1. Railway regulation: a comparative analysis of a diverging reality

Juan Montero and Matthias Finger

1.1 INTRODUCTION

Railway regulation is being evaluated all around the world. Over the past 40 years, regulatory reform has transformed railway systems in the most developed nations. Despite divergences in railway systems and the resulting divergences in regulatory reform, many countries are now evaluating the results and thinking about ways forward. The nations that lead in railway reforms, such as the United States, the United Kingdom and Japan, are now evaluating the results. The European Union, in turn, is completing the liberalization of the rail sector. Other countries are considering whether to implement similar reforms, at the same time as experimenting with their own transformations, with interesting lessons for everyone.

The aim of this book is to provide readers with a comprehensive perspective on the debates about railway regulation across the world. Leading academics and practitioners from the main railway countries present the results of 40 years of railway regulatory reform. The chapters describe the current challenges in railway regulation, as well as suggestions for how to address them. Specific chapters are devoted to the most relevant challenges in railway regulation – namely, high-speed services, financing, tendering, technical specifications, digitalization, consumer protection, and the application of competition law.

The scope of this introductory chapter is to provide an analytical perspective on the evolution and the challenges in matters of railway regulation in the world’s main railway countries. To this end, we naturally rely heavily on the different contributions to this book. In the book we will identify common themes, but more importantly, explain the divergences in railway regulation. We posit that these divergences emerged mainly after competition from road and later from aviation took place. Subsequently, many countries undertook railway reforms, the nature of which can typically be explained by the policy choices made by the different countries. In other words, and especially in railways, one size does not fit all.

The remainder of this introductory chapter is structured as follows: in the next section we will describe the evolution of railways from its origins in the UK to “peak rail” in the early 20th century. In most countries, railways were initially built and operated by private companies: governments sometimes financially supported their construction, but most often granted concession to private companies who were in the lead. As railways became increasingly important for national development, whether in terms of transporting raw materials (such as coal), manufactured goods, or people, the national rail network grew denser and railways became subject to government regulation due to market abuse, high rates, and discriminatory rebates. In some countries, railways were even nationalized. The main differences between countries during this period concern the degree to which the network was developed.

Section 1.3 describes how railway systems declined due to competition, first from road and then from aviation. Around the 1940s and 1950s, railway companies across the world were
nationalized, either because they went bankrupt or because they were instrumental for the war efforts. The US was the main exception. This second period can be seen as the key moment of divergence in railways, as some systems evolved into mostly freight systems (North America and Australia), while others evolved into primarily passenger-serving systems (Western Europe and Japan). Objective reasons explain this divergence: communist countries supported railways as the most efficient collective transportation means for both freight and passengers. Emerging countries, which had not developed a dense railway network to begin with, struggled to manage heavily used networks.

In Section 1.4 we describe how the different reform policies originating in the 1980s actually reflected the diverging realities in Northern America and Europe. Horizontal separation and competition was the policy in dense freight systems in North America. Vertical separation and downstream competition in the transport service (but not in infrastructure) was the policy in most European passenger systems.

Section 1.5 will discuss the current challenges to railway regulation, devoting special attention to the latest debates in the US and the UK, both pioneering countries. However, in both countries, following decades of reform, some negative effects become evident and fine tuning appears to be necessary. The situation in the EU is also of interest, as 2020 is the year of full liberalization of passenger commercial services all over Europe. Finally, we will also discuss the latest challenges, namely decarbonization and digitalization.

1.2 FROM THE ORIGINS TO “PEAK RAIL”

Railways have had quite an exciting ride over their almost two centuries of operation. The obvious advantages of railways compared to previously existing transport alternatives prompted a construction boom led by private investors, who were then also operating the railway lines. Gradually, governments supported investments and also started to supervise operations, safety, and rates. During the last decades of the 19th century, operations were rationalized, private companies were consolidated, and more coherent networks were formed upon the isolated railway lines. In some countries, nationalization took place; in others, railway behemoths triggered regulatory reaction in the form of rate controls.

1.2.1 The Origins: Private Rail Lines

The use of “rail” tracks in order to guide wheeled vehicles can be traced back to coal mines in Germany and the United Kingdom as early as the 16th century. Low adherence reduces the energy required to transport heavy loads (Leboeuf, Chapter 23 in this book). Over the centuries, different models of increasingly sophisticated rail transport emerged: tracks transitioned from wood to iron and horsepower evolved to stream-powered locomotives. During these times, railways were mostly reserved to mines, short urban itineraries, and emerging industrial sites.

It is commonly accepted that the first modern railway line was the “Liverpool and Manchester Railway” line, inaugurated in September 1830. It was a double-track intercity railway that provided both freight and scheduled passenger services. It was constructed and operated by a private company with private shareholders, albeit with state support for the acquisition of land.
In 1840, the “Railway Regulation Act” was enacted by the “Queen’s most Excellent Majesty” for “the due Supervision of Railways.” Railway lines would be developed and operated by private companies, but prior notice to government was required, as was the approval of the bylaws of companies intending to run a railway line. Rates would have to be publicly displayed. Safety obligations, inspections, and accident reporting were introduced. In other words, the bases of railway regulation were already defined at that time.

“Railwaymania” exploded, first in the United Kingdom and Western Europe. An unprecedented accumulation of capital empowered by modern institutions such as corporations, stock exchanges, and a sophisticated financial industry allowed for the multiplication of private corporations exploiting small, non-standardized, and mostly disconnected rail lines. Many of these ventures were highly speculative, rapidly leading to a rail bubble.

This British model was exported all around the world. Concessions were granted by governments to private companies, financed by British, French, and German investors: Spain started in 1848 (García Álvarez, Chapter 6 in this book), India in 1853 (Gangwar, Chapter 11 in this book), Turkey in 1856 (Ardiyok and Sesli, Chapter 14 in this book), China in 1876 (Tjia Yin-nor, Chapter 9 in this book), Korea in 1899 (Karlsen, Chapter 10 in this book), and so on.

1.2.2 “Peak Rail”: From Rail Lines to Consolidated Networks

National railway networks were assembled out of the pre-existing isolated, non-standardized railway lines. In this way, railways matured into a complex system of sprawling infrastructures, connecting the most densely populated territories, thus monopolizing transportation of goods and passengers. Reaction to the market power of the network managers triggered the early regulation of this system.

