THE INFRASTRUCTURE – SUPERSTRUCTURE DISTINCTION IN THE RAILROAD INDUSTRY AND THE REGULATION OF NATURAL MONOPOLIES*

Guy NUMA§

Do not quote without permission

Keywords: infrastructure, natural monopoly, regulation, railroads, superstructure.
J.E.L. Classification : B10 ; D42 ; L51 ; L92.

Abstract:

This paper retraces the origins of the distinction between the infrastructure, which is a monopoly, from the superstructure, which is subject to competition. Using the case of the railroad industry in France and the US from the perspectives of 19th century theorists, I examine how both economic theory and legislation deal with these subjects over the last thirty years.

I argue that the origins of this infrastructure-superstructure distinction principle emerged in the 19th century with legislation pertaining to transportation networks. This was particularly the case for the railroad industry due to pricing, competition and financing rationales. I analyze the writings of Dupuit, Walras and Hadley and show that they all agreed that the distinction principle had to be implemented for the inland waterways. In contrast, they drew upon different justifications to defend the monopoly at the infrastructural and superstructural levels for the railroad industry.

In the late 1970s, the concept of the natural monopoly was reformulated by contestable markets theorists with the concept of subadditivity of the cost function. Thus, from their perspective, a theoretical basis justifies the distinction of the infrastructure which remains a monopoly, and the superstructure which can be subject to competition. I show that this theory influenced how the European Commission deregulated the utilities networks since the early 1990s.

The first section of the article explores the origins of this distinction in the 19th century. The second section analyzes its re-emergence by the contestable markets theory since the late 1970s. Throughout, I highlight the interplay between the work of the theorists and legislation.

* This study has received the support of the French Railway Historical Society (AHICF). A previous version of this article has been presented at the 11th Summer Meeting of the Pôle d’Histoire et d’Analyse des Représentations Economiques (PHARE) in Paris, France, September 1-5 2008.
§ Université Paris Dauphine, Laboratoire d’Economie de Dauphine, pôle Stratégies et Dynamiques Financières (LEDa-SDFi), Place du maréchal de Lattre de Tassigny 75775 Paris Cedex 16 ; E-mail: guy.numa@dauphine.fr.
Introduction

Since the 19th century, an industry is considered as a natural monopoly when, due to economies of scale, it is more efficient to have one company serving all of the market instead of several ones which would serve only a part of it. But in network industries like railroads, energy or telecommunications, the monopolist abuses its power by selling at higher prices and producing lower output in comparison with prices and output in a competitive situation. From the 19th century, Jules Dupuit was perfectly aware of this situation when he argued that the “ways of communication whose construction and operation need tremendous expenses are necessarily monopolies, and the proprietor of a capital monopoly can draw a revenue from it that is superior to that of capitals submitted to competition” (1853a, p. 340). Thus, as Stigler reminded more than a century later, “the main methods of controlling economic activity alternative to the market are public regulation and ownership” (1982, p. 6). That is, the state can supervise a natural monopoly via either infrastructure ownership (by nationalization for instance) or regulation.

This paper contributes to research in the areas of Industrial Organization and the History of Economic Thought on natural monopoly consequences. While the origins of natural monopoly have been retraced by Mosca (2008), no specific research has been made to analyze the origins of the principle that distinguishes the infrastructure, which is a monopoly, and the superstructure, which is subject to competition, though this distinction was developed in conjunction with the concept of natural monopoly. This paper fills this gap in the literature by retracing the origins of this principle. Using the case of the railroad industry in France and the US from the perspectives of 19th century theorists, I examine how both economic theory and legislation deal with these subjects today. The focus on railroads in the 19th century lies in the fact that, throughout this period, the railroads were created and progressively surpassed inland waterways in terms of passenger and freight traffic. Moreover, at that time the theorists’ work on the distinction between the infrastructure and the superstructure specifically focused on the railroad industry in comparison with the inland waterways and roads. This distinction was specific to the transportation industries and did not concern other network industries such as energy or telegraphs. In the railroad industry, the infrastructure consists of land, architectural works, railroads, platforms, regulation stations and electric facilities. The superstructure consists of the trains and the commercial operation of the railroad activity.
The first section of the article explores the origins of this distinction in the 19th century. The second section analyzes its re-emergence by the contestable markets theory since the late 1970s. Throughout, I highlight the interplay between the work of the theorists and legislation.

1. The origins of the distinction in the 19th century

1.1. The rationales of the distinction: a threefold explanation

In French railroad legislation, the infrastructure-superstructure distinction draws upon three complementary rationales: pricing, competition and financing. The pricing is closely related to the competition rationale which first appears in 1835. The concession book imposed to the contractor of the line from Paris to Saint-Germain-en-Laye (and to all the future concessions) distinguishes in the 33rd clause a toll (which represents 2/3 of the total price) from a transportation fare (which represents 1/3 of the total price) (Ribeill, 1997, p. 32). These “split fares” were established with the anticipation that the toll would be paid by other competitors using the tracks of the main contractor upon the principle of free competition contained in clause 42. Tax system is another reason which corroborates the distinction of the toll from the transportation fare: an act of July 2nd 1838 creates a tax on each ticket, this tax has to cover the transportation fare only (Ribeill, 1997, p. 32).

