1 Introduction

1.1 Let’s Start with Some Intuition

Every time you turn on the light, take a shower or ride on a train, you are using a regulated service. Do you ever wonder how it gets funded? Are you paying for this service as a user or are you paying as a taxpayer? For sure, someone has to pay for it. Whether the service is funded by users or taxpayers, it has to raise enough revenue to ensure it is financially and/or fiscally sustainable. But revenue cannot be the only dimension against which these prices should be assessed. The price should also send the right signal to users. If the price you pay for water is too low, you are likely to stay longer than necessary under the shower. And if the price is lower than the costs involved then water companies may not have a strong incentive to invest. This might not hurt you in the short run, but it could eventually result in water rationing. Farmers may not be able to irrigate enough and as a result food prices might increase. Equivalently, if the price of public transportation is too high or there are no tolls on highways, you may prefer to use your car. Air pollution is likely to increase as a result. And if the price of water or electricity is set high in order to get people to consume less, that may exclude some poorer users, who need the service as much as the richer users (think of water).

Clearly, getting the price right is not simple (and it is even more complex when subsidies are an option to finance the service). Adding to these dimensions the strong role of monopolies (or quasi-monopolies) in the production and delivery of these services explains why, in most countries, electricity, water or transport prices tend to be subject to regulation by the government or a regulatory agency.\(^1\) How theory argues this regulation should be done and how policy experience suggests it can be done is what this book is all about. In the rest of this chapter, we look in more detail into each of these dimensions before revisiting some of the intuitions more conceptually in the following chapters.

1.1.1 Every Time You Turn on the Light, Take a Shower or Ride on a Train, You Are Using a Regulated Service

Significant investments are needed to allow you to use these services. A firm, public or private, has to invest in the facilities needed to produce the electricity and then rely on

\(^1\) Similar characteristics also hold for other public services such as hospitals.
transmission lines to bring it close to your location. A firm, maybe another one, has to maintain the local delivery lines to your home to allow you to plug in your phone or laptop safely when it suits you. The water you use to take your shower or to drink has to be pumped from ground sources or rivers. Before you use it, it has to go through expensive treatment facilities to make it safe for consumption. Then it is channelled to your home, often from quite far away, through large pipes which have to be maintained and repaired regularly. Similar investment and maintenance needs apply to the train stations and tracks. The point is that most of the regulated services you enjoy are heavy users of capital and labour and other basic inputs to ensure their operation. Many are also set in public space and rely on complex planning and coordination between the various users of shared public assets, which can include air, land and water assets.

1.1.2 Someone Has to Pay for the Regulated Services

The investment, maintenance and operating costs of regulated services are significant in most circumstances. These cost levels, however, do not only depend on the cost of the physical inputs. Because these services involve long-lived assets amortized quite slowly, they are particularly sensitive to financial costs associated with long-term borrowing requirements. The expected return for public or private banks or investment funds providing or lending the capital is an important cost driver in these activities. These costs show up, somewhere, in the price paid by users or in the subsidies the government needs to allocate to the sector.

If, on average, the price paid by users fully recovers these costs, it is essentially a user fee. But in many of the sectors providing services of public interest, the price is often set below costs. In that case, subsidies are needed and these are financed by current taxes or public sector borrowing (i.e. future taxes). For instance, in most countries, rich or poor, the investments made to ensure that we all have access to water are usually largely paid for by current and future taxpayers rather than users. The price we pay often only covers the costs of operating and maintaining the assets. It seldom recovers the amortization of the investments made. Even more rarely does it reflect the scarcity value of water, which, in an ideal world, should be accounted for in the price we pay for these services. Such formal or informal subsidies make these services more affordable. But, as will be discussed in various chapters, while they are well intended, and certainly necessary for the poorest populations, they also tend to lead to

---

2 For a given service, costs can differ a lot across service areas since they depend also on factors such as the technology used or on the geographical and geological characteristics of the region in which the service is provided. Building rail tracks in the Swiss mountains is quite different from doing so on the flat lands of the Netherlands.