Standardization was the precondition for the creation of a national railway network. An early example was the so-called “battle of the gauges” in the United Kingdom. The Liverpool and Manchester Railway was built with a narrow gauge of 1.435 mm. The Great Western Railway engineered by Lord Brummel was built with a “broad gauge” of 2.134 mm. Gloucester was the first town in which a narrow gauge and a broad gauge converged. The break of gauge obviously hindered the flow of train traffic from one line to the other: passengers and goods had to transfer from one train to another. Finally, “an Act for regulating the gauge of railways” was enacted by Parliament in 1846, imposing the “narrow gauge” as a standard. The “narrow gauge” would become the international standard, but different gauges are prevalent in the Iberian Peninsula (García Álvarez, Chapter 6 in this book), Russia (Saakyan, Chapter 13 in this book), and India (where four different gauges are in operation; see Gangwar, Chapter 11 in this book).

The definition of a standard “railway time,” taking as a reference the time at the Greenwich Royal Observatory (Greenwich Mean Time, GMT), is another example of standardization. Initially, each railway line would define a time of operation, overcoming the traditional time set by the sun in the church tower in each town. However, each line used a different reference, which led to accidents. Companies gradually agreed to use GMT as the common reference, but it was only in 1880 that GMT became not only the standard railway time, but the official time for the Great Britain and later for time zones all around the globe. Similar debates took place in all the countries (Pittman, Chapter 15 in this book for the US).

Over the second half of the 19th century, railways were massively deployed by private companies all over Europe and the United States. Railways reached all the populated areas, span-
ning tens of thousands of kilometers in the largest countries. The leading private companies, always with the support of public authorities, constructed the first transcontinental railroad in the US in 1869, while the Orient Express from Paris to Constantinople was inaugurated in 1883 and the Trans-Siberian line from Moscow to Vladivostok was completed in 1904. The European Alps were crossed for the first time in 1882 thanks to the 15-kilometer long Gotthard tunnel, the longest tunnel in the world at that time (Finger, Chapter 7 in this book).

Thus, networks were gradually constructed from isolated railway lines. As standards were implemented, the connection of the previously isolated railway lines provided obvious benefits for operators and the public in general. Consolidation of the corporations managing isolated lines led to concentration; that is, a reduced number of operators. Such operators were in the position to rationalize the deployment of infrastructure to cover all the territory and to coordinate services so as to provide well-structured time-tables to meet demand in a more systematic way.

Consolidation was also a response to fierce competition between railway companies. In the US, the financier JP Morgan built his fortune by consolidating competing railway companies in financial distress due to competition (Strouse 2000). Railways were the first industry to be “Morganized,” although Morgan also applied this strategy to the emerging telephone and electricity industries, creating such behemoths as AT&T and General Electric.

National economies, at least those in the “railway countries,” gradually became dependent upon the transport services provided by the railway. Only inland waterways provided a competitive alternative to railways, but such waterways were only available on very specific routes. Consolidation somewhat reduced the fierce competition that existed between railway companies and shippers became increasingly dependent upon the monopolistic railway companies. Populist movements, such as the Grangers in the US (Pittman, Chapter 15 in this book), complained about the concentration of economic and political power, that railway tycoons such as Jay Gould (Morris 2005) were considered to be the new “Robber Barons.”

The growing market power of the railway companies was increasingly subjected to regulatory control. The first statutes regulating rail rates were adopted at a state level in the US in the 1870s (Ely 2001), leading to the creation of specialized regulatory agencies. Independent regulatory agencies were created in the US specifically for the regulation of railway services. The Interstate Commerce Commission, created in 1887, was the first such federal independent regulatory agency, with railway services being the main object of regulation (Pittman, Chapter 15 in this book).

Railways reached their peak during the first two decades of the 20th century. A reduced number of consolidated operators, like the Big Four in the UK, and their peers in the US and elsewhere, managed the most extensive railway networks, reaching each corner of the territory. Railway behemoths were often the only available transport solution in a specific area, enjoying a monopolistic position. However, such market power was subject to control by public authorities, which often set maximum and minimum prices, prohibited rebates under strict non-discrimination rules and required governmental approval to abandon particular lines (Pittman, Chapter 15 in this book). Nationalization was proposed and partially implemented during the First World War, not only as a response to private monopolies, but also in light of the role that railways had come to play in transporting troops. In most countries, however, such nationalization was reversed after the war.
1.2.3 Objective Underlying Circumstances

Objective circumstances defined railway systems during the first decades of the 20th century and these circumstances would determine how the different railway systems would cope with the emerging competition from road and later from air transportation.

It is obvious that the larger a country is, the larger the railway network tends to be in terms of the length of the railway infrastructure. The largest countries tend to be those that had the most railway infrastructure. Countries that pioneered industrialization in the 19th century also developed a large and denser railway network. Railways were the most competitive transport system, with basically no alternative, so infrastructure was rapidly developed and completed before 1920. This was the case first in the UK and then in Western Europe (particularly Germany, Benelux, and France) and the US (particularly the East Coast).

Countries that developed at a later stage, once road and air transportation were available, tended to have smaller railway networks, even if they were large in size and had a high GDP. This is the case of Australia, for example. Some emerging markets have not developed a railway system, as is the case of the Philippines, given that they rely on road and air transportation both for passengers and freight.

Policy is also relevant. Communist countries invested heavily in railways, as it was the most efficient transport mode and have thus avoided individual transport by road. This is why former communist countries such as the Czech Republic and Hungary have more developed networks than Italy or Spain, and are actually among the world leaders in terms of rail density.

The use of the infrastructure also depends upon some objective variables. The larger a country, the more intense the use of railways for freight transportation. This is why large countries such as Russia, the US, and China have more than ten times more tons/km than smaller countries (of the European Union) (Saakyan, Chapter 13 in this book).

Large distances also play against rail, where aviation is more competitive. Rail passenger services, in turn, are more competitive in small and densely populated countries, where traveling distances between cities are too short to be covered by plane, where roads are often congested, and where rail can provide high frequencies (Messulam, Chapter 17 in this book). On the contrary, railways can hardly compete with road for passenger transportation in rural areas, as population density is too low. This is why passenger services are popular in Western Europe, Japan, and South Korea.

As nations differ in size, geography, population density, and industrial development, their national rail systems become different, and railway decline has not occurred in the same way in all countries, as we will see now.

1.3 DECLINE: DIVERGING RAILWAY SYSTEMS

As the 20th century evolved, competition from road and then from aviation caused the decline of most rail systems. However, each rail system declined in its own way. Decline led some rail systems to evolve into primarily freight transport systems, as was the case in North America and Australia. To the opposite, other national systems evolved primarily into passenger transport services, as in Western Europe, Japan, and South Korea. Finally, decline was not as relevant in other countries, such as the communist countries, where alternative transport modes were not massively deployed, and emerging countries, where the rail network was not dense
and alternative transport modes evolved slowly. Competition, regulation, and then the Second World War led to the decline of the industry, as well as to its widespread nationalization.