The competition rationale states that “to reimburse the company for the construction works and expenses that it plans to do, the state allows the company to receive the tolls and the transportation fares” (ibid.), the latter are received “in case the company makes the transportation at its own expense and by its own means” (ibid.). Furthermore, the 41st clause specified that “any future construction of road, waterway, railroad, in the area where the railway is located, will not be considered as a reason for reimbursement by the company” (Ibid). Finally, the 42nd clause—called the “free travel” clause—was aimed at “preventing any monopolies by the law or fact from the companies” (Ribeill, 1997, p. 29). This clause enabled the state to allow the construction of parallel infrastructure akin to the first railroad, but this threat of potential competition has not been implemented. Thus, it clearly appears that in the 19th century the legislation of the railroads anticipated a form of intermodal competition (clause 41) and the two forms of intramodal competition, that is, the possibility to build parallel infrastructure (clause 41) along with the possibility for a competitor to use the

---

1 For a specific analysis of French railroad legislation and concession contracts in the 19th century, see Numa (2009).
railways used by the main contractor (clause 33). Unfortunately, this will to preserve competition at the superstructural level did not become a reality, as some theorists and managers of the railroads considered that there could have been security problems.\(^2\) One may assume they feared that regulation techniques at that time could not correctly regulate large traffic generated by competitors’ trains.

The financing rationale is at the source of the « Railroad Charter Act » of June 11th 1842, an original model of Public-Private Partnership (PPP). This act clearly distinguishes the infrastructure—whose the expenses are paid by the state—from the superstructure, which is in charge of the railroad companies. The infrastructure is mentioned in the 5th clause which states that “\textit{lands and facilities, excavation works, architectural works and stations will be paid by the state}”.\(^3\) The superstructure is referenced in the 6th clause as “\textit{the tracks […]}, the equipment and operation expenses, maintenance expenses and reparation of the railroad […] will be in charge of the companies whose operation of the railroad will be contracted out”.

But we shall notice that the tracks which are normally a part of the infrastructure were paid for by the railroads companies, a detail not given by Caron when he argues that “\textit{the state had to build the infrastructure, the [railroad] company [had to build] the superstructure}” (1997a, p. 217). However, this form of PPP was not carried out, as the state directly contracted with the companies for both construction and operation of the railroads, due to a legislative amendment initiated by Prosper Duvergier de Hauranne (Picard, 1884, p. 279-81).

Overall, only the first two rationales clearly convey the possibility of competition at the superstructural level. The financing rationale helps to distinguish more clearly the infrastructure from the superstructure and clarifies the responsibilities of the state and the companies. The engineer Alexis Legrand is responsible for integrating these three rationales into the French railroad legislation. In 1835, as director of the Ponts et Chaussées et des Mines, he is charged with writing the concession book and naturally was one of the three signatories of this document along with the Minister of Interior Adolphe Thiers and the concessionnaire Emile Péreire (Chemin de fer de Paris à Saint-Germain, 1835). Most projects pertaining to the ways of communication submitted to the Chambre des députés (the French equivalent of the House of Representatives at that time) have his imprints. Between 1839 and

\(^2\) See Bricka (1894, p. 397), Colson (1903, p. 136) and Picard (1918, p. 499). These three authors do not dwell on the nature of these security problems. We assume they considered that switching techniques were not elaborated enough to manage a heavy traffic generated by the circulation of competitive trains.

\(^3\) Initially, the central government was in charge of one third of the expenditures, the départements and cities had to cover the remaining two third. This financial participation of the departments and cities was removed three years later due to an act passed in July 1845.
1847, he is sous-secrétaire d’Etat aux Travaux Publics, a cabinet-level position for the Public Works specially created for him. After interminable debates among lawmakers, it is under his tenure that the national railroad network with Paris as its center is created. This project that he designed, named “Legrand’s star”, eventually served as basis for the 1842 « Railroad Charter Act ». As a result, the free market orientation contained in the legislation since 1835 was not initiated by an economist but by a pure engineer. This is an important fact, especially because counter to several fellow engineers trained at the Ecole des Ponts et Chaussées, Alexis Legrand was not affiliated with a learned society and he was not a member of the French Liberal School. Legrand did not publish books or articles about the economy of transportation although it was his area of competence.

1.2. The analysis of theorists

How did the 19th century theorists deal with the distinction principle after identifying a natural monopoly situation?

Several French, British and American authors identified natural monopolies, especially in the network industries. But few of them raised the question of the distinction between the infrastructure and the superstructure. Three authors differ from the others: Dupuit, Walras and Hadley. For the roads and the inland waterways, Dupuit (1853a, 1853b) argues that the infrastructure should be owned by one operator only while there should be a free competition at the superstructural level between waterways companies. He is more ambiguous regarding the railroads, as he advocates an unlimited competition at both infrastructural and superstructural levels and, at the same time, he believes that this industry is more effectively managed under a monopoly situation. Examining the debates of the Political Economy Society allows us to clarify his positions. While Walras (1875) and Hadley (1886) defend the same viewpoints as Dupuit for the roads and the inland waterways, they argue that the infrastructure and the superstructure cannot be subject to competition. On one hand, Walras advocates that the railroads are a double natural monopoly, on top of high fixed costs the railroad industry is a public utility. Therefore competition should not operate at both infrastructural and superstructural levels. On the other hand, using the example of large scale companies, Hadley posits that larger firms experience greater economies of scale and thus

---

4 This is the case of Mill (1848) and Marshall (1890). The former is considered as the creator of the concept of natural monopoly (West, 2008), the latter used the term « indivisible industries» to characterize a natural monopoly in gas, water and railroad industries among others (Marshall, 1890, p. 628).
exercise their market power more effectively. This is the case of railroad companies. Thus, these companies have to be state monopolies or placed under state supervision.