3 Some developing and emerging economies rely on a third de facto financing source. In Mexico, for instance, during the 1990s, poor regions started to rely on voluntary community work to maintain roads and thereby reduce the need to depend on cash subsidies. Similarly, in the slums and many poor rural parts of Africa, Asia and Latin American countries, the basic and light repair works on water systems are performed by trained local citizens committed to contribute in kind to the maintenance of what they see as public goods. As a result, studies show that most citizens on earth contribute to public goods, either through the taxes they pay or through their in-kind contribution.
overconsumption by users who do not face the real cost of their consumption and this adds to the demand for underfunded investment.

1.1.3 The Price Paid Needs to Account Jointly for Financial/Fiscal Viability, Efficiency and Equity Concerns

Regulated prices have multiple purposes. Combined with subsidies, they are essential to ensure the long-term economic, social, financial and fiscal viability of the service. The discussions between regulators and regulated firms tend to focus on the role they have in the determination of the financial returns to public or private investors and the residual costs to taxpayers if subsidies are needed. Regulators, and increasingly many stakeholders in civil society, also spend (or at least should spend) time analysing the role of prices in signalling the effects of consumption decisions on the use of natural resources. Many stakeholders think as well about the impact of prices on the affordability of the service, to avoid penalizing the poorest users. Moreover, many analysts highlight that prices have to be fair to tomorrow’s stakeholders. It is indeed easy to forget – or to choose to forget! – that if investment is not made today services will not be available tomorrow. From a financial perspective, the prices need to ensure that the investors get a return consistent with the risks they are taking – accounting for subsidies. And from a fiscal perspective, they need to deliver on many policy responsibilities, including accounting for the distortions associated with the need to raise taxes to pay for subsidies, as will be discussed in some detail in subsequent chapters.

Ideally, prices should also account for the difference between the social return and the financial return from the regulatory decisions. For instance, the net costs of environmental externalities should be reflected in the price of the services you are consuming if they are not subject to a specific tax. This is seldom the case. This explains, for instance, why it is still much cheaper to travel short distances by plane than by train. In a nutshell, getting the price right implies addressing the often conflicting concerns of many different stakeholders, including some that are yet unborn.

1.1.4 Pricing Is Influenced by the Strong Role of Monopolies in These Industries

Ensuring that the price of a service accounts for all these concerns would be challenging in any market, but it is particularly hard for industries in which the services are

---

4 The decision to omit the difference between social and private returns to decisions has dominated many policy activities, and not just regulatory decisions. For instance, most of the private sector development vice presidencies of international organizations work on the assumption that the private returns are a lower bound for the social return. This implies that they rule out the possibility of a negative impact of environmental or other externalities on social welfare when assessing the price of a service. The pricing is usually disconnected from the information provided by the environmental impact assessments required by these organizations for the projects they finance. This disconnection has been a long-lasting debate within these organizations. For now, in practice, the private sector perspective, anchored in a socially costly conceptual omission, continues to define the policy perspectives. This perspective also spills over in debates on the optimal design of regulation of private firms or commercialized state-owned enterprises.
often best provided by local, regional or national monopolies, as is the case for many of the regulated public services. Monopolies are a source of concern if they enjoy enough discretionary margins on their production decisions to underprovide quantity and/or quality and to overcharge. Monopolies also enjoy an informational advantage over all other stakeholders that they can use to influence the setting of prices or quantities to their benefit. For instance, they know more about their actual costs, and their potential cost savings from changes in technologies or improved management efforts, than consumers do. They can use the informational advantage to inflate their costs report and hence secure higher compensation for their services than they deserve, as a way of increasing their profits.

Regulation needs to arbitrate between all these concerns, and it is not simple. It has to account for the fact that not all participants have access to the same information and that the abuse of market power is a serious threat in these sectors in most countries. This is why, historically, the price of basic services has usually been, directly or indirectly, controlled by the State in some form. In other words, this is why the prices charged by monopolies are typically being regulated. Regulation is needed to minimize the risks that the de facto power a firm has over a market is not abused to the disadvantage of users (domestic and industrial) or taxpayers.⁵

### 1.1.5 What Does Regulation Deal With?

Ultimately, in very broad terms, regulation could be defined as a policy tool used to ensure efficient price, quantity and quality outcomes in utilities services that are fair to all stakeholders when markets on their own fail to do so. And this applies whether the tool is implemented by a national, regional or local public authority. In practice, regulation ends up being an essential policy tool. To make the concept as practical as possible, it may be useful to point to some of the most basic questions it is expected to address when competition is limited:

