist” as a company that has substantial control over prices and outputs, as common in antitrust law (Posner 2001, pp. 195).

4.1 The rationale of the “internalizing complementary efficiencies” theory

The Chicago School’s neoclassical economic theory argues that a monopolist does not have reasons to monopolize a complementary market, given that the applications (e.g., energy management services) that are complementary to its product (e.g., AMI and smart meter infrastructure) are competitively supplied and used in fixed proportions with the monopoly product (Bork 1993, pp. 372, Posner 2001, pp. 198). This reasoning is known as the “one rent monopoly theorem” (Bowman 1957, Posner 1976, pp. 200, Bork 1978, pp. 372). The theorem suggests that the monopolist can extract the complete monopoly rent by the pricing of its primary good and cannot, therefore, gain any additional profit by capturing the secondary market (Whinston 1990, Farrell and Katz 2000). When a monopolist owns the product and only one monopoly rent is available in the final product’s market, the monopolist has no incentive to engage in exclusionary behavior, as it can capture the complete monopoly rent in the primary market.

If the value of a monopolist’s product increases because of unaffiliated third parties’ complementary applications, the monopolist in principal has an incentive to spur independent producers’ entry, which will allow it to capture an additional surplus. In other words, the monopolist can internalize complementary efficiencies (ICE) (Farrell and Weiser 2003). This logic is illustrated by the following numerical example.

Consider, for instance, a mobile phone platform, owned by a company with substantial control over prices and outputs in its market. Imagine that the phone’s stand-alone value is $80 and that, because of available third-party applications, the original consumer value of $80 increases by $20. In total, the product’s new value is now $100. If, in this situation, the monopolist seeks to monopolize the secondary applications market, this would result in the phone’s value decreasing by $20 (this additional value includes the available applications’ variety, quality, price, and usefulness, measured when the product has already been purchased). Consequently, by monopolizing the secondary market, the monopolist would either lose $20 because the phone’s price has been reduced to its original value of $80 due to the lack of independent applications, or sell fewer phones at a maintained price of $100. Therefore, a rational monopolist in this situation would not seek to monopolize the secondary applications market. Instead, the monopolist will try to capture the additional $20 value with the phone’s pricing (Farrell and Weiser 2003). In this situation, the monopolist’s gains are higher if independent producers provide complementary applications.

Thus, the monopolist has incentives to facilitate easy access to the product by providing, for example, third parties with its product’s interface information. ICE thus argues that the monopolist will choose a pattern that ensures that its product and customers are
provided with independent applications. Consequently, no anticompetitive problems will arise (Farrell and Weiser 2003).

The one rent monopoly theorem therefore suggests that a monopolist’s exclusionary conduct will not result in a monopoly rent’s increase. The ICE theory goes even further by suggesting that such behavior would decrease a monopolist’s profits. In practice, however, a monopolistic product’s owner often has an financial interest in integrating into the market offering complementary applications, and is therefore likely to occupy a dominant position in this market (Farrell and Weiser 2003).

A monopolist will hence often choose to abandon an arm’s-length relationship and integrate vertically. Regardless of the particular intentions associated with vertical integration, such as decreasing the coordination costs or strengthening incentives for product deployment (Coase 1937, Coase 1960, Williamson 1979), this will lead to competitive disadvantages for independent providers. It will also increase policy concerns (Farrell and Katz 2000, Farrell 2003). In respect of this situation, the ICE theory also argues that a monopolist’s decision to engage in the secondary market will be efficient. Even when integrating vertically, a monopolist has reasons to continue promoting independent applications (Farrell and Weiser 2003).

Hence, in summary, the Chicago School claims that no anticompetitive concerns are associated with vertical integration, as a monopolist would not increase the overall monopoly profits through discrimination, because it can always charge customers a higher price for its product (Bowman 1957, Bork 1978, pp. 288). Consequently, a monopolist will welcome third-party applications. It has no reason to extract profits from unaffiliated producers or even to impede or exclude rivals from access to the complementary market. It might even prefer not to enter or withdraw from the applications market, as its presence might increase independent producers’ concerns.

However, contrary to the ICE, recent research shows that there are situations in which a firm might monopolize the downstream market (Whinston 1990, Whinston 2001, Farrell and Weiser 2003, van Schewick 2007, van Schewick 2010). In these cases, the monopolist might very well benefit from the presence of an independent applications market. However, the profits associated with monopolizing the downstream market are greater than the losses associated with the lack of independent applications. In the next section, we elaborate on these exceptions.