### 1.3.1 Decline and Nationalization

Competition from road and air transportation eroded the monopolistic position of railway companies and both traffic and revenue diminished, damaging their financial situation. However, regulation actually increased rather than reduced. Indeed, as public authorities started to subsidize railway operations, they increased their control over the service. Eventually, railway companies were nationalized in almost all countries in the world, and state-owned monopolies were put in charge of the national railway systems.

Road transportation increasingly provided an alternative to railways. As roads were improved, first around cities and later in-between cities, and as vehicles were cheaper and more reliable, both passengers and freight started to migrate to road transportation. Road transport is better suited to serve low-density routes, particularly the last mile of a door-to-door journey. Road transport is also more flexible. Passengers and shippers do not have to adapt to the routes and schedules services decided by a railway operator. Because road transport companies were lightly regulated, they had the flexibility to manage their rates, offer rebates, and so on. On the other hand, rate regulation imposed on railway companies did not allow them to react to such pricing strategies.

Air transportation was popularized after the Second World War, but mostly during the 1960s and the 1970s. It provided an attractive substitute for long-distance rail trips as it dramatically reduced traveling times. As passengers fled to airplanes, rail frequencies were reduced, which made railways even less attractive and ignited a vicious cycle. Freight was not affected by air transport, as costs were too high.

Railway companies started to rely on public subsidies. Railways had had the support of governments from the outset, in terms of use of eminent domain, provision of land for constructions, and other instruments. However, competition from road and air transportation required even more support, this time mainly in the form of subsidies.

The combined effect of competition from road and air transportation, the rigidity imposed by regulation to meet such competition, and events like wars and revolution led to the nationalization of railways all over the world. The Civil War in Spain triggered nationalization and the creation of the national railway monopoly (RENFE) in 1941 (García Álvarez, Chapter 6 in this book). The French SNCF was created in 1938. The Second World War triggered the creation of British Rail in 1948 and the state-owned monopolies in China in 1949 (Tjia Yin-nor, Chapter 9 in this book). Independence in India led to the creation of Indian Railways in 1951 (Gangwar, Chapter 11 in this book).

Some countries had nationalized railways at an earlier stage, for various reasons. Prussia had nationalized railways in the late 19th century in an effort to promote industrialization. Japan had created a state-owned company managing most of the network as early as 1906, after poor coordination between private companies in the Russo–Japanese War was blamed for the Japanese defeat. Switzerland had also nationalized its railways at the turn of the century for national coherence purposes (Finger, Chapter 7 in this book).

The United States was the exception. Railways have always been under private ownership and operation, with the brief exception of the First World War. US railways certainly suffered the same challenges as in the rest of the world. Financial difficulties were widespread in the
1960s and 1970s. In 1970 the Rail Passenger Service Act allowed railroad companies to hand off their passenger services to a newly formed state-owned company: the National Railroad Passenger Corporation – Amtrak. Still, the infrastructure was privately owned, and the freight services, the most relevant part of the system, were privately operated.

“The best of all monopoly profits is a quiet life,” remarked Nobel laureate Sir John Richard Hicks in 1935 (Hicks 1935, p. 8). In railways, tranquility might have been a profit, but monopoly did not lead to economic profits for the carriers. As decades passed, passengers and cargo shifted to other transport modes, and the costs of maintaining the system increased. Networks were shrunk and infrastructures in the less dense areas were dismantled. The less popular services were terminated. Railways entered into a decades-long decline.

1.3.2 Railway Systems Focused on Freight

Some railway systems have evolved to provide primarily freight services, with extremely limited passenger services, mostly restricted to large metropolitan areas. This is the case in North America, Argentina, and Australia. The United States, Canada, and Australia are large countries with low population density and competitive road and air transport systems. Rail passenger services have just not been able to compete with aviation for long-distance travel, perhaps with the exception of some specific routes such as the Boston–New York–Washington DC corridor.

Railways could also not compete with road in rural and scarcely populated areas. Even transport in metropolitan areas became a challenge, as sprawl made it difficult to serve the last mile in suburbia.

Consequently, during the second half of the 20th century, most passenger services were simply terminated. Passenger services became limited to some long-distance and suburban services in the large metropolitan areas that had developed during the 19th century. Rail services are scarce, even in large metropolitan areas, notably those that had developed at a later stage, such as Los Angeles.

Freight was not subject to the same dynamics. The economies of the above-mentioned countries generate cargo that is particularly well-suited for rail transport, namely coal and petroleum, iron, steel, vehicles, and so on. This cargo is very heavy, so it cannot travel by air and is even challenging to transport by road. Such cargo is also concentrated, both in origin (mines and factories) and in destination (mostly industrial areas and ports for exports). At the same time, the rail network in these countries is dense enough and it was originally developed to reach this kind of extractive and heavy industry sites.

Finally, long distances work in favor of railways in these countries. Despite the large territory, national markets evolved during the 19th and 20th centuries, largely thanks to railroads. Consequently, it is common to transport goods for very long distances, even coast to coast. Railways are very competitive for this kind of cargo and it could hardly be shifted to road.

As Pittman (Chapter 15) describes in this book, the decline of the US railway from the 1920s onwards, and particularly between the 1950s and the 1970s, did not so much affect the volume of freight hauled, but the revenue such transport generated. Rigidity in regulation was an obstacle for the review of prices and even for the termination of uneconomic services. However, railways were always popular with shippers, even if it was not a good business for the carriers.
1.3.3 Railway Systems Focused on Passengers

The opposite evolution can be identified in Western Europe, as well as in other small and densely populated countries such as Japan and South Korea. The national railway systems declined due to competition from road and aviation, but passenger services coped better with competition, and it was freight services that took the worst hit.

During the 1950s the modal share in freight transportation was similar in the US and in Western Europe: around 60 percent. However, by the year 2000 rail had declined in Europe to 8 percent in terms of ton/km, while in the US it was stable at around 38 percent (Vassallo and Fagan 2007).

Shipment distance seems to be the most relevant difference between the US and Europe. While cargo travels around the continent in the US, cargo in Europe tends to travel inside each country, with distances being relatively small. Thus, rail could not compete with road in the transportation of freight for such short distances. Rail was only competitive with very specific cargo (coal, iron, etc.), which is not as prevalent in Europe as in the US.

Short distances, on the contrary, were a protection for rail passenger services, especially against aviation. In some countries, no domestic flights were even available due to the short distances. Some countries prohibited the provision of coach services in parallel with rail passenger lines (France and Germany) to protect railway services from competition. However, no protection would be provided against private vehicles. As roads improved significantly after the Second World War and people could afford private vehicles, a significant number of passengers migrated from trains to private vehicles.