- **What is a fair price?**
- **How can a price possibly be fair to all stakeholders?**
- **Should it be based only on the level of costs claimed by the service providers?**
- **Is the cost data reported by the firms the right one to compute the price?**
- **How much margin is there in adjusting quality to influence cost levels and hence prices?**
- **How are prices, subsidies and expected returns connected?**
- **How well are social concerns and externalities accounted for in prices?**

⁵ The same concerns apply if the market is only catered by a few firms able to collude (i.e. if the market structure is an oligopoly), and in that case regulatory rules also need to be designed to make sure that an incumbent or a group of incumbents do not manage to block potential entrants to the market.
1.1 Let’s Start with Some Intuition

**How is investment taken into account in short-term regulation?**

**How much does the nature of consumers and investors matter to the way prices are set?**

**Should regulation be implemented by the government or by an independent regulator?**

**Should regulation be local, national or international?**

**Does the prevalence of corruption in a country influence the answer to these questions?**

**More generally, is the answer to these questions the same for developed and developing or emerging economies?**

By the end of the book, you should have an informed view on how to address each of these essential questions because the economic theory of regulation offers answers, or at least guidance, on how to deal with them. In many cases, there are various ways of answering any of the questions. In each chapter, we summarize the options available to regulators suggested by theory and practice, as well as the limits of our collective knowledge on these issues, and therefore the need for future research.

1.1.6 The Choice of the ‘Optimal’ Regulation Also Depends on the Institutional Context

The degree of adoption of technical guidelines and rules derived from theory varies significantly according to the specificities of the informational, institutional and political environments in which regulatory interventions have to be designed and implemented. The preferred approach to regulation depends on the extent to which countries trust markets and institutions involved in the design, financing, implementation and enforcement of regulation. When trust among consumers, taxpayers and firms is strong, rule-based regulation tends to prevail over politically driven regulatory decisions. When trust is weak, political solutions tend to prevail over technical rule-based decisions, often deepening the distrust of many stakeholders.

The option to be somewhat flexible with rules may be useful when significant economic shocks such as brutal devaluations, changes in interest rates or social crisis hit a country, making extreme policy reversal such as renationalization less likely. But it also makes the regulatory process much more subject to political interference and sensitive to electoral cycles. For instance, a change in government can lead to price freezes resulting in financing gaps despite commitments to cost recovery made by the previous political administration. This is what happened in 2001 in Argentina, for instance. It led to legal battles between the government and service providers that were sorted out by international arbitration courts, sidelining national regulators and overruling previous regulatory commitments.

The dual characterization of countries between rules versus discretion is clearly extreme. Many countries sit somewhere in the middle of the trust ladder and of the
scope for discretionary leverage to change rules. Moreover, regulatory preferences can be a moving target as they tend to evolve over time with the political context. But as discussed throughout the book, the risk of abuses resulting from the difficulty of achieving full protection of regulators from undue political interference is serious in any country. When regulators are not independent and free from political pressure, they can indeed be pushed to adjust to changes in political preferences. This is true whether the country is a young, unstable democracy or an old, established one.

Some of the instability of regulation is the result of ideological fads. For instance, less than thirty years ago, when the privatization wave started with the British reforms, nobody focused on the fact that the privatization of a public electricity or railway company could actually mean the acquisition of a national public enterprise by the public enterprise of another country. Hardly anyone really worried about this possibility in the 1990s when Prime Minister John Major cleared the transformation of the sector. In 2019, in a context in which the British population had voted to exit the European Union and the perceived German influence on the region, the fact that Deutsche Bahn (DB), the national German railway company, controls five British rail services (through its Arriva subsidiary acquired by DB in 2010) had become an issue in regulatory discussions.

In addition to the effects of ideological swings on preferences, there are also more technical reasons why countries may differ in their choice of design and practice of regulation. For instance, size matters. Mali’s regulatory challenges are quite different from those Brazil has to address. And those Portugal must deal with are quite different from those of the United States, Canada or Australia. In other words, the regulatory needs and constraints of a small market are not the same as those of a much more complex large market in which regions have a lot of autonomy to decide on how they regulate. Market size defines the distribution of regulatory responsibilities across government levels and sometimes across government agencies within the same government level. Market size defines the margin a country has to attract domestic and international private financing. It also often impacts the level and composition of demand. Differences in income levels may explain differences in the willingness to pay versus the ability to pay and demand different regulatory designs. Circumstances and context clearly matter too, as illustrated throughout the book. But before getting into the technical details of what can be learned from economic theory and policy

6 Full independence does not mean that regulators are not subject to audits and accountability. The design of regulatory institutions is discussed in detail in Chapter 14.