For all three of these authors, in the railroad industry the infrastructure – superstructure distinction is closely related to the natural monopoly because it turns out that economies of scale are significant, moreover railroads are considered as utilities. One should bear in mind that the French legislation of the inland waterways and canals allowed a free competition between companies (in exchange of the payment of navigation taxes removed in February 1880) counter to the case of the railroads for which the operation is made via long term concessions (99 years).5

1.2.1. Dupuit’s ambiguity

Dupuit is one of the first authors to identify a natural monopoly (although he does not use this term). He analyses it in two articles published in the second volume of the Dictionnaire de l’économie politique, the entry “Tolls” and entry “Ways of communication”.6 From this perspective, what we currently refer to as economies of scale constitute barriers of entry to the railroad industry for new entrants. Thus, the incumbent company has a monopoly power which enables it to sell less quantity and at higher prices compared to a competition situation, which produces a rent:

“What I say about a railroad, I would be able to say the same for a canal, of some bridges; in fact their operation is always a monopoly, if not a legal one...First the enormous quantity of capital required to establish a new [infrastructure] limits the possibility to do so to a small number of people, then as the incumbent company is the only one within the market, the new one cannot compete with the first one, and the profit made by one is not enough for two” (1853, p. 340). For Dupuit, it is more efficient to have one company serving all the market instead of two serving only a part of it because “the new company would have significantly harmed the incumbent one [...] instead of a good deal, there would be two bad ones” (ibid.). A century later, the idea that economies of scale are barriers of entry is also used in the work of Bain (1956), Modigliani (1958) and Sylos-Labini (1962). However, these authors do not relate it with the natural monopoly. And if Dupuit acknowledges that other industries are characterized by high fixed costs,7 he underlines the importance of this type of costs in the

---

5 From 1823 to 1830, the four concessions granted were perpetual.
6 He also wrote the entries “Weights and measurements”, Ponts et Chaussées” and “Highways and roads”.
7 He cites the case of book publishing (1853, p. 342).
railroads. He states that “there is almost no industry in which this phenomenon does not exist, but maybe nowhere else it happens more remarkably than in the railroads” (1853, p. 342).

To distinguish the infrastructure and the superstructure, Dupuit mentions that “a way of communication [...] is made of the route, the vehicle and the engine” (1853, p. 847). Even though he does not use the words “infrastructure” and “superstructure”, the route (that is to say the railroad) is for him an element of the infrastructure while the superstructure consists of vehicle (the car) and the engine, which are indivisible because the engine makes the vehicle work. Then Dupuit considers that “for the canals [...] the building expense [...] is significant enough that the capital interest required is a notable proportion of the transportation price [paid by the client]. But the owner is not the operator; anyone can settle and use his boat; the competition works on the whole part of the transportation price [paid by the client] which corresponds to the transportation fare” (ibid.). This means that given the importance of infrastructure expenses in the inland waterways, there is only one owner, one should recognize a natural monopoly situation. Instead “the owner is not the operator”, competition can and have to operate between boatmen for the transportation activity, that is, the superstructure. In term of pricing, the consequence is the distinction between the toll and the transportation fare. For the railroads, Dupuit’s positions are much more ambiguous. He begins by praising competition virtues advancing that “there would be yet another way to take this industry [the railroads] away from state control: it would be to acknowledge [...] complete competition and freedom in regard to the building of communication routes. This is may be the true economic solution. Thus every company possessing the necessary capital to build a railroad would be allowed to do so where and how it chose [...]” (1853, p. 853). “Complete competition and freedom” means that for Dupuit competition must prevail both at infrastructural and superstructural levels. Furthermore, he corroborates this idea of generalized competition on both levels concluding that “unlimited competition, regarding the transport means, would not cause more drawbacks than it would cause elsewhere [...]” (1853, p. 854). He admits that “the parceling in a high number of operators significantly increases the required funds for the output, makes waste a lot of time and therefore multiply the expenses” (1853, p. 853), but this “drawback is much less serious than monopoly and perhaps even less so than state operation” (ibid.) especially because “it is present in all kinds of industry” (ibid.).

Dupuit anticipates the Austrian conception of competition in the way that his arguments strongly suggest that certain elements of monopoly were in fact natural to the competitive
process (Ekelund and Hébert, 1999, p. 166). Competition is envisioned by Dupuit as a dynamic process, thus monopoly is inevitably a temporary and transitory phenomenon. Therefore, monopoly is just a stage of the competitive process. In Dupuit’s mind, monopoly is not only a phenomenon natural to the competitive process but essential to the actualization of that process (Ekelund and Hébert, 1999, p. 167). Monopoly, even natural, cannot be used as a reason to restrict competition. Here, Dupuit posits that competition is the normal rule both at infrastructural and superstructural levels in the railroad industry. This is the reason why he thought that competition in railroad industry would be no more injurious or wasteful than that between butchers and bakers.