4.2 Exceptions to the internalizing complementary efficiencies theory

The ICE theory provided the basic guideline for analyzing vertical integration’s effects for a long time (Farrell and Weiser 2003). However, ICE’s claims do not always hold, as we will illustrate in the next section. With regard to the research context, we identify four elementary exceptions (Baxter’s Law, complementary products reduce outside revenue, potential competition in the primary market, and regulatory strategy) to the theory’s reasoning without claiming to be inclusive.

**Baxter’s Law**

The one monopoly rent and ICE theory’s fundamental premise collapses when a monopolist is unable to extract additional consumer value from independently provided applications via the monopoly product’s pricing. This occurs if the monopolist’s product is more regulated (revenue cap or rate of return) than the complementary market’s products (Laffont and Tirole 2000). While metering providers’ revenues in the primary market will be subject to regulation, the secondary market for applications will not be regulated. In this constellation, a monopolist is unable to extract the full monopoly rent in the primary market. As a result, the monopolist’s conduct will change from “application-promoting” to “application-impeding” (Farrell and Katz 2000). This differs radically from ICE’s claim.
Given that the price of the monopolist’s product is set under a price cap regulatory regime, which does not respond to changes in the consumer value over time, and the proportions between the primary and secondary market are fixed, a monopolist might want to increase the profits that originate from the applications market. Therefore, it might be tempted to abuse its dominant position in the primary market. According to ICE’s logic, such behavior would result in the primary product’s price decreasing due to the monopolist’s surcharge in the applications market. However, since the platform price is regulated below the profit-maximizing level, the impact on the platform profits will be less significant. Hence, a monopolist can compensate for some of the “lost” monopoly profits in the (regulated) primary market by generating additional profits in the secondary market. This would, however, be inefficient if the primary market is unregulated (Farrell and Weiser 2003).

Presume regulated prices were to change over time, mimicking a price cap or a rate of return. Again, the monopolist will increase the prices in the monopolized complementary market. This will result in a corresponding revenue decline in the primary market. However, even in the short term, the monopolist will benefit from this behavior, as the regulatory process will eventually restore its losses in the primary market (Farrell and Weiser 2003).

The antitrust suit, United States v. AT&T, which was settled by William Baxter, is a prominent example of Baxter’s Law (see Joskow and Noll 1999). In this lawsuit, the vertically integrated telecommunications company AT&T was accused of abusing its market power via the integrated Bell Operating Companies, which controlled local access to the telephone network. AT&T leveraged this monopoly to the (potentially competitive) markets for long distance services and telecommunications equipment by refusing equal access to the essential input (Joskow and Noll 1999). Consequently, AT&T’s affiliate, Western Electric, could rent and sell consumers telephone equipment at inflated prices without being penalized—as ICE would suggest—since the prices for the local exchange telephone services were regulated.

In general, Baxter’s Law can be applied in any industry in which a vertically integrated incumbent is active in both the primary (regulated) and secondary (competitive) markets (Joskow and Noll 1999). With regard to the research context, regulated DSOs can deter entry by raising rivals’ costs through practices such as exclusive dealing, refusals to deal, tying, or defining of proprietary protocols and standards to artificially increase rivals’ transactions and consumers’ switching costs (see Salop and Scheffman 1983, Krattenmaker and Salop 1986, Salop and Scheffman 1987). These practices do not only cause customer lock-ins, but also reduce consumer and total welfare (Choi and Christodoulou 2001, Carlton and Waldman 2002).

Complementary products reduce outside revenue

In some cases, complementary applications are a source of outside revenues for unaffiliated providers, but they reduce monopolists’ outside revenues. Usually, revenues come from products’ direct sales or from fees for a good’s provision. If, however, complementary products are a source of outside revenues (e.g., advertising revenues), these products are often offered for free or below marginal costs since profits are derived from outside sources. This is a common principle in two-sided markets (e.g., print media, internet search engines, or credit cards), in which one side generates the profits while the other side is subsidized (Rochet and Tirole 2001).

This principle applies for applications (of which demand response is the most prominent) that threaten outside revenues of a DSO’s parent company. These outside revenues originate from other segments, such as electricity sales or generation. Thus, complementary applications can be a source of outside revenues for rivals in the downstream market, while negatively impacting a monopolist’s (outside) revenues (van Schewick 2007).
Under these conditions, a monopolist is unable to extract all potential revenues from outside sources unless it monopolizes the complementary market.

According to ICE’s reasoning, a monopolist would usually try to force rivals to lower the complementary good’s prices in order to extract the consumer surplus that would result in the secondary market. However, this is not feasible, since the price for this class of complementary applications is already zero. Furthermore, the primary market is regulated. A monopolist might, however, threaten independent producers with exclusion or discrimination if they do not pay inflated access charges (Farrell and Katz 2000).