High speed was the response in Western Europe, and in some countries in Asia, to competition in long-distance passenger services. Shorter travel times made railways more competitive against both aviation and road transportation. It has been said that high-speed services are particularly competitive for trips of around three hours, which accounts for distances of approximately 500–600 km. Private vehicles are more competitive for shorter distances, while aviation becomes increasingly competitive as distances become longer. Five to six hundred kilometers is a standard distance for the most popular high-speed train routes in Europe and Eastern Asia.

The first high-speed service was the Japanese Shinkansen from Tokyo to Osaka (397 km), which was inaugurated in 1964 (Kurosaki, Chapter 12 in this book; Leboeuf, Chapter 23 in this book). In 1981 France became the first European country to launch a high-speed service, between Paris and Lyon, and then from Paris to Marseille (661 km). High-speed services were later launched in Spain (with Madrid–Barcelona (505 km) being the most popular route) and from London to Paris and to Brussels. In Italy, the high-speed service between Rome and Milan (479 km) became the most popular route as it reduced aviation modal share to less than 20 percent against more than 80 percent for rail. Currently, 21 countries have high-speed services, totaling 46,483 km of lines (Leboeuf, Chapter 23 in this book).

High-speed services have allowed railways to face competition from other transport modes when it comes to passenger services. Ridership and modal share have increased, particularly against aviation. However, high-speed services are very expensive, as it is very expensive to deploy the necessary infrastructure and to acquire the specialized rolling stock. High-speed is also expensive to maintain and operate. Costs are certainly higher than for conventional passenger services, which has limited the deployment of high-speed to heavily used routes in countries with the available resources.
Speed, which is a requirement for all passenger services, not only for high-speed services, has a further impact on the operation of the service. As priority has to be given to passenger services to ensure a minimum speed, freight services become problematic because they slow down traffic and will therefore be left for the paths that are not used by passenger services, often at night. They are also subject to delays and uncertainty due to incidents in the operation of passenger trains. High-speed exacerbates this negative effect, as high-speed services are essentially passenger services exploited over dedicated infrastructure over which no freight service can be operated (Leboeuf, Chapter 23 in this book). Consequently, the more passenger services improve, the more they negatively impact freight services.

For all these reasons, rail services in Western Europe, Japan, and Korea are increasingly specialized in the provision of passenger services. Freight services face increasing difficulties and have been limited to a marginal role.

1.3.4 Railway Systems with Mixed Use

Competition from other transport modes had diverging effects in the rest of the world. In the communist countries, planning policies supported railways against other transport modes, meaning that minimal impact was experienced up until the 1990s. The effect was not as relevant in the emerging countries as it was in developed countries, as alternatives developed slowly.

Thus, railway systems in Soviet countries suffered little competition from other transport modes. Planners did not support the use of private vehicles or the massive use of aviation. On the contrary, collective rail transport was considered the most efficient transport mode and a symbol of industrialization and development. This was the case of Russia and China, which developed some of the largest railway networks in the world.

Competition from other transport modes was also limited in emerging countries. On one hand, the rail network was not very dense and it was deployed on the routes where demand was strong. On the other hand, alternative transport modes evolved slowly, as resources to invest in road development were not available and both private vehicles and air travel were out of reach for the vast majority of the population.

Decolonization in emerging countries posed an even more relevant challenge. Countries such as India struggled, but ultimately managed to adapt the railway system to the new territorial reality after independence (Gangwar, Chapter 11 in this book). This was not the case in Africa, where railways mostly collapsed after the individual countries gained their independence. In Africa, rail had been built by the colonial governments, mostly for the exploitation of natural resources (mines and crops), and was not really designed to respond to the needs of the countries.

1.4 REFORM: DIVERGENCE IN RAILWAY POLICIES

Reform was necessary to confront the decline of the railway industry up to the 1990s. Passengers and cargo were rapidly shifting to alternative transport modes. The financial situation of the railway companies was deteriorating swiftly. Further public subsidies were necessary, as railway companies were not properly equipped to face the competition from the other transport modes.
However, policies to reform the railways diverged across the world. In North America, horizontal separation was promoted as a remedy: vertically integrated railway companies operating infrastructures and freight services would compete with each other without the traditional constraints of rate regulation. In the European Union, by contrast, vertical separation was promoted: infrastructures would be state-owned and operated as a monopoly at the national level, while competition would be introduced downstream among the providers of passenger and freight services. In other countries, reform was rather focused on the improvement in the operation by the state-owned monopolies.

1.4.1 Deregulation and Horizontal Separation in the US

Deregulation and the promotion of competition was the US remedy against the decline of the railway industry, as regulation was identified as the key obstacle for existing private railway companies to be able to adapt to the growing competition from alternative transport modes. Regulation prevented companies from abandoning unprofitable passenger services and from focusing on freight services, which had more commercial potential. Rate regulation was an obstacle to adapting pricing strategies to competition from the other transport modes. Deregulation would eliminate such rigidities. More competition would strengthen the industry.

The first criticisms of “entry and rate” regulation were issued by the so-called Chicago School. Applying microeconomics to existing regulation by independent agencies, the Chicago School posited that regulation was leading to inefficient results. The Chicago School was highly influential and triggered the deregulation policies that were subsequently implemented during the 1970s and the 1980s in the US. Legal entry and exit barriers were lifted, competition was promoted, and rate regulation was relaxed and even terminated. Air transportation, telecommunications, and electricity were deregulated.

Railways were identified as a leading example of the inefficiencies generated by regulation in the US (Kahn 1970). Railway companies had traditionally complained about the difficulties of competing with new transport modes (road and air transport) due to rigid rate regulation. They also complained about the exit regulation, which did not allow them to stop the provision of unprofitable long-distance passenger services. New economic literature was supporting these claims.

Therefore, deregulation was seen as the policy response to railways’ decline in the US. Railway deregulation in the US started in 1970 with the Rail Passenger Service Act, which allowed private railroad companies to abandon unprofitable long-distance passenger services. Such services would subsequently be provided by Amtrak, a state-owned company, making use of the infrastructure owned by the private railway companies at favorable access rates. However, this relief proved insufficient (Pittman, Chapter 15 in this book).

Rate regulation was particularly damaging in the 1970s, as high inflation was not followed by an immediate increase in rates. As a result, the financial position of the private rail companies deteriorated rapidly and a significant number of carriers filed for bankruptcy. The Staggers Act, adopted in 1980, deregulated railway freight services. Thus, the requirement for private railway companies to have rates pre-approved and publicized was thus terminated. Railway companies were allowed to abandon the operation of unprofitable sections of track and to stop services to unprofitable shippers.