7 Not all changes are politically driven. The fast-changing digitalization of many regulated industries is leading some regulators, notably in the transport sector, to consider adaptive regulation. The evolution of the legislation adopted by various countries to frame the use of autonomous public transport vehicles is anchored in ‘soft laws’ designed to make the adaptation of rules easy enough to match the fast-evolving technologies. These soft laws are more flexible than the stricter approaches used in technologically more stable environments. They also allow the industries and the regulators to converge towards realistic and fair regulatory standards in a much more transparent way.

8 See Clifton et al. (2006, 2016) for a detailed discussion of this phenomenon in Europe.
practice, it is useful to set the stage in broad terms. That is the main purpose of this first chapter.

To complement the intuitive discussion, we provide some basic statistical information on the relative and absolute economic importance of infrastructure industries to show how large the stakes are in terms of affordability, investment needs and associated costs across the world. Next, we summarize the main case for regulation made in the economic literature. This is followed by a first discussion of the continuous gap between the theory and practice of regulation and the need to reconcile the two more systematically to account for a more realistic sense of what regulation can and should deliver. The need for more pragmatism in the development of theories and a faster adjustment capacity is reinforced by a brief discussion of the evolution of the context in which regulation is expected to deliver, suggesting that theory continues to lag behind rather than lead real-world discussion on some key issues. The chapter concludes with an overall presentation of the content of the book and a discussion of how it can be used in different ways by the diversified audiences interested in regulation.

1.2 Some Basic Statistical Evidence on the Importance of Regulated Services

One of the main challenges for the design of regulatory policy is the need to work with partial or incomplete data on the various economic characteristics of the sector. It is, indeed, often only when frequency and duration of traffic jams explode, when a bridge collapses or when there is a significant power outage that additional audits are conducted to collect the data that would have helped identify the risks of technical failure well ahead of the incident. Despite these limitations, there is enough data collected by national governments and international organizations to demonstrate that the services provided by regulated industries are economically, socially and politically important. For instance, we know how much consumers spend on infrastructure services — that is, the sum of expenditures on electricity and gas, information and communications technologies (ICT), transport (including ownership of vehicles, their use and the use of public and private transportation) and water and sanitation (W&S) services. We also have a good sense of how much countries actually invest against how much they need to invest in these activities to sustain growth and meet demand. Any of these indicators illustrates the importance of these services and can be used to highlight the role of regulation in the effective delivery of public services. In what follows we focus on household infrastructure expenditures and investment in the sector.

While this book takes an economics perspective on regulation, the legal public sector management and political perspectives are just as important. They will only be addressed in this book when they point to dimensions ignored by economists in their assessments. For useful sources to give a better sense of these approaches, see, for instance, Baldwin et al. (2012), den Hertog (2012), Gomez-Ibanez (2006), Levi-Faur (2011) or Lodge and Wegrich (2012).
1.2.1 How Much Do We Spend on Infrastructure as Consumers?

This is an important question, as how much we spend on a regulated service is one of the core indicators used to assess the extent to which regulation accounts for social concerns. It turns out that we all spend a lot on these regulated services and that some income classes end up spending a lot more than others in relative terms, as will be discussed in detail in Chapter 9.

In 2017, on average, depending on where we lived, between 16% and 26% of our total consumption expenditures went to pay for infrastructure.\textsuperscript{10} For developed economies, the differences across countries varied from around 16% for Australians on average to around 25% for American and European households. Canadians were closer to Europeans at about 22%. For developing, emerging and transition economies, the range was somewhat narrower than for developed economies. The average household spend on all infrastructure services was between 16.3% in South Asia and 21.7% in Latin America.