Completely excluding rivals might still be more profitable than extracting (some) of the outside revenues. The monopolist might therefore choose to exclude its rivals entirely from the complementary market. These kinds of applications might reduce revenues in other business segments in which the monopolist has a financial interest to make as many sales as possible. Alone through the SMT’s mere presence in households electricity consumption is expected to drop considerably (BMWI 2006). Moreover applications like demand response, virtual power plants, or e-marketplaces seek to provide customers with incentives to adjust demand to electricity’s current availability.

Consequently, consumption is supposed to increase during the periods of the day in which electricity is cheaper and sales will decrease at “expensive” times of the day. The sales subsidiary of a DSO’s parent company might try to compensate these losses by increasing electricity prices. However, owing to retail competition, it is only partially possible to increase prices. Furthermore, if the DSO’s parent company has a stake in generation facilities, it has an additional interest in engaging in exclusionary behavior. Applications such as demand response are likely to lower the electricity price’s level, as well as the energy exchange’s volatility. Moreover, if a DSO is affiliated with a VIU that offers balancing power, complementary applications may negatively affect this VIU’s revenues by increasing competition in the market for balancing power. Additionally, with regard to revenues from distribution grids’ operation—even though they are RPI-X-regulated (retail price index minus expected efficiency savings)—DSOs still have a (small) incentive to increase, or at least stabilize, the amount of electricity delivered to consumers (Diekmann et al. 2006).

Under particular conditions, however, DSOs may also welcome some of the complementary application effects, such as demand response, which help increase the reliability of grids, prevent investments in grid expansion, intensify consumer relations, and improve power quality. Especially power quality will become more important in the near future, as it will be included in the network charges’ calculation. In sum, however, DSOs’ profits—earned by making as many sales as possible, either through their retail, generation, or distribution activities—will in most cases outweigh the benefits of complementary applications that can be also offered by DSOs themselves, if profitable and beneficial.

Potential competition in the primary market

Another situation in which the monopolist’s actions may not conform to the ICE theory emerges when it fears that a rival might attack its monopoly in the primary market. Thus, even though a monopolist may profit from complementary applications, it will opt for a lower profit that accrues from a two-level monopoly, namely a monopoly in the primary and secondary market, rather than to risk losing its dominant position in the primary market (van Schewick 2007).

To keep potential competitors from entering the primary market, a monopolist has to capitalize on entry barriers. A monopolist might therefore prevent potential rivals from having sufficient supply of complementary applications. This requires a monopolist to offer its applications only to its customers. Owing to its exclusionary conduct, the number of independent application providers is limited. In this case, a competitor would have to enter the primary and secondary market at the same time. For this exclusionary strategy to succeed, a
two-level entry must be implemented, which is more expensive, riskier, and more difficult than entering one market alone (van Schewick 2007), which applies to the research context: First, given that a competitor only has experience and competencies in one of the two markets it intends entering, it would face increased capital costs. Other investors will charge higher risk premiums, because deploying the necessary infrastructure causes high sunk costs, which additionally increase capital expenditures (CAPEX). Second, economies of scale differ considerably between provisioning of the last mile in the primary market and offering applications in the complementary market. While only a small number of customers may be needed in the applications market in order to break even, the required number of customers in the primary market will be significantly higher. Hence, a rival has to decide whether to operate at an unnecessarily small size in the applications market, or at an inflated size in the metering services market. Both cases would increase operational expenditures (OPEX). Third, two-level entry requires a sufficient supply of complementary applications for the competitive platform. If supply is lacking, a rival cannot compete with a monopolist’s primary product, which already offers complementary goods and services. However, as the metering applications market is subject to indirect network effects, a rival’s entry will be additionally complicated, because independent producers—if present at all—will prefer to provide applications for the larger network (Katz and Shapiro 1994). A rival would therefore have to “invest” in convincing independent third parties to offer applications for its platform.

One may argue that a single DSO does not carry enough weight to deter rivals from entering the secondary market. However, given the current regulatory provisions and the market characteristics, most DSOs will have a financial interest in discriminating against unaffiliated parties. Consequently, they might succeed in forcing independent firms to operate at a less efficient scale, or with a smaller customer base.

Another possible threat to DSOs is that unaffiliated firms might get too powerful in the complementary market. It might therefore become feasible for rivals to enter the primary market (Shapiro 2000). By entering the primary market, a rival in the complementary market would not only benefit from safeguarding its access to consumers, but also from lowering the platform’s price. Therefore, it can increase its complementary product’s sales. Consider, for instance, a demand response service provider that generates profits by selling balancing power that is subject to increasing returns to scale. The provider has a financial interest in contracting with as many consumers as possible and might therefore choose to enter the primary market and offer the monopolist’s product at a lower price. Consequently, the demand response service’s outside revenues will increase.