The fundamental objective of deregulation was to strengthen railway companies by allowing them to compete, first with each other, and second with the other transport modes. This
policy can be labeled as “horizontal separation,” in the sense that each company would own and manage the railway infrastructure and would operate the rolling stock over such infrastructure to provide transport services to final customers.

A pre-requisite for this policy to succeed is the existence of a very dense railway infrastructure. Only high density ensures that several companies are in the position to connect two points via parallel infrastructure. Thus, this policy can only be implemented in a country where railways are mostly devoted to freight services, and hardly in a country where passenger services are prominent. On one hand, freight transport is not so dependent upon time; consequently, different routes – some faster and some slower – can be used to reach destinations. However, passengers are time-sensitive and demand to use the most direct route to arrive at their destination. On the other hand, infrastructures devoted only to freight transport have lower construction and maintenance costs, as they are less demanding in terms of speed and safety. Therefore, it is possible to maintain two or more competing infrastructures. On the contrary, infrastructure supporting passenger services, particularly high-speed infrastructure, is more expensive to build and to maintain. As a result, it will typically be uneconomical to operate two parallel infrastructures supporting passenger services.

Finally, competition between horizontally separated railway companies requires a strong demand for railway services, to ensure that there is sufficient demand to cover the high costs of several competing infrastructures. If demand is weak, it might not even cover the costs of one railway infrastructure, let alone more than one. This is the case of the US; where there is strong demand to transport heavy bulk cargo such as coal, steel, and so on, while objective geographic circumstances limit competition from on-shore maritime and inland waterways services, which, in the case of Europe, are good substitutes for the transportation of heavy bulk cargo.

Horizontal competition certainly delivered results in the US. Market forces triggered a deep transformation of the industry, which improved productivity and regained profitability (Pittman, Chapter 15 in this book). As a result, the decline in freight transportation was reversed. From 1980 to 2015, railways doubled the cargo transported in terms of tons/km and their modal share increased by 10 percentage points (US Department of Transportation n.d.). Certainly, rail companies have improved their financial situation over this period.

1.4.2 Liberalization and Vertical Separation in the European Union

The European Union implemented a multi-faceted strategy against railway decline. On one hand, major investments have been made to improve the infrastructure for high-speed passenger services – a success in terms of ridership – and to build a European network on top of the traditional national networks. On the other hand, competition is being introduced in order to increase efficiency, albeit only in the downstream layer, not at the infrastructure layer (as in the US), both for passengers and freight. Rail infrastructure is considered to be a natural monopoly and regulation is focused on controlling the conditions at which railway undertakings, while competing against each other, have access to infrastructure. Vertical separation between infrastructure managers and railway undertakings – also called “unbundling” – is considered to be the most appropriate tool in order to ensure non-discriminatory and fair access conditions.

Prior to the initiatives of the European Commission to reform rail, two Member States – Sweden and the UK – had already made significant progress in this regard. Sweden had unbundled its state railway company in 1988. As of October 2010, services were offered on
open access grounds (for a complete overview see Rosa and Finger 2012). In his doctoral thesis, Alexandersson (2010) argued that the Swedish reform was actually more “accidental” than deliberate and that Sweden’s railway reform was a fortunate outcome, rather than the result of a deliberate policy.

In the UK, the state-owned monopoly was unbundled into tens of companies back in 1994: (1) the infrastructure was privatized and sold to a private monopolist; (2) rolling stock was sold to a few competing companies (so-called ROSCOs), which then leased it back to railway undertakings; (3) freight services would be provided under competitive terms; and (4) more than 40 (now down to 21) franchises were tendered for the operation of passenger transport services by railway undertakings under exclusive rights. In other words, competition would take place during the tendering procedure, after which the winner enjoyed a monopoly.

This model was supposed to be temporary, limited to a transition period after which full competition would be introduced, although this has never materialized and the tendering of monopolies continues today. Still, reform in the UK has delivered obvious benefits: ridership has doubled and public subsidies have been reduced. However, costs have not been curbed (Nash and Smith, Chapter 4 in this book), but merely passed on to passengers thanks to an aggressive yield management system, granting steep discounts in order to attract passengers, while tripling prices for the passengers with little alternatives (peak times, tickets acquired just before traveling, etc.). This has been possible because England is a small and very densely populated island, with distances too short to fly and very congested roads.

At a larger scale, the efforts of the European Commission constituted the first systematic initiative to create a Single European Railway Area. Railway reform in the whole of the EU started in 1991 with Directive 91/440/EEC. This was a rather programmatic Directive that outlined the main dimensions of European rail liberalization and has not changed since. However, the execution of this program has taken 30 years and still has some shortcomings. Four packages of legislation have been necessary for the implementation of the Commission’s plan and total market opening – both in freight and in passenger transport and both cross-border and national – will only take place in December 2020.

First, infrastructures in Europe are considered to be natural monopolies; as such, they are mostly managed by state-owned companies. This is the most fundamental difference from the US approach, which promotes horizontal competition at the infrastructure level.

Second, a fundamental objective of the European Commission is the creation of a European rail network by harmonizing the different national infrastructures. Common technical standards (see Chavanel, Chapter 16 in this book) and a long-term, capital-intensive strategy to make all rail infrastructure interoperable are the tools with which to build this Single European Railway Area. The benefits seem obvious, particularly in terms of increased efficiency: lower costs for the development and the deployment of signaling technology, for rolling stock, the possibility to make freight transport more competitive, as it can be provided cross-border in longer routes, and so on. However, massive investments are necessary and national infrastructure managers have been reluctant to engage short-term expenses for a project that will only deliver long-term benefits. Projects such as ERTMS – a standardized European Rail Traffic Management System – have so far brought mixed results and national specifications continue to proliferate.

Third, vertical separation and access regulation are the main elements of European railway regulation (see Fetzer, Chapter 2 in this book). So far, two models of railway organization have emerged. Some Member States (such as Denmark, Portugal, Spain, Sweden, and the UK)
have fully unbundled the railway infrastructure from the transport operations, while others (such as Austria, Germany, Greece, Italy, Poland, France, and Switzerland) have implemented a holding model. However, there is currently no general consensus as to the optimal model for the organization of railways, and the current regulation still allows both the fully separated as well as the holding models. The argument from the supporters of a more integrated railway system is that vertical separation impedes overall system optimization, causes the duplication of fixed costs, reduces the potential for innovation at the systems level, and increases coordination costs overall.