For a vast majority of families, this is a significant share to allocate to services that are hard to compress, in particular for the middle- and lower-income classes. This is also a big part of the reason why these services are so politically sensitive at all stages of development. To make this relevance as concrete as possible, think of what it would mean if the costs of these services were cut by, say, 10–20% because regulation managed to make the service providers become more efficient. It would be equivalent to an increase in the total purchasing power of families of 1.5–5%. This is significant and can come from various sources, ranging from improvements in the management of regulated companies leading to cuts in operational costs to the adoption of cheaper technologies. Chapter 10 will show that the equity payoffs of these efficiency gains can make a difference to the political viability of regulatory reforms, but this requires a design of regulation which ensures that the gains are shared fairly. The basic data on expenditure shares allocated to regulated services suggests that the efficiency–equity trade-off should be a core focus of the design of regulation at all stages of development.

To get a more precise sense of the relative importance of these trade-offs across sectors, it is useful to take a look at the disaggregated expenditure data. Once again, the most obvious observation is the strong heterogeneity across countries and regions. Nevertheless, it is on transport (including all expenditures linked to vehicle ownership) that people spend the most. In developing, emerging and transition economies, on average, in 2017, households spent from 6.7% of their income in South Asia to 12% in Latin America on this sector. Energy absorbed from 4.5% of the household expenditure in Sub-Saharan Africa (SSA) to 6.9% in Eastern Europe. ICT expenditures vary from 3.2% in Sub-Saharan Africa to 4.4% in Eastern Europe. Finally, W&S absorbs from 0.4% of income in South Asia to 1.5% in the Middle East.

\textsuperscript{10} The data for Organisation for Economic Co-operation and Development countries is from the OECD website on household expenditures and the data for developing and emerging economies is from the World Bank household expenditure website. All of the data is for the year 2017.
In developed economies, the patterns were somewhat different. As in the rest of the world, transport dominates with on average 10–11% of household expenditure but the share spent on energy is also quite significant. In Europe, for instance, it stands on average at 7.5%. And the share spent on W&S in this region is also much larger at about 4%. At 3–4%, Europeans’ expenditure on ICT is closer to the figures noted in poorer regions. Australia is somewhat of an outlier since its households spend on average only about 1.2% of their disposable income on W&S, 2% on energy and 2.2% on ICT (it is in the average range for transport at 10.2%).

Table 1.1 summarizes these figures to make the comparison across country groups and sectors easier. If the shares noted in developed economies serve as a leading indicator of what is the likely evolution of expenditures in developing and emerging economies, the demand for energy and water should be expected to grow the most. This is consistent with the evolution of demand resulting from the fast-emerging middle class in these regions. This is not a minor concern in practice since both are central to the debates on the environmental sustainability of growth.

Regulation is not neutral to the evolution of these shares and what they imply for the environmental sustainability of household infrastructure consumption. The expected evolution means that the regulatory pressure is likely to have to adapt, and possibly increase, as demand management becomes a necessity, unless new resources and technology solve the problem. For now, the most likely scenario is that demand is going to be rationed unless regulators come up with politically sustainable ways of managing it. In Europe, the European Commission has been pushing for a national regulator to take on the challenge. In all cases, regulated prices (including subsidies) and the definition of harmonized standards will be central to the solution but there is evidence that not all countries will adapt at the same speed. In the USA, states differ in the pace at which they adapt, mainly based on local needs. For instance, California has a long tradition of being ahead of other states in testing and mainstreaming regulatory innovations to improve the sustainability of water and energy consumption, simply because these resources are scarce locally. States with more abundant resources are less keen on using regulation for conservation and optimal rationing.

Table 1.1 Order of magnitude of share of total household expenditures in infrastructure services

<table>
<thead>
<tr>
<th>Total infrastructure (%)</th>
<th>Electricity and gas (%)</th>
<th>ICT (%)</th>
<th>Water &amp; sanitation (%)</th>
<th>Transport (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developed economies</td>
<td>16.2–26.5</td>
<td>2–7.5</td>
<td>2.2–4</td>
<td>1.2–4</td>
</tr>
<tr>
<td>Developing and emerging economies</td>
<td>15–25</td>
<td>4.5–6.9</td>
<td>3.2–4.4</td>
<td>0.4–1.5</td>
</tr>
</tbody>
</table>

Source: Authors’ compilation from various sources.