In sum, a DSO can safeguard its dominant position in the primary market by establishing a two-level monopoly.

Regulatory strategy

Owing to regulatory considerations, a monopolist might decide not to provide access to its product in a particular context, because this could result in the regulator imposing additional obligations in other contexts (Farrell and Weiser 2003). Some DSOs might welcome third-party applications, for example, those that provide security or assistance services. These would add value to the platform and increase DSOs’ profits (e.g., through certifying or licensing) to some extent. However, some applications decrease DSOs’ revenues. Therefore, if a DSO assumes that opening its platform to one class of applications is likely to involve granting access to other (profit-decreasing) application types, it might rather choose not to open its platform at all. A DSO might come to the conclusion that if it provides access to its platform, its future strategic scope will be limited, since returning to a closed, or fully integrated, platform would give rise to anticompetitive concerns. Hence, a monopolist’s regulatory strategy might result in the paradoxical situation of keeping its platform closed, although some independently provided applications may increase its profits.
Profitability of discrimination without monopolization

The exceptions that we outlined are based on the implicit assumption that a DSO has to monopolize the complementary market to make exclusion or discrimination a profitable strategy. However, given the electricity and metering market’s special characteristics, a DSO might still be motivated to abuse its market power in the primary market, even though it might not be able to entirely exclude rivals from the secondary market.

First, given that an application’s market price is considerably above marginal costs, any sale of additional applications generates an increase in profit (Shapiro and Varian 1999). This means that discriminating against independent providers is a profitable strategy, although rivals cannot be completely excluded from the secondary market. Second, suppose complementary applications were to threaten a DSO’s outside revenues. In this case, discriminating against these application providers might still be the best available strategy, because at least some consumers, who would have used a competitor’s application, will choose a DSO’s rivaling application instead. Consequently, a DSO can compensate at least some of its lost outside revenues. Furthermore, the DSO can “customize” the applications according to its economic interests. Third, DSOs do not have to monopolize the market for complementary goods to prevent a firm with considerable market power and an existing consumer base emerging. Discriminating against unaffiliated firms in the secondary market may suffice to prevent threats to DSOs’ monopoly in the metering services’ primary market. Given how effective discrimination without monopolization is, the likelihood that DSOs will engage in discriminatory conduct will increase, because they do not have to monopolize the entire complementary market to gain additional profits.

Our analysis was based on the hypothesis that a monopolist does not face competition in the primary market. Even if this assumption proves to be invalid, van Schewick (2007) shows that ICE’s rationale does not apply automatically. Whether all four presented exceptions will occur in each distribution area is an empirical question. Moreover, the analysis has shown that there are several reasons for DSOs to leverage the market power that arises from their control over the last mile and the meter data in the complementary applications market.

With regard to the second criterion of the European Commission’s three-criteria test, the analysis of the German electricity and metering market showed that competition is very unlikely to constrain DSOs’ substantial market power. This section has additionally shown that there are several reasons for DSOs to engage in anticompetitive conduct, which is likely to result in a decreasing number of potential rivals in the primary market. With regard to the third criterion of the European Commission’s test, the incentives and opportunities to exploit essential facilities justify ex ante regulation. Without appropriate regulatory provisions in place, potential competitors would be deterred from entering the market, because competition law is always associated with significant time lags.
5. Regulatory instruments

Recent experimental analysis backs the rationale that underlies Baxter’s Law by showing that incentives to engage in anticompetitive vertical strategies prevail over possibly greater efficiency gains (Martin et al. 2001, Elliott et al. 2003). It also shows that regulators often develop intermediate regulatory approaches that fall somewhere between “quarantine” and “vertical laissez-faire” (Farrell and Weiser 2003).

Quarantining is a classic structural remedy. It prohibits the monopolist from engaging in vertical integration by enforcing ownership unbundling. However, the monopolist often has the best opportunities and greatest economic interest in a vibrant complementary market (Farrell 2003). Unfortunately, structural remedies preclude any of these integrative efficiencies (Joskow and Noll 1999). Regulators therefore seek to develop compromise approaches to have the “best of both worlds” (Farrell and Weiser 2003). On the one hand, these approaches allow a monopolist to integrate vertically. On the other hand, they aim to ensure that a monopolist does not abuse its position through conduct remedies. In the next section, we present and discuss remedies that may prevent critical bottlenecks’ emergence (section 5.1) and assure non-discriminatory access to these facilities (section 5.2).