The regulation of access by railway undertakings to the rail infrastructure is the cornerstone of the whole European regulatory construct. Access is mandatory for infrastructure managers, and access has to be provided under non-discriminatory conditions, so as not to discriminate against newcomers. Furthermore, access has to be provided under strict pricing conditions, charged at direct cost, plus a mark-up when the market can bear it. There are significant divergences in access charges across the Member States. Many countries have very low access charges, particularly for freight and suburban services. Charges tend to be higher for long-distance high-speed services and some specific freight services (heavy bulk cargo) with a low elasticity of demand. As the Italian example shows, competition can be favored when access charges are reduced, meaning that the underlying infrastructure becomes more subsidized.

Fourth, as is the case in the US, competition is considered the most effective instrument for increasing efficiency in railways operations. However, in Europe, competition is limited to the provision of transport services. In Europe, such competition has been possible for freight services since 2007, for international passenger services since 2010, and will be for passenger national services starting December of 2020.

In freight services, liberalization has allowed for market entry of a large number of private companies. Newcomers have increased their market share, to an average of 39 percent in 2016 (European Commission 2019). However, liberalization has not increased the modal share of rail in freight transport, which is stagnating below 20 percent, albeit with large differences across Member States. Countries with low international trade tend to have lower shares, which confirms that rail is particularly competitive for cross-border longer distances (European Commission 2019). The EU is far from the green goal of shifting a significant amount of freight currently transported by road to rail (shifting as much as 30 percent of road freight transported over distances greater than 300 km to other modes of transport by 2030 and more than 50 percent by 2050).

In passenger services, the impact of liberalization can only be measured in countries that have liberalized their services ahead of the EU mandated schedule, namely in Italy (described in Bert and Guglielmi, Chapter 5 in this book), the Czech Republic (Tomeš et al., Chapter 8 in this book), Austria, Sweden, and minor experiences in Germany (described in Ramos, Chapter 22 in this book). The conclusion is that competition seems to be limited to a low number of routes in each country (sometimes just one) with the best infrastructure (high-speed), and the highest volume of passengers. Competitors tend to be limited to the incumbent and a newcomer (Czech Republic being the exception, with three companies competing against each other). Competition benefits consumers in terms of price reductions (up to 50 percent), frequency increases, improvements in on-board services, ticketing, and so on. As a result, ridership has increased in routes with competition. More negative is the impact on the finances of the companies, both the incumbents and the newcomers, but results are mixed. Competition
also has negative effects in other routes, namely in terms of frequencies and termination of some services, as these routes can no longer benefit from cross-subsidies from the most profitable routes in the country.

Competition for the market is mandatory for those services that are financed by public subsidies under so-called public service obligations (PSOs). Starting in 2023 (with some transitory exemptions), Member States will be obliged to put the contracts for the provision of these services up for tender, as required in Regulation 1370/2007. Competition in the market is considered unfeasible, as these services are heavily subsidized, but competition can be introduced at least in the form of a mandatory tender. As a result, public authorities will no longer be allowed to directly assign the contracts to the incumbent state-owned companies. Some Member States (Germany, Sweden, and the Netherlands) have already experimented with such tenders. Costs tend to be reduced when tendering is introduced, but the right tendering scheme still has to be found (see Perennes, Chapter 18 in this book). Particular conditions exist when it comes to the tendering of mass-transit systems (Messulam, Chapter 17 in this book).

Fifth, as in the other network industries, a key element of European railway governance has been the creation of independent national regulatory authorities (NRAs) to enable and control the functioning of the market. NRAs have powers to control the access by the railway undertakings to the infrastructure and other facilities, such as maintenance facilities. Antitrust authorities have a parallel role to play, notably when it comes to the abuse of their dominant position by the incumbents and the collusion among competitors (Ganino, Chapter 19 in this book). With the creation of the European Railway Agency (ERA) in 2004, fostering interoperability across national networks has made a significant step forward.

The EU model of vertical separation and competition limited to the transport layer has inspired the reform of the railway industry in other countries. It has been exported to former Soviet countries (Poland, Hungary, Romania, etc.) as they have joined the EU, as well as to associated countries (Ukraine and Turkey, as explained in Ardıyok and Sesli, Chapter 14 in this book), in both cases with mixed results.

South Korea has implemented a duopoly in high-speed passenger services, with a second company operating services over the infrastructure managed by the state-owned incumbent company and in competition with the incumbent. No vertical separation has been implemented and the incumbent is actually a shareholder in the newcomer. However, competition is expected to discipline the behavior of the incumbent, to reduce costs and to increase efficiency (see Karlén, Chapter 10 in this book).

1.4.3 Reform in Other Countries

State-owned monopolies are still in place in large railway systems such as Russia, China, and most emerging countries. In these countries, limited competition has been introduced and reform has basically been limited to the restructuring of the state-owned monopolies.

In the former communist countries, railway companies have evolved into a state within the state, offering housing, education, health care, and even tourist services to their employees. They were often the largest and among the most powerful organizations in the country. The Fall of the Berlin Wall in 1989 challenged these companies, but the state-owned railway companies often continue to be among the largest employers in the country and citizens still benefit from numerous discounts (veterans, elder, students, etc.), which threatens the financial viability of these companies. Freight services are often competitive, as roads are of poor quality,
leading to freight services cross-subsidizing passenger services. On the other hand, investments in infrastructures and in rolling stock have been reduced significantly over the past 30 years, leading to incidents that affect the reliability of the services. Eastern European countries of the European Union have reformed their railway sector along EU policies by closing lines, laying off employees, and terminating services, as companies are urged to become financially sustainable. Furthermore, investments in roads and airports have reduced passengers and cargo patronage in railways (in Romania, cargo dropped 46 percent from 2007 to 2017).

In the non-EU countries, reform has mainly focused on the re-organization of the state-owned monopolist and on the review of rates. Reform has sometimes been undertaken in order to centralize the operation of the company, while on other occasions it has been undertaken to decentralize the organization. In some countries, the state-owned railway company has been transformed into a joint stock corporation, as is the case in Russian RZD (Saakyan, Chapter 13 in this book), Turkish TCDD (Ardiyok and Sesli, Chapter 14 in this book), and China (Tjia Yin-nor, Chapter 9 in this book).

Some competition has been introduced in India for container freight services (Gangwar, Chapter 11 in this book), for freight services in Turkey (Ardiyok and Sesli, Chapter 14 in this book), and for railcars in Russia (Saakyan, Chapter 13 in this book). The aim was always to attract private investment.