Next, in the same writings, Dupuit defends state intervention in the railroad industry. First he states that “on the railroads, the safety of operation requires that all the transportation expenses must be reunited and concentrated under a unique supervision. The route, the vehicle, the engine are the responsibilities of the same person, who has almost the whole monopoly of transportation. Thus the state had to intervene to repress its abuses setting up transportation price caps. It is obviously a necessity because competition is in fact removed and sometimes it is so on a legal basis” (1853, p. 851). He also declares that “communication routes whose construction and operation require high expenses are necessarily monopolies”. (1853, p. 340). After advancing that the state should intervene by regulating prices, next Dupuit goes further declaring that “every transport means that is a monopoly should be controlled by the state, and every transport means that is accessible to competition should be operated by private industry” (1853, p. 854). But this assertion does not clarify what should be done for railroads.

Concerning the natural monopoly and its consequences for the ways of communication, Dupuit therefore expresses ambiguous ideas in the two articles of the Dictionnaire he wrote. This ambiguity was not stressed by the contemporary and modern authors who studied his writings. We found two eloquent cases. Ekelund and Hébert (1999), in their study of Dupuit’s work on monopoly, extensively cite Dupuit’s writings from the Dictionnaire. However nowhere they explicitly conclude he was ambiguous which is so evident for the reader.

---

8 The Austrian School is far from being homogenous on competition and monopoly (Mosca, 2007). As example of arguments developed by this school, Dilorenzo (1996) argues that “competition is an ongoing process, and that market dominance was always necessarily temporary in the absence of monopoly-creating government” (p. 46). “When monopoly did appear, it was solely because of government intervention” (p. 48). In addition, the Austrian School underlines the importance of the information level of the agents on the market (Arena and Festré, 2002).

9 See in particular chapter 7 and chapter 12.
Béraud points out the opposite ideas used by Dupuit on the natural monopoly but surprisingly finds that on the same topic “his conclusion leaves no ambiguity” (2004, p. 21). Indeed in the same writings, Dupuit advocates an “unlimited competition” on both infrastructural and superstructural levels in the railroad industry and, at the same time, defends public interference in the same industry that can go till state operation. How can we explain this ambiguity?

It turns out it is related to the evolution of his thought in conjunction with his “status”. He is a pure engineer till 1849, an engineer-economist between 1850 and 1860, and finally a confirmed economist after 1860 (Etner, 1983). Dupuit was first an engineer that recommended state intervention in the railroad industry showing that a public monopoly is preferable to a private one (Dupuit, 1849). Let us recall that Corps des Ponts et Chaussées—which Dupuit belongs to—does not really approve the monopoly of railroads companies (Caron, 1997b). Then he can be described as an engineer-economist between 1850 and 1860. This period is somewhat a hybrid stage in Dupuit’s thought in which he is torn between state intervention and pragmatism based on facts that show that competition does not work in the railroad industry on one hand and his free market convictions on the other. We are particularly interested in this second stage of Dupuit’s life because before 1850 and after 1860 one shall notice that the concept of natural monopoly does not appear in his writings neither explicitly nor implicitly. Yet it turns out that the two entries “Tolls” and “Ways of communication” in which he uses the concept of natural monopoly are dated from 1853, a year which corresponds to the period we are interested in. At the end of his life, he became a free market economist who naturally defended free trade with no restrictions (Dupuit, 1861). Dupuit went so far in his free market logic that he was even criticized by other free market economists apparently quieted down by their age like Dunoyer (1861, p. 281).

After explaining Dupuit’s ambiguity on competition and natural monopoly, how can we clarify his positions?

1.2.1. The engineer-economist Dupuit and the Political Economy Society

10 He reiterates his position in the *Dictionnaire* at the entry “Tolls” (1853, p. 341).
11 In the *Dictionnaire*, Dupuit rightly reminds us that two companies operated between Paris and Versailles, one via the left bank and the other via the right bank, this attempt to introduce competition failed (1853, p. 853). Both companies ended up merging in 1855.
Dupuit’s positions can be clarified examining the internal debates of the Political Economy Society (PES), a learned society founded in 1842 regrouping J.-B. Say’s disciples. They regularly meet to discuss economic subjects more or less related to current events. Among the hot topics, there is the monopoly of railroad companies. This matter becomes critical in June 1852 when the Duke of Morny, half-brother of Emperor Napoleon III, announces before the Chamber of Representatives that “the railroad is itself a monopoly, its only limit is the concession book or its own interest: when it’s badly built, badly maintained, badly managed by a poor company, overwhelmed by overhead costs, then it becomes a true and dangerous monopoly, because it fights with its own misery, looks for profits in tariffs abuses and damaging public interest. Instead if the company is powerful, if it is widely recognized, it is able to freely attempt improvements, keep expanding, make sacrifices to go get travelers and merchandise further away[...].” (quoted by Carmona, 2005).