5.1 Interoperability and data

As outlined in section 3.2, meter data is an essential input for facilitating numerous business processes, as well as new applications’ efficient and seamless functioning. Hence, the data access mode should enable any authorized market actor to compete on a level playing field. Traditionally, DSOs provided metering services and meter data. Therefore, DSOs had exclusive access to the data. Other authorized actors were only granted access upon request, or on a pre-scheduled basis (ERGEG 2007). In an end-to-end smart grid, however, meter data’s reliable and close to real-time 24-hour availability is crucial to enable new business models to emerge in the downstream market. As shown, DSOs have reasons and opportunities to prevent efficient complementary markets from emerging. For instance, they could leverage their control over the data to increase rivals’ transaction costs. They could also define incompatible data formats or interfaces for each distribution area, or they could intentionally delay data access and provision. Hence, to enable an efficient applications market in a future smart grid requires that all authorized parties are guaranteed equal access to an (online) data platform to recall data in

1) as close to real time as possible,
2) a standardized and machine-readable format, and
3) the same granularity in which it is collected (ERGEG 2007).

Furthermore, consumers should have access to this data and determine the respective parties’ data access rights if the information needs go beyond essential data for billing, or essential technical information (Anderson and Fuloria 2010).

Today, data’s availability to independent third parties is still unsatisfactory, due to incomplete unbundling (ERGEG 2007). Several regulatory agencies have recommended establishing an independent data platform accessible to third parties, or have already established such a platform. Others have suggested that the function of data collection, management, and access should be completely decoupled by establishing an independent and neutral data service provider (ERGEG 2010, FCC 2010, OFGEM 2010). Either approach
could be effective to guarantee efficient and non-discriminatory access to meter data. Moreover, an independent single platform provider may be able to provide the data more cost-effectively, due to economies of scale. This provider can also perform tasks such as meter registration and consumer switching (OFGEM 2010).

Data’s seamless exchange requires open and non-proprietary standards and communication protocols that allow each component and actor within the smart grid to communicate end-to-end. As mentioned before, protocols and standards can resemble essential inputs (Renda 2004, Renda 2010). Whenever standards are regarded as essential, they point to a market with intra-system competition. In such a market, firms compete with each other on the level of components within a particular system. Dependent on the degree of interface information availability, systems are distinguished as either open or closed. Open systems benefit modular innovation, the number of potential market entrants, and market dynamics (Nelson and Winter 1977, Langlois 2001). If intra-system competition is to work efficiently, it requires a degree of openness and modularity (Langlois 2001). In respect of the research context, DSOs may use protocols and standards as “strategic weapons” to build closed systems in which they safeguard interface information. In order to prevent this threat ex ante, there is a wide consensus among policy makers, regulators, and scholars that smart grids should be open and modular (Brown et al. 2010, ERGEG 2010, NIST 2010).

Hence, governments around the globe are fostering the emergence of open smart grid standards to ensure interoperability between components. These efforts are mostly coordinated by standard developing organizations in an attempt to identify or develop open and non-proprietary standards and protocols (see NIST 2009, DKE 2010, ENSG 2010, METI 2010, NIST 2010). The majority of these standardization processes rely on a consensus-driven approach. The aim is for various stakeholders, such as experts from industry, academia, governments, and associations, to agree on standards and protocols (Brown et al. 2010). While these attempts and standardization in general are contentious issues within the literature (Farrell and Saloner 1986, Buxmann et al. 1999, Picot et al. 2008, pp. 54), the social benefits are very likely to outweigh the costs as far as smart grids are concerned (ERGEG 2010).

5.2 Last mile

As outlined in the previous sections, once a smart grid’s last-mile infrastructure is rolled out, it becomes an essential facility that competitors cannot replicate practically or reasonably within an acceptable time frame. This results in a lack of competitive entry in the complementary market, which negatively affects investments in smart grids regarding developing more efficient and sustainable energy systems. High entry barriers (as a result of economies of scale and scope, plus high irreversible costs), as well as DSOs’ non-transitory, substantial market power erode the prospects of new entrants replicating the infrastructure to offer metering services and develop new markets for novel services and products.

Thus, leaving access to the essential facility unregulated (which would result in negotiated access) runs the serious risk of monopolization, or of inefficient investment (Cave and Vogelsang 2003). If access is unregulated, the essential facility’s owner can strategically manipulate potential entrants’ build-or-buy decisions through its access conditions (Bourreau and Dogan 2004). Hence, regulatory intervention, in the form of open (or mandated) access, is needed to secure transparent and non-discriminatory third-party access to a smart grid’s last-mile infrastructure.