Financing of infrastructure construction and maintenance, as well as rolling stock, is a permanent challenge. Japan has been particularly innovative when it comes to financing. In times of rapid population increases, railway undertakings were able to finance the expansion of the network with the development of land around the station, benefiting from income from real-estate development and from the increase in ridership. However, as population has stagnated, more public involvement has been necessary (Kurosaki, Chapter 12 in this book).

1.5 REGULATORY DEBATES TODAY

In this section we will discuss the current challenges to railway regulation. Special attention will be paid to the latest debates in the US and the UK, both pioneering countries. However, after decades of reform, some negative effects have become evident in both countries and fine tuning appears to be necessary. The situation in the EU is also of interest, as 2020 is the year of full liberalization of passenger commercial services all across Europe. Finally, we will also highlight the latest challenges, which are decarbonization and digitalization.

1.5.1 Reevaluation of Reform in the US

In the US, the 1980 Staggers Act deregulated rail freight rates. For 20 years, rates were reduced, cargo volumes increased, productivity improved, and the financial situation of the railway companies was first stabilized and then improved.

However, horizontal competition was weakened by the reduction in Class I railroads from 39, when the Staggers Act was adopted, to 19 in 1990 and to seven in 2000 (Pittman, Chapter 15 in this book). This has actually been a common trend in most industries in the US, as antitrust enforcement of mergers has been reduced as part of the same deregulation trend that had led to the adoption of the Staggers Act. Further mergers were discouraged in 2001 by the
Surface Transportation Board (SFB), the regulatory agency that had replaced the Interstate Commerce Commission (ICC).

Consolidation has reduced competition. Duopolies exist in most areas of the country, and some shippers now depend upon a single railway company. However, shippers can often rely on alternative transport modes, which poses a competitive constraint on the potential abuses by the railway companies. Thus, the regulatory debate in the US is focused on the situation of the so-called “captive” shippers; that is, shippers that are served by a single railway company and cannot benefit from alternative transport modes because waterways are not available and their cargo is too heavy to be transported by road. This is mostly the situation of coal and bulk goods shippers. Consequently, captive shippers have been questioning the existing regulation and calling for more rate regulation.

Different alternatives have been proposed, such as benchmarking thanks to sophisticated regression analyses to identify high rates, as well as specific rate ceilings for particularly captive cargo categories. Rate-of-return regulation for some specific assets has also been proposed, along with intra-modal competition whereby railway companies could compete by having access to the rail infrastructure following the EU example.

The access model already exists for passenger services. Amtrak provides passenger services that make use of the freight infrastructure owned by private companies. Amtrak pays very low track access charges (marginal cost) and, in most corridors, ticket sales do not even cover variable costs. However, passenger services are still essentially perceived as disruptors of freight services. The success of high-speed services in Europe and Asia has generated interest for similar services in California and Texas, even if the completion of these projects is highly uncertain.

1.5.2 Reevaluation of Reform in the UK

A reevaluation of reform is also underway in the UK (see Nash and Smith, Chapter 4 in this book). The introduction of competition for the market, in the form of tendering of franchises, led to a clear increase in ridership and ticket revenue, as well as to a reduction of state subsidies.

However, the railway reform has failed to reduce costs, both at the infrastructure and the transport layer. In addition, the regulation of the rail infrastructure has been particularly challenging. Private ownership by shareholders of Railtrack was the weakest link in the scheme that was introduced back in 1994. The transformation of the infrastructure manager into a company that had no shareholders and was able to borrow from the private sector did not deliver either. The latest transformation into a government-owned company, controlled by the Department of Transport, is making borrowing more difficult, while providing no solution for the cost challenge. Different strategies have been used to curb costs, with minimal results so far.

At the transport layer, costs have not been reduced either. Tendering conditions were defined to maximize the exploitation of the exclusive rights by increasing revenue, not by reducing costs, particularly labor costs. Consequently, the Competition and Markets Authority is proposing to increase “open access” competition; that is, introducing competition in the market and by reducing the scope of the exclusive rights granted as a franchise, taking inspiration from the EU model. Overall, the franchise system might be too complex to ensure optimal coordination among all the players of an ever more complex system.
1.5.3 Completing Reform in the EU

The 30-year process of reform of railways in the European Union is set to be completed in December 2020. The market will be open for both commercial services (both passengers and freight) and passenger services subject to PSOs in 2023 (competition for the market).

Overall, the experience in the European countries that have already liberalized passenger services is mostly positive, even if competition has been very limited. Still, competition in the market has reduced prices and increased frequencies, which has led to a sharp increase in ridership.

However, competition has been limited in all EU countries in terms of the number of competitors. Also, the routes with competition are very limited. In Austria, Sweden, and initially the Czech Republic, competition was limited to one route, usually the route with the best infrastructure and the highest volume of passengers. In Italy, competition has been limited to high-speed routes connecting major cities. Even if competition was allowed in countries such as Germany, no significant market access has taken place. In most cases, only one newcomer has entered the market to compete with the incumbent.

Competition has also had some negative consequences. As could be expected, it has put an end to internal cross-subsidies from the most profitable routes (where competition has emerged) to the non-profitable ones, which had continued to be served by the incumbents. Consequently, some non-profitable services have been terminated or had their frequencies reduced. An unintended consequence of competition has been the reduction of access charges to make competition viable, as has been the case in Italy, where access charges were reduced by 35 percent in 2014. Low access charges seem to foster competition, while high access charges act as a deterrent. Of course, the costs of lower access charges have to be borne by taxpayers. Countries like France and Spain have recently expressed an interest in competition in high-speed rail.

Competition in freight services is a well-established reality in the European Union. There is a high number of active companies, newcomers have a large market share (more than 40 percent) in the main markets such as Germany, France, Switzerland, and Italy. Cross-border freight services account for more than half of the trips.

However, competition has not delivered the expected results in terms of an increase in competitiveness against road transport and in terms of modal shift more generally. Railways still only account for 17 percent of modal share in freight transport in the European Union, and this share is not growing. Cross-border services – those where railways are supposed to be more competitive – face interoperability challenges, capacity constraints, and poor coordination. Freight transport is unreliable, as priority in track allocation is given to domestic passenger services. Incentives are not aligned for infrastructure managers to make freight transport grow, as track access charges are lower than for passenger services. The policy of reducing access charges for freight service ends up working against freight transport. Major changes will be necessary in order to foster freight services on a European scale.