After the address of the Duke of Morny, the members of the PES decide to examine the status of the railroad companies. Thus the March 10th 1853 session is about the “railroads monopoly”. This session takes place during the same year that the second volume of the *Dictionnaire* appeared, which included the two articles authored by Dupuit. The debate opposes the proponents of the railroad companies’ monopoly to those who advocate free market in the railroad industry. In the first group, there are Horace Say, member of the chamber of commerce, the trader Alphonse Courtois and the economist Joseph Garnier who is a professor at the Ecole des Ponts et Chaussées. The second group includes a member of the Board of the railroad company of Orléans named Benat Saint-Marzy and Wolowski, the CEO of the Crédit Foncier. Jules Dupuit attends the session and at the time, he is an engineer working for the city of Paris. He joins the first group declaring that railroad companies “are both legal and de facto monopolies […] legal monopolies because the legislator decided so, excluding competition and setting up transport prices” (Courtois, 1889, p. 390), but also de facto monopolies. For these ones, he gives four justifications: first of all because once formed with enormous capitals “it is almost impossible to see rival firms […] emerge because such amounts are always difficult to have; second, it is likely that the firm company choose the more advantageous way; third, it had the time to take all the clientele and to create commercial habits; fourth, the founders would only have half of the advantages of the first,

---
12 F. Caron relies on this address to portray the Duke of Morny as a theorist of the natural monopoly (1997b, p. 208). Even if “he is one of the first who realized the interest of mergers in order to create powerful monopolies” (Rouart, 1995, p. 174), Caron seems to be wrong. There is no attempt of analytical construction of the concept (except the non justified assertion that « the railroad is itself a monopoly ») and no references to economists who used it.
and would risk superseding one good deal with two very bad ones” (Courtois, 1889, pp. 390-91). However “for most other companies, competition is easier” (ibid., p. 391). The fourth argument refers to the common justification of natural monopoly. During the session, the concept is used to describe either “a limited property, or restricted qualities and advantages, privileges [...] that comply with the nature of things or justified by necessity and utility” (Courtois, 1889, p. 392). He argues that examples of natural monopolies are « lands, mines, canals, railroads », because they “are not concerned with the same competition operating in the factories, as noticed by Mr. Dupuit” (Ibid.). Regarding the railroads specifically, Garnier concludes “it is impossible not to recognize the status of monopoly in the sense used by Mrs. Horace Say, Courtois and Dupuit” (Ibid.). Although he distinguishes the infrastructure and superstructure, Dupuit finally becomes a proponent of a unique monopoly at both levels in the railroad industry relying on the concept of natural monopoly. I assume that speaking in front of his peers forced him to clarify his viewpoints on the railroad industry. Another French author defends the railroads monopoly, Léon Walras. However, he uses different arguments.

1.2.2. Walras and the double natural monopoly

Walras (1875) dedicates a whole article to natural monopoly.13 As in indicated by the title, “The State and the Railroads”, he draws upon the railroad industry. The author distinguishes the “moral monopolies” and “economic monopolies”. The former designate “state monopolies related to services and products of public interest which are justified by law” (Ibid., p. 188), in other words to public utilities. The latter are “private monopolies turned into state monopolies or monopolies contracted out by the state” (Ibid., p. 189). For Walras, natural monopolies fall within these two categories. The adjective “natural” is used several times, however he does not give an accurate definition. In Walras’ thought, railroad is more than a commercial activity—it is also a public utility. It is useful for national Defense, for national unity and is essential agent of civilization and progress. Consequently, railroads must be nationalized because the owner of the routes is in monopoly situation which means it will manage at the detriment of the public, reducing consumer utility. Thus railroads are “natural and necessary” monopolies (Ibid., p. 208). Indeed the risk for society is to be dominated by railroads operated by private companies (the number of trains), a private monopoly produces less (it does not use its profit to open lines with uncertain return) and imposes higher prices

13 Walras may have been the earliest writer to employ the term “natural monopoly” in its modern sense (Ekelund and Hébert, 2003, p. 666).
than under competitive situation. State management and operation of railroads would have two benefits: lower prices and faster establishment of the network—things that the market cannot accomplish under competition.

An important fact to mention about Walras’ article is that, according to him, the railroads would constitute a double natural monopoly. In the last paragraph of the second part of the article entitled Railroads as public utilities and economic monopolies, Walras starts out distinguishing two types of transportation means. For the first one, he implicitly introduces an infrastructure – superstructure separation: at the infrastructural level, competition is not possible owing to wasteful duplication of infrastructure, and because railroads are public utilities. The superstructure which consists of the commercial part of the activity (the transportation itself), can be subject to competition. Roads and canals fit with this first type: “If the road and canal [...] are a natural monopoly, at least the transport that operates fits with the conditions of competition due to the fact that an indefinite number of cars and boatmen can drive on the road or ship on the canal. The toll [...] is paid to a monopolist; but the freight is paid to competitive entrepreneurs” (Ibid., p. 200).