Open access implies competition based on services, because several companies offer their services using a single infrastructure (van Gorp and Middleton 2010). If there is no infrastructure-based competition, each firm competes by using its own infrastructure. Facilities-based competition is generally considered to better stimulate innovation and competition (van Gorp and Middleton 2010). However, there is a broad consensus that potential entrants should initially be granted favorable access conditions to enable the
emergence of sustainable infrastructure competition, but that these conditions should be gradually adjusted over time. Thereby, entry barriers are lowered because if a rival’s market entry (based on the incumbent’s infrastructure) did not work out, the rival can withdraw without having to write off irreversible costs for infrastructure investments. If the entry does work out, the transitory entry assistance can be gradually withdrawn to increase the entrants’ economic and strategic incentives to invest in their own infrastructure (Cave and Vogelsang 2003). This approach has proved to be effective in the telecommunications sector in which it is known as the “investment ladder” or the “stepping stones” model (Cave 2006). For example, by applying the investment ladder approach in the broadband services market (see Figure 6), competitors could gradually “climb up the rungs” by expanding their consumer base and revenues (Cave 2006). Although competition in the broadband market was solely service-based at the beginning, and the entrants’ business models relied purely on resale, rivals have now replicated everything but the local loop. Rivals have progressively added more value to the product and decreased their reliance on the incumbent’s infrastructure (Cave 2010). Consequently, within the EU15, the majority of new entrants’ preferred form of access has switched from resale to either bitstream access (17%) or local loop unbundling (56%) (Cave 2010).

Figure 6. Ladder of replicability for smart metering (broadband in parentheses). (based on Cave 2006)

While the investment ladder model has substantially increased the number of competitors, it has been criticized for causing inefficient entry (Renda 2010). Indeed, empirical studies suggest that an aggressive access policy leads to excessive service-based competition, which results in lower consumer prices. It also distorts the incumbent’s and competitors’ incentives to invest in their own infrastructure (van Gorp and Middleton 2010). In contrast, countries that place stronger emphasis on infrastructure competition are found to have higher prices, but also a better infrastructure (Renda 2010). Hence, open access policies should not be an argument for low prices on a “carte blanche basis” (Cave 2006). Rather, open access policies must balance between encouraging investment and innovation on the infrastructure level, and promoting service-based competition and application-level innovation in the short run.

The telecommunications sector’s experience suggests that the primary focus with regard to the smart grid’s last mile should be on attracting a reasonable number of entrants in the downstream applications market to promote service-based competition. Consumers will benefit from new applications, which will increase energy systems’ efficiency and sustainability. According to findings regarding the development of competition and the degree of replicability, both of which depend on how technology develops and how much it costs,
Regulators have to time the raising of bars to entrants carefully by means of dynamic pricing or sunset clauses. This practice will stimulate investments in progressively less replicable assets (Cave 2006). Similar to what occurred with broadband services, the result of increasing access prices should allow entrants to gradually acquire more of the capital assets of the smart grid’s communications infrastructure, as illustrated in Figure 6. With their decreasing reliance on the incumbent’s infrastructure, the entrants’ differentiation potential increases progressively, since they are able to invest in innovative technologies that may offer higher service quality or increased cost-effectiveness (Bourreau and Dogan 2004). Thereby, the ultimate goal of access regulation can be achieved: the “emergence of self-sustaining effective competition and the ultimate withdrawal of regulatory obligations” (ERG 2004). However, regulators should always keep in mind that removing compulsory access rules too early may negatively impact initial achievements.
6. Discussion and conclusions

Seamless end-to-end communication is a prerequisite for the improved coordination of electricity generation, transmission, distribution, and consumption, as well as for the emergence of new business models. This paper sought to identify facilities that can be classified as essential for smart grids (RQ 1). We examined whether the firms that own the bottlenecks have reasons to engage in exclusionary behavior (RQ 2). We based the analysis on theoretical arguments and empirical observations. Furthermore, we presented and discussed the applicability of regulatory instruments which might help establish equal access to such essential facilities and prevent incumbents’ discriminatory behavior (RQ 3). We subsequently discuss the findings regarding the three research questions that guided this paper.

We identified three critical bottleneck areas that serve as essential inputs for competitors in the downstream market and may be used anti-competitively. In order to qualify as essential facilities, three criteria have to be met (European Commission 2003). The first criterion refers to high and non-transitory entry barriers, which applies to the smart grid’s last mile. Once DSOs have rolled out the new metering infrastructure, any new entrant would be confronted with significant and irreversible costs that the incumbents do not have to bear (Stigler 1968). Furthermore, new entrants to the metering market would face short to medium-term drawbacks, such as variations in the economies of scale, higher advertising spending, and capital costs, which the incumbents would not (e.g., Bain 1956, Schmalensee 1989: 968). Hence, the first condition of the three-criterion test applies to the research context, as duplicating the facilities’ functionality would be uneconomic and unfeasible for competitors in the complementary market (European Commission 2002).