In developed economies, the patterns were somewhat different. As in the rest of the world, transport dominates with on average 10–11% of household expenditure but the share spent on energy is also quite significant. In Europe, for instance, it stands on average at 7.5%. And the share spent on W&S in this region is also much larger at about 4%. At 3–4%, Europeans’ expenditure on ICT is closer to the figures noted in poorer regions. Australia is somewhat of an outlier since its households spend on average only about 1.2% of their disposable income on W&S, 2% on energy and 2.2% on ICT (it is in the average range for transport at 10.2%).

Table 1.1 summarizes these figures to make the comparison across country groups and sectors easier. If the shares noted in developed economies serve as a leading indicator of what is the likely evolution of expenditures in developing and emerging economies, the demand for energy and water should be expected to grow the most. This is consistent with the evolution of demand resulting from the fast-emerging middle class in these regions. This is not a minor concern in practice since both are central to the debates on the environmental sustainability of growth.

Regulation is not neutral to the evolution of these shares and what they imply for the environmental sustainability of household infrastructure consumption. The expected evolution means that the regulatory pressure is likely to have to adapt, and possibly increase, as demand management becomes a necessity, unless new resources and technology solve the problem. For now, the most likely scenario is that demand is going to be rationed unless regulators come up with politically sustainable ways of managing it. In Europe, the European Commission has been pushing for a national regulator to take on the challenge. In all cases, regulated prices (including subsidies) and the definition of harmonized standards will be central to the solution but there is evidence that not all countries will adapt at the same speed. In the USA, states differ in the pace at which they adapt, mainly based on local needs. For instance, California has a long tradition of being ahead of other states in testing and mainstreaming regulatory innovations to improve the sustainability of water and energy consumption, simply because these resources are scarce locally. States with more abundant resources are less keen on using regulation for conservation and optimal rationing.

11 UNSW City Futures Research Centre (2019).
Overall, while the data reviewed so far is useful to highlight the social and political relevance of regulation, it ignores some of the other concerns that regulators must address. One of these is the impact of regulatory decisions on investment levels. It is frequently the case that the apparently high costs stated by regulated firms hide higher profit margins allowed by creative cost accounting and corporate financing practices. This implies that getting a firm, public or private, to cut costs is equivalent to cutting its profits. If, in the process, the efficiency gains demanded by regulators cut too much into these profits, its incentive to invest or its ability to do so may also decrease. Unless this can be compensated by subsidies, service coverage and quality may drop. This is why it is also important to track the evolution of investment, subsidized or not.

1.2.2 How Much Do Countries Invest in Infrastructure and How Much Should They Spend?

Tracking investment is needed to assess the extent to which there is a fair regulatory treatment of current and future generations of consumers. Regulatory decisions designed to protect the weakest users in the short run may not always have the intended consequences in the long run. This can be avoided in the design of regulation. For now, however, the investment data available at the global level suggests that there is a regulatory failure. Investment in many regulated industries is lagging demand in most countries of the world. This lag tends to penalize the poor in the short run and all users in the longer run.

Consider the case of infrastructure investment.\textsuperscript{12} Since the mid-2000s, a large number of estimations conducted by consulting firms and multilateral organizations such as the World Bank and the regional development banks suggests that investment levels have ranged from less than 1% of GDP in some of the most developed economies to 10% in some of the developing countries with the lowest initial infrastructure asset stocks or with the strongest commitment to develop their capacity (notably China). The levels for most developing countries are actually much lower: their infrastructure investments are estimated at about 4% of GDP on average. In most cases, these figures are well below what is needed to meet the demand for these services that is expected to come from economic growth. These needs are summarized in Table 1.2. The growing middle class in developing and emerging economies may not get what it wants even if it is willing to pay for it. This is why it is important to make sure that the incentives to invest to meet the growing willingness to pay are one of the mandates assigned to regulation.

The regulatory challenge is not minor. According to the World Economic Forum, if investment trends follow the current paths, the world is expected to face a US$15 trillion gap in infrastructure investment by 2040.\textsuperscript{13} That is close to 20% of the world’s GDP in 2018 to be spread over a twenty-year period, with much higher gaps in the

\textsuperscript{12} For a review of the evidence on the social rate of return of investment in infrastructure see Pender and Torero (2018)

\textsuperscript{13} www.weforum.org/agenda/2019/04/infrastructure-gap-heres-how-to-solve-it/