For the second type of means of transport, Walras cites the railroads which, counter to the roads and canals, form a double natural monopoly. Thus for the railroads, “the route is a natural monopoly and the traction is another essentially linked to the former because [...] an indefinite number of companies cannot use the [same] tracks [...]. In this case, the rent of the track, cars and engines, the toll, and the freight are all paid to a monopolist. It is a real aberration to invoke freedom of industry in the railroads” (Ibid.). Walras organizes network industries into a hierarchy. In his mind, railroads distinguish from other industries in the sense that they constitute a double natural monopoly that requires state intervention because it is “the most powerful and dreadful monopoly” (Ibid.). In fact, our expression “double natural monopoly” designates a natural monopoly caused by a twofold explanation. It does not define a double explanation to economies of scale, both at infrastructural and superstructural levels. Walras just states that competition is impossible on both levels. By “double natural monopoly”, one should understand a double reason not to conceive competitive freedom, the first one is technological: it is explained by a ruinous competition (infrastructure duplication is inefficient). The second one is because the railroad industry is a public utility. This double natural monopoly explains why railroads must be placed under the supervision of one

---

14 Ekelund and Hébert’s translation into English is truncated and erroneous (2003, p. 666), it does not faithfully transcribe Walras’ words on the double natural monopoly of the railroads.
authority only, that of the state. There is only one natural monopoly but justified in two ways. Admittedly Walras is not the only theorist to defend state intervention in the railroads but the justifications he brought are original thanks to the double natural monopoly thesis.

1.2.3. Hadley: natural monopoly and large companies

Hadley’s contribution is of note for two reasons: the role played by the large size in the definition of natural monopoly and the relative similarity of positions with Walras. The American economic context of the late 19th century is marked by the emergence of large companies and concentration. Hadley relates the natural monopoly to large size though he does not provide any definition of what a large firm is. He stresses the key role played by transportation industries (especially the railroads) to characterize a natural monopoly. For Hadley, it is in these industries “that the monopoly of large concerns developed itself most sharply; it was here that legislators [… ] first came to recognize it as inevitable; it is here that there have been the most costly, but at the same time most fruitful, experiments in legislative control” (1886, p. 28). For Hadley, “the irregular and imperfect action of competition is most marked in the case of railroads” (p. 39). The natural monopoly is caused by the presence of high fixed costs spread over a greater output of units which produce scale economies. The bigger the firm is, the greater these types of economies will be, and the more the firm will expand in comparison with a smaller company that will not be able to do so; hence the more it fits with a natural monopoly. Hadley clearly relates his conception of natural monopoly to the large corporation.

However let us remark that this conception of the firm is reductive because it implies that the costs of a firm are connected to the importance of fixed capital. But the modern theorists of the firm rightly point out that the level of effectiveness in managing capital and labor does affect the unit cost. As noticed by Chandler (1990), large modern companies appeared in the US in the early 1900s obtained lower production costs than their smaller size competitors. It was possible not really due to economies of scale but especially owing to “economies of speed”. For a given plant, this word means the speed of transformation of inputs to output, in other words the velocity of the “throughput”. These economies are enabled because of a good programming and standardizing of the production process taken from the organizational

---

15 In the railroads, the most emblematic merger is that of the Southern Pacific with the Central Pacific effective in 1885. Their collaboration started in 1870 (Stover, 1997).
capacities possessed by an “army” of competent executives.\footnote{16} That being said, for Chandler it is true that the “first movers”\footnote{17} that massively invest in physical and human capital to become modern large corporations obtained quasi-monopolistic positions because of their size.

At first sight, Hadley’s analysis is similar to Walras. Drawing upon the facts, Hadley considers that transportation history consists of two periods. In the first which matches the roads and waterways, “the highway and the carrier were sharply distinguished. The highway, whether road or river, would have constituted a natural monopoly, if it had been made the subject of private ownership” (1886, p. 29). The second period matches the railroad industry and telegraphs. For these two industries, “it was an obvious advantage, if not a necessity, to have the line and the business owned and controlled by the same head” (1886, pp. 30-31). In one hand, the canals and roads, the infrastructure is a natural monopoly which makes impossible any form of competition while the superstructure is subject to completion between operators, it is what Hadley call “the general law of transportation”. On the other hand, “the railroad was at once highway and carrier” (1886, p. 31).

Nevertheless, we cannot conclude to a perfect similarity between Hadley and Walras. Thus the justification of natural monopoly in Hadley’s article differs from Walras on the superstructural level. Nowhere Hadley invokes the railroads as public utilities; the only justification of state regulation provided is the absence or the impossibility of competition. Overall there is no trace of the walrassian thesis of double natural monopoly. Moreover, one important fact is that Walras was consistent (it is easy to say so because he does not talk about the natural monopoly elsewhere than his 1875 article) while Hadley was not. One year before the publication of his 1886 article, he published his first major book, Railroad transportation, in which he set a quite different tone. As Cross and Ekelund (1980) remind us, although Hadley was aware that railroads were characterized by decreasing costs due to the high component of fixed costs, he did not believe them to be natural monopolies.\footnote{18} Along with water, electric, and telephones, railroads are called “particular monopolies” (1885, p. 63). In Hadley’s view, regulation of railroads or other industry was unnecessary and disastrous for

\footnote{16} Another modern theorist of the firm, Williamson (1985), analyzing vertical integration causes, implicitly the topic of the natural monopoly asking why a large firm could not do the same or better than a group of small ones (p. 131), why industries are not organized in monopolies (p. 133) and why the output does not come from one large firm (Ibid.). For a critical analysis of Williamson arguments, see Gabrié and Jacquie (1994).
\footnote{17} Chandler uses this expression to characterize firms which were pioneers in taking market opportunities and invested in several levels so that they beat their competitors.
\footnote{18} But Cross and Ekelund do not mention his opposite ideas on natural monopoly and regulation expressed one year later in the 1886 article. In addition, they do not indicate that this article marked an evolution in his positions that he maintained in later editions of Railroad transportation.
society in the long term. He even defended the combination of railroads as a source of prices stability (1885, p. 76-77). This position was clearly the opposite of that exposed in the 1886 article previously analyzed. This complete U-turn in Hadley’s thought is problematic as we do not have an explanation. Did he change his mind after the merger between Central and Southern Pacific in 1885?