The second criterion refers to market structures that do not tend towards effective competition in the relevant time horizon. As outlined in section 2.2, the German electricity market is characterized by a high concentration in all segments. This limits the number of potential entrants. To date, there is no consumer demand for the SMT. Therefore, the entrance of a sufficient number of rivals in the metering provider market is extremely unlikely. Furthermore, these rivals would not only have to enter the upstream market, but also the downstream market, as they will be confronted with an insufficient supply of complementary applications. Entry into the complementary market will be complicated even further by indirect network effects. Hence, it is very unlikely that competition will constrain DSOs’ market power and, consequently, will be stable in a foreseeable future. The second criterion is therefore also met.

The third criterion deals with competition law’s capability to correct market failures. Competition law serves to justify ex ante regulation. As stated in section 4.2, DSOs have various incentives to engage in exclusionary and anticompetitive behavior, such as refusals to deal with certain actors, exclusive dealing arrangements, and predatory pricing. Hence, the likelihood of inflated access charges and discrimination is very high. In addition, competition law is associated with a significant time lag. Consequently, the application of competition law alone will not suffice to address market failures in order to guarantee rivals’ reliable, efficient, and non-discriminatory access to the facilities within a reasonable time frame.

Although the three-criteria test is controversially discussed in the literature, if properly applied, it provides good guidance to identify facilities that need to be ex ante regulated (Blankart et al. 2007, Renda 2010). However, with respect to data access and the definition of a basic set of open and non-proprietary interface standards and data protocols, one could argue that ex ante regulation is not indispensable. Competition law might suffice to correct possible market failures. However, an excessive emphasis on competition distracts from the aim to increase energy efficiency and environmental sustainability (Hertin 2004, Kemfert...
Similar objections can be raised with regard to entry barriers’ non-transitoriness. As replicability is generally not a binary variable (Cave 2006), one can argue that the last mile in a smart grid is replicable if entrants can find other technical ways to bypass the facility. However, similar to telecommunications (Wernick 2007, pp. 190, Picot 2009, Renda 2010), DSOs’ market power alone already justifies (asymmetric) regulatory intervention.

According to the public-interest theory, the paramount societal interest is to realize the environmental benefits that can be gained from SMT’s widespread diffusion. Therefore, we argue that new market entrants have to be guaranteed a transparent and stable regulatory environment. Access rules regarding essential inputs are important elements of such a regulatory framework which also facilitates the emergence of intra-system competition (de Bijl 2005). As illustrated in section 4.2, if there are no effective regulatory provisions in place, DSOs might discriminate against complementary products’ unaffiliated producers, or even prevent them from gaining access to essential inputs. The absence of complementary applications would then negatively affect the amount of independent innovation at the application level, since independent third parties would face (1) significant uncertainty about the future competitive environment, (2) threats of discrimination, which will reduce profits, and (3) the risk of DSOs imitating third parties’ innovations (van Schewick 2007). From a social welfare perspective, a decrease in independent applications is only relevant if DSOs cannot offset this reduction. Owing to a smaller number of innovators, the amount and quality of innovation are also likely to be reduced (van Schewick 2007). Furthermore, DSOs have no economic interest in developing applications that decrease their outside revenues. However, for independent innovators, such applications would be very compelling. Application level innovations would also spur intra-system competition, which is crucial to increase consumers’ interest in adopting and using the SMT.

A sufficient condition for justifying regulatory intervention is met if societal benefits outweigh the costs. Thus, regulators have to trade off regulatory interventions’ benefits and the associated costs. As already outlined, the benefits gained from regulatory intervention include increased competition and application level innovation. From a public interest perspective, this increase in competition and innovation is only relevant if it increases social welfare. While this relationship is theoretically ambiguous (Tirole 1988, Katz 2002), in the study’s research context, the presence of uncertainty and uncompensated spillovers is likely to result in a supply level below the social optimum. Furthermore, a smart grid can be considered a general purpose technology that will be required to drive future economic growth (Bresnahan and Greenstein 2001, Larsson 2009). Regarding the costs, regulatory intervention is associated with a distortion of incentives to invest and innovate in a smart grid’s communication layer. Furthermore, regulation itself incurs costs. While the latter may be negligible, the former needs particular attention.

