CITY
UNIVERSITY OF LONDON
EST 1894

CCRP Working Paper Series

COMPETITION, ECONOMIC REGULATION AND AFFORDABILITY IN INFRASTRUCTURE INDUSTRIES: AN ECONOMIC HISTORY 1840-1980

Jon Stern
Centre for Competition and Regulatory Policy (CCRP)
City, University of London

CCRP Working Paper No 30
(November 2017)

COMPETITION, ECONOMIC REGULATION AND AFFORDABILITY IN INFRASTRUCTURE INDUSTRIES: AN ECONOMIC HISTORY 1840-1980

Jon Stern
CCRP, City, University of London

ABSTRACT

This paper discusses the role of affordability in the development of economic regulation from the emergence of railway regulation in the 1840-60 period. It also covers the regulatory responses to the development of network telecoms, electricity and the water and sewerage industries, primarily in the UK and US. These network infrastructure industries were initially established as competitive industries but by 1920 had almost everywhere become vertically and horizontally integrated monopolies. This period saw the emergence of legally established independent regulatory agencies at municipal, legal and national levels. Affordability was a major issue in these developments and continued to be so until into the 1950s and after. The paper ends in 1980 after which competition and efficiency dominated infrastructure regulation. Since 2005, there has been a sharp revival of interest in affordability in regulated industries and the paper concludes with some observations on the historical record and comparisons with recent years.

Keywords: Infrastructure industry regulation, affordability, economic history.

JEL Codes: N40, L51, L97.

1. Introduction[1]

The main infrastructure industries (railways, electricity, water and sewerage, telecommunications and now information communications technology) have transformed our economies and the quality of peoples' lives. Each of them was introduced and each developed widespread coverage and use over a relatively short period, around 30-50 years.

Once in place, these industries became essential parts of economic activity and of everyday life. Hence, for each of them, governments have taken an active interest in their roll-out and conditions of service – particularly on the prices charged to household and small businesses and on general affordability. These issues led to the introduction of economic regulation to protect the customers while promoting the development of infrastructure network industries. This has been the consistent pattern from railways through electricity, gas, water and sewage, telecommunications and ICT (information and communications technology).

The focus of the paper is on the economic history of infrastructure industries with unavoidable use physical networks – railways, telecommunications, electricity, natural gas and water. The paper sets out the development of these industries from railways in the 1840s, via telecoms and electricity from the 1890s up to the 1970s. The main focus is on the development of economic regulation and the role of affordability. The paper stops in the 1970s as it was in that and the following decade that the focus of regulation was strongly reoriented towards efficiency concerns with affordability relegated to a very much lesser concerns. Affordability has returned as a major regulatory concern since 2008, but that is another and different topic that I discuss elsewhere[2].

Economic history repeatedly shows the key role played by political economy concerns regarding affordability and fairness which have affected the concerns and policies of governments of all persuasions – and regulatory agencies - since the introduction of railway regulation in the UK and US after 1850.

A second theme of this paper is the role of competition and markets in infrastructure industries. With the exception of water, all of these infrastructure industries initially developed as competitive industries supplying businesses and well-off households. It was only as they became more generally used, with larger networks, that they became vertical and then horizontal monopolies. That arose in the 20th century, particularly the period 1920-60, as the commodities they provided became 'essential' goods for urban households and then universally. However, with the rise of computers and digitisation, it became not just possible but economically more efficient to unbundle the industries and to reintroduce competition and markets in non-network segments. That began in the 1970s with telecoms and the unbundling of AT&T in the US. This

1 I am grateful to Richard Cadman, Martin Cave, Nick Crafts, Amelia Fletcher, Martin Lodge, Geoffrey Myers, Michael Pollitt, Agustin Ros and many others for encouragement and assistance with this paper. However, the views expressed in it are solely my responsibility and should not be ascribed to others or to any organisation with which I am or have been associated.

2 See Stern (2017).

was followed by related reforms and privatisation in the UK and elsewhere, typically starting with fixed line telecoms before being followed in the network energy industries from the 1980s.

The developments listed above were a watershed. It is very doubtful whether the reintroduction of markets and competition in these industries could have been achieved in a pre-computer age. They also gave further impetus to taking efficiency – particularly dynamic efficiency – as the dominant purpose of regulation rather than affordability. That is a further reason why this paper does not pursue post-1980 developments

In what follows, Section 2 introduces the arrival of economic regulation via the mid-century railways. Section 3 discusses why economic regulation was needed for the railways and successor network infrastructure industries and section 4 discusses their economic impact on growth, productivity and living standards (including health and wider welfare considerations). Section 5 discusses the rise of independent regulation and the role of affordability concerns via the evolution of railways, their increased monopolisation and their regulation from 1850-1900 in the UK and US. Section 6 discusses regulation, affordability and telecoms from 1890-1970 and section 7 covers electricity over the same period. The paper ends with a summary and some remarks on the implications for infrastructure regulation with the revival of concern over affordability since 2007.

2. Infrastructure Industries: Initial Competition, Natural Monopoly and the Revival of Competition

In the 19th and early 20th centuries, governmental concerns over supply conditions and retail prices in network infrastructure industries led to the development of economic regulation of various types in the UK, the US, and more widely. This occurred first for railways and then for the other network infrastructure industries. Railways were the pathfinder, technological frontier infrastructure industry from 1830 until around 1890. They provided the only means of transport at more than 10 miles per hour. However, they progressively lost their role – and much of their market power – with the arrival and growth of petrol and diesel-powered lorries, cars and buses.

Telecommunications became the pathfinder, technological frontier infrastructure industry from around 1890 and (now with ICT) it remains so. Telephony was also the post-1890 regulation pathfinder, particularly in the US. (In the UK and most other OECD countries, the national telephone company was state-owned and managed – most often along with postal services.) It was the post-1970 developments in telephony/telecoms that also led to the post-1970 industry privatisations and modern economic regulation. Throughout the twentieth century, the regulatory activities and concerns of electricity, natural gas and other infrastructure industries have followed those of telecoms, typically with a 5-10 year lag.

For railways, independent regulatory agencies were established progressively from the 1860s in the UK and the US and in other richer countries with substantial private investment in railways and other infrastructure industries. The same happened for

electricity and town gas in the US from the early twentieth century and for US telecoms from the 1930s.

Outside the US (including the UK and most Continental European and other comparable countries), governments - national, regional or local - provided economic regulation of infrastructure industries by government agency from the 1880s to around 1980. It was governments that established, monitored and enforced investment programmes, quality standards and prices – the key economic regulation outputs