Therefore, in the 19th century, the theorists unanimously distinguish a monopolistic infrastructure from a superstructure subject to competition for the inland waterways. In fact, the French legislation of the inland waterways and canals preceded the economists. It allowed a free competition between companies counter to the case of the railroads for which the operation is made via long term concessions (99 years). Regarding the railroads, the authors who developed the concept of natural monopoly all converge to consider that both infrastructure and superstructure should be managed under monopoly. Dupuit sets himself apart by advocating an unlimited competition on both levels, infrastructural and superstructural, and at the same time, defending the necessity of a monopoly for the railroads. The intermediate position that consists of distinguishing a monopolistic infrastructure from a superstructure subject to competition is not endorsed by any economists in the 19th century. This proposition becomes theoretically justified in the 1970s.

2. The re-emergence of the infrastructure – superstructure distinction in modern theories.

Modern theories of regulation can be summarized in two points: the proposal of alternative state intervention modalities and the definition of the exact frontier of the natural monopoly (Cartelier, 2007). The contestable markets theorists contributed to this last concept.

2.1. The contribution of the contestability theorists

Public interference in the economic world expanded during the 1930s. The economic crisis which began in 1929 created distrust toward free markets ideas and state intervention became inevitable in industries like telecommunication, transportation and energy--industries described as natural monopolies. Several regulation agencies were created in the United States

---

19 From 1823 to 1830, the four concessions granted were perpetual.
and a revival of interest in regulation arose in the 1970s and the 1980s. This time, however, public intervention was met with distrust due to the poor performance of public monopolies. Economists contributed to the rehabilitation of the market through developing more satisfying regulation theories. The underlying idea is that a market failure is more costly for a country than a government failure. Consequently, government intervention cannot improve the social welfare (Baumol and Willig, 1981). The work of Baumol (1977, 1982), Baumol, Panzar and Willig (1982) are particularly known for their contribution in showing the virtues of potential competition or ex post competition through the concept of contestable markets. In case of abuse by a company, the threat comes from the market itself. The idea is to use the potential competition to force the monopolist to reduce its market power. Thus, contestability theory, in its pure form, asserts that potential competition is as effective as actual in controlling market performance (Gilbert, 1989, p. 107). Potential competitors are to a status comparable to that of actual competitors. The threat of competition then encourages costs and price minimization without the expensive duplication of infrastructure required by effective competition (Foreman-Peck, 1987, p. 106). Potential competition becomes like effective competition because it plays a perfect control of monopoly power (Baumol, Panzar and Willig, 1986). A market is perfectly contestable under two conditions: there would have no barriers to entry and no barriers to exit (with no sunk costs). Contestable markets are characterized by “hit and run” entry. If a firm in a market with no entry or exit barriers raises its prices above marginal cost (average cost for a monopolist) and begins to earn abnormal profits, potential rivals will enter the market to take advantage of these profits. When the incumbent firm responds by returning prices to levels consistent with normal profits the new firms will exit. In this manner even a single-firm market can show highly competitive behavior and has the same efficiency properties than a competitive market.

In their work, Baumol et al. have also shown that the natural monopoly has a wider perimeter than economies of scales. In the case of a single-product monopolist, economies of scale are a sufficient condition but not necessary. Indeed a single firm can produce at a lesser cost than several ones on a range of output in which there are diseconomies of scale. This proposition is illustrated below (figure 1): $A_c$ is the U-shape average cost curve for one firm. Economies of scale are said to exist at all outputs less than $q^0$ and diseconomies at all outputs greater than $q^0$. Let us consider a unique technology for a given product on the market and identical costs characteristics between competing firms. In order to examine the least-cost solution for
outputs beyond \( q^0 \), we introduce the minimum average cost function for two firms, \( AC_2 \) shown on figure 1. \( AC_2 \) is obtained by construction from \( AC \) in the following manner. For least-cost production, each firm must produce at the same output rate and thereby have the same marginal cost. Hence, for a given point on the \( AC \) curve, simply double the output rate to obtain a point on the \( AC_2 \) curve. At the minimum average cost point \( M \) on \( AC \), double \( q^0 \) to get \( 2q^0 \), which corresponds to the minimum point \( M' \) on \( AC_2 \). The intersection of \( AC \) and \( AC_2 \) at output \( q^* \) defines the range of subadditivity. For all outputs less than \( q^* \), a single firm yields least-cost production. Hence the cost function is subadditive for outputs less than \( q^* \).

Although diseconomies of scale obtain between \( q^0 \) and \( q^* \), it would be in society’s interest to have a single firm produce in that range.