While the literature suggests that incentives to invest in a general purpose technology prevail over those for application-level innovation (Bresnahan 1998, pp. 10, Weiser 2003, pp. 79), we state that, with respect to the research context, the investment ladder approach (Cave and Vogelsang 2003, Cave 2006) provides an adequate regulatory instrument. However, studies report negative correlations between mandatory sharing of essential facilities and investment incentives (Grajek and Röller 2009, Wallsten and Hausladen 2009). This means that once rivals have been granted “easy” access, this assistance should be gradually withdrawn to encourage firms that profited from low entry barriers to invest in their own infrastructure. With an increasing number of independent firms investing in the last mile and in the entire smart grid communications infrastructure, the threats of DSOs’ discrimination will be gradually superseded. Thus, through infrastructure competition, the primary metering provider market will assume a structure in which the last-mile infrastructure’s owners will have an economic interest in providing independent producers favorable access conditions, since these producers can internalize complementary efficiencies. Once dynamic market
forces have been stimulated (Schumpeter 1934), regulation can be progressively removed as it was already partly done in the telecommunications sector (Cave 2010).

In addition to the study’s limitations that we have already mentioned, other shortcomings have to be considered when interpreting the findings. Although the analysis is grounded in an extensive literature review, and is based on empirical evidence from various scientific domains, our normative research approach can only establish the basis for future research. Our analysis was grounded in public interest theory. Therefore, our aim was to produce a positive theory based on a normative analysis. Accordingly, we proposed regulatory measures that can correct market failures and prevent discrimination in a future smart grid. Some scholars, however, criticize public interest theory because it claims that “regulation occurs when it should occur because the potential for a net social welfare gain generates a public demand for regulation” (Viscusi et al. 2005). However, empirical evidence suggests that this proposition is not always true, as regulatory policy is sometimes “captured” by the industry it should regulate (see Stigler 1971, pp. 3, Picot and Landgrebe 2009). However, public interest and capture, as well as other regulation theories (e.g., economic, credible commitment) have been condemned for what Christensen (2010) calls “plausible logic, questionable validity” (see also Viscusi et al. 2000, pp. 330). In contrast to capture theory, however, the shortcomings of a normatively oriented research approach based on public interest theory can, in terms of validity, be addressed by involving a broad range of insights and stakeholder interests, as done in this study. Nevertheless, further studies are needed to apply other theoretical and methodical approaches. This will help scholars generalize and further develop the propositions.

Our investigation was based on current regulatory provisions and assumptions on the German metering market’s future development and the roll out of an AMI infrastructure, with smart meters. Although the assumptions rely on empirical evidence, they entail a certain degree of uncertainty. Therefore, our propositions may need to be realigned if certain hypotheses do not apply. The examination of DSOs’ incentives to discriminate has highlighted several situations in which DSOs may engage in discriminatory practices. Whether all of these conditions will occur in the real world and all DSOs will behave accordingly is an empirical question.

Despite these limitations, our study provides an in-depth analysis of potential monopolistic bottlenecks that can reduce the socially optimal amount of innovations at the smart grid’s application level from where—similar to the internet—innovations are expected to come. This study thus contributes to the political and scientific discussion on whether regulatory actions are required to ensure essential facilities in a smart grid and the instruments required to help address market failures (Pérez-Arriaga 2009, ERGEG 2010, Hempling 2011).

In sum, our analysis shows that the presence of numerous trade-offs provide no simple answer to the question of whether ex ante regulation is necessary, and it shows that it is impossible to find an easy solution to the problem of configuring regulatory remedies. The proposed regulatory instruments can be compared to the successful regulation of the telecommunications sector. They seek to find a “third way” between quarantine and vertical laissez-faire, in which integrative efficiencies are allowed to emerge through open access rules. However, regulators might consider structural separation between the distribution grid operation and metering service provision a more effective remedy for discriminatory practices.

Based on the study’s findings, future energy regulation should reconsider current regulatory barriers to remove problems that stem from misaligned incentives, as highlighted in section 4.2. In particular, DSOs, which are the most affected parties in energy supply systems’ transition, should be provided with appropriate economic incentives to promote upgrading to smart grids. DSOs should also be incentivized by decoupling revenues from the amount of electricity delivered to consumers and fostering a more efficient systemic and
commercial DER integration by more extensively including measures for energy losses and quality of service in RPI-X regulation than is currently done (Cossent et al. 2009, Langniß et al. 2009, Niesten 2010). Moreover, in order to encourage more R&D and risk taking with new smart grid approaches, national regulatory authorities should consider following OFGEM’s example by creating an “Innovation Funding Incentive” that allows DSOs in the UK to spend .05% of their regulated return on R&D projects, of which 80% can be passed on to consumers (Bauknecht et al. 2007, OFGEM 2009).
References