1.5.4 New Challenges: Decarbonization and Digitalization

Railway regulation will also be affected by two trends that are transforming the economy: decarbonization, and digitalization. Not only will modal share be changing, but a closer inte-
Decarbonization is perceived as a great opportunity for rail, albeit with some important challenges. Railways are clearly the transport mode with the lowest CO₂ emissions. The countries with the most ambitious plans to decarbonize their transport modes are planning a major modal shift from road and air to rail. Freight transport should migrate from road to rail so as to reduce emissions from trucks. Commuters in metropolitan areas should migrate from the private car to commuter railway services. Short-haul flights should be substituted by rail services. These policies, which are actively promoted by the European Commission, will increase ridership in passenger services and cargo in freight services, even if this evolution will reduce coal and other fossil fuels as the most popular cargo in the rail systems.

Digitalization offers obvious opportunities for railways to increase efficiency (Montero, Chapter 24 in this book): costs can be reduced in the construction and the maintenance of the infrastructure and of the rolling stock, mainly thanks to sensors and predictive maintenance. The operation of the system can also be improved thanks to digital communications technologies and automation.

The full exploitation of the potential of digitalization requires not only the digitalization of the operation of each actor in the system, but also digitalized coordination among all these actors. Railway networks are highly fragmented systems and reform has introduced further fragmentation because of the horizontal and vertical separation of the activities. Still, the right data governance model has yet to be identified and common data standards are overdue.

Digitalization offers obvious opportunities, but also poses challenges, as it could disrupt railways, just as it is disrupting all industries. Furthermore, digitalization is also empowering the other transport modes. The example of car-pooling platform BlaBlaCar in Europe (Montero, Chapter 24 in this book) proves that digitalization can increase the competitiveness of road transport: thanks to digitalization, road transport can combine the traditional flexibility with the network effects of a coordinated system. Automation can offer further efficiencies to road transport against railways, both for passengers (Leboeuf, Chapter 23 in this book) and freight.

Furthermore, mobility platforms have the potential to transform transport into a multi-sided market, with platforms coordinating all transport modes for the benefit of shippers and passengers. As a result, railway companies might well be “platformed” in the future; that is, transformed into commodified service providers with no direct relationship with their customers (Montero and Finger 2017).

1.6 CONCLUSIONS

Railways evolved into national networks during the 19th century, becoming the prime providers of passenger and freight transport services. The more a nation industrialized, the denser the railway network became. Peak rail was reached at the beginning of the 20th century. New transport modes developed during the second half of the 20th century (road and air transport), forcing the decline of the national railway systems. As we have seen, decline took different forms, notably as a result of the different underlying objective circumstances in each national system.
In the Americas and Australia, railways evolved mostly into freight networks. Large distances and low population density made other transport modes more competitive for passenger services. The same factors made rail services competitive for freight transport, particularly for the transport of heavy bulk goods – coal, petrol, steel, agricultural commodities, and so on – which are prevalent in these countries.

In Western Europe, national railway systems evolved mostly into passenger networks. Distances between cities are often too short for aviation and high population density makes rail competitive against congested roads. High-speed trains also increased the competitiveness of rail passenger services. Small distances within each nation worked against the competitiveness of rail for freight transport, which, combined with more competition from water transport (both maritime and inland waterways) and smaller extractive industries than in the Americas, reduced the modal share of rail below 20 percent in freight transport.

Communist countries were the exception both to decline and specialization in freight or passenger services. Centralized planning and support for collective against individual transport limited the development of road transport. Railways were always the backbone of the freight and passenger transport systems, and this is still the case in countries such as Russia and Ukraine and, to a lesser extent, in China.

In less developed countries, no further railway deployment was undertaken, but alternative infrastructure such as roads and aviation were slow to take off. If the existing infrastructure was well designed, it often became congested. By contrast, in Africa, where railways served colonial interests (mines, etc.) rather than the local needs, rail infrastructure often fell into oblivion after decolonization.

Different underlying realities led to different regulatory reform strategies. Railway reform gained traction during the last decades of the 20th century. Deregulation was a pervasive policy that affected all industries, particularly the network industries. Railways were also subject to such regulatory reform, sometimes even leading the way. However, the reform of the railways industry took very different shapes across the world.

In the Americas, horizontal separation and competition were reinforced, as rate regulation was relaxed and as privately owned, vertically integrated (infrastructure and services) freight carriers competed with each other. Reform led to improvements in productivity, larger volumes of cargo, and a more solid financial situation for the carriers. Passenger services, which were limited in scope, were nationalized in the US. They would be provided over the infrastructure owned by the freight carriers.

In the European Union, vertical separation became the rule. Rail infrastructure was considered a natural monopoly, run by state-owned organizations, while competition was introduced in the service layer. Liberalization was led by the UK in 1994, implementing a “competition for the market” model for the provision of passenger services over the monopolized infrastructure. Italy, the Czech Republic, Sweden, and Austria experimented with competition in passenger services, with positive results, even if limited to only a small number of routes. However, the reform of the industry took more than 30 years and is still not fully accomplished, partly due to opposition from Member States and national railway undertakings. Freight transport has not improved its small modal share in Europe despite it being fully liberalized back in 2007.

Reform in the rest of the world, particularly in the large railway systems of Russia, China, and India, has not taken the form of the introduction of competition. Horizontal and vertical separation for competition has been considered, but not implemented. Reform has basically focused on better organization of the monopolistic provider, improvement of infrastructure,
and sometimes the introduction of competition at the fringe of the system, as is the case in Russia.

Looking to the future, there is an interesting debate in the US about whether consolidation has reduced competition making it necessary to return to a stricter rate regulation, particularly to protect “captive” shippers. In Europe, pioneers such as the UK are reconsidering why reform has increased ridership, but has not curbed costs, and in particular, what the best governance solution for the regulation of the monopolistic rail infrastructure is. The full passenger liberalization taking place in the EU in December 2020 has yet to demonstrate whether it can increase the efficiency in the whole network, and not only in a few cherry-picked highly profitable routes. After years of competition, freight services remain a challenge, particularly as decarbonization requires a major modal shift from road to rail. Interesting developments in other countries also deserve attention: massive investments into high-speed services in China, the high efficiency and innovative funding of rail projects in Japan, and the reform of the national railway companies in the largest rail countries (Russia, China, and India) offer valuable lessons for everyone.

Finally, new trends such as digitalization and decarbonization are transforming the balance between the different transport modes. Furthermore, these trends are forcing the different transport modes to converge for the most efficient provision of door-to-door services. Sector integration, along with the integration of rail transport into intermodal solutions, often under the coordination of digital platforms, requires new regulatory strategies (Montero 2019).

The following chapters will present the different national experiences of railway regulatory reform. Cross-cutting chapters will also discuss financing, technical interoperability, competition, digitalization, mass transit, and passenger rights.

REFERENCES