Figure 1: **Natural monopoly and average cost curves**

![Cost vs. Quantity Graph with AC and AC2 Curves](image)

*Source: adapted from Viscusi, Vernon and Harrington, 2000, p. 341.*

In the case of multi-product monopolist, the economies of scale for each product are not a sufficient condition, because if economies of scope\(^21\) are small, it is likely to be more advantageous to split the output in several plants (Sharkey, 1982). Thus, if economies of scale are neither necessary nor sufficient to make the monopoly the best productive organization

\(^{21}\) Economies of scope operate in case of joint production. In this case, it is more efficient to have a single firm producing any given quantities of at least two goods instead of two firms producing each one of these goods. That is: \( C(X_1, X_2) < C(X_1, 0) + C(0, X_2) \) where \( C \) represents the cost function, \( X_1 \) and \( X_2 \) are the different goods.
with the lowest costs, then it is not a relevant criterion to characterize a natural monopoly situation (Baumol, Panzar and Willig, 1982, p. xvi). In fact Baumol et al. redefine the concept of natural monopoly. The monopoly is natural if the cost function is subadditive for all the possible levels of output—a situation which can occur even in case of diseconomies of scale. The subadditivity means that it is more costly for two firms or more to produce the quantity $Q$ of a given product than in comparison with a single firm, for the total output sold on the market. We have the quantities $Q^1 \ldots Q^k$ produced by different companies, $j = 1 \ldots k$ with $\sum_j Q^j = Q$. The cost function is subadditive if and only if:

$$C(Q) < \sum_j C(Q^j)$$

This means that the average cost of a firm serving all the market is smaller than the average cost of two firms or more that would serve only a fraction of it.

As a result, the contestable markets theorists argue that in network industries, which were mostly public monopolies in the early 1980s, the perimeter of the monopoly was bigger than the zone of economies of scale (Sharkey, 1982). It is then possible to distinguish the superstructure, which can be subject to competition, from the infrastructure for which the fixed costs are high and justify the presence of a single firm.

2.2. The legislative translation

These analyses clearly influenced the European Commission in its liberalization policy of the public utilities. Thus, in the railroads, the legislation 91/440 imposes the distinction of the infrastructure from the superstructure, the legislation 96/92 does the same for electricity, 98/30 for gas, 90/387, 90/388 and 96/19 for telecommunications imposes the same principle of vertical de-integration. The legislation 91/440 clearly uses the term "infrastructure", while the superstructure is described as "the operation of transportation services of railroad companies". Consequently, the Réseau Ferré de France (RFF) has been created to manage the infrastructure and the Société Nationale des Chemins de Fer (SNCF) became a simple railroad operator by an act of February 13th 1997. More recently, a report of the Cour des Comptes (Accounting Court) published in 2008 underlines the limits of this distinction. The report shows that the responsibilities of RFF and SNCF are not respected. In fact, the former subcontracts a large part of its responsibilities to the latter. According to an act of February 1997, the SNCF has a double mission of both infrastructure manager and railroad company.

22 According to the term 1, there are three different systems: the accounting distinction is compulsory, the institutional (no capitalistic link between entities) and organic (possibility of subsidiary) distinctions are optional.
The Accounting Court report sheds light on the difficulties of implementing the distinction principle in France given the presence of a historical operator that was a monopolist for years. More generally, infrastructure and superstructure are complementary activities for all network industries, the question remains as to whether this principle could be efficient under these circumstances (Barale, 2000, 2002).

**Conclusion**

One of the solutions proposed by the modern economists to regulate the natural monopoly is a vertical de-integration by distinguishing the infrastructure, which is a monopoly, from the superstructure, subject to competition. But already in the 19th century, the theorists identified natural monopoly situations. For the inland waterways, they limited it to the infrastructural level and the superstructure could be subject to competition. This was already the case in France since 1820, as legislation pertaining to canals allowed free entry and free operation while the infrastructure was held by public or private monopolies. On the contrary, economists defended different measures for the railroads. Well-known authors like Dupuit, Walras and Hadley identified natural monopoly situations. Drawing upon different arguments, they recommended the monopoly at both infrastructural and superstructure levels. In reality, the origins of the distinction principle go back to the 19th century with the railroad industry on pricing and competition rationales. This distinction has been made by the legislation of the railroads, in particular the act of June 11th 1842 made this distinction on financing rationale. The theoretical debates on the distinction of the infrastructure from the superstructure came after the legislation. Therefore the legislation preceded the economic theory, and the distinction principle is not a new idea.

From the late 1970s, the work of the contestability theorists redefined the perimeter of the natural monopoly. They draw upon the concept of subadditivity of the cost function. From this perspective, the monopoly is located at the infrastructural level and the superstructure can be subject to competition. This distinction has been used by the European legislation since the early 1990s and implemented in all the network industries, including railroads but also in gas, telecommunications and electricity. For instance, in France a distinction was made between the infrastructure manager, Réseau de Transport d’Electricité (RTE), and the distributor,
Electricité de France (EDF) in the electricity industry. This time, the economic theory preceded and influenced the legislation.

In the 19th century, I did not find any legislative or theoretical origin of the distinction principle in other network industries like telegraph, water or gas. This corroborates the idea that the railroad industry generated the highest number of writings and legislation on the infrastructure – superstructure distinction. The modern theorists justify the implementation of the distinction principle by economic efficiency. However, this justification is much more the product of “identifiable social and historical forces rather than of extra-societal economic precepts” (Dobbin, 1994, p. 27). In other words, the distinction principle is much more influenced by the dominant ideology at the time it emerges.

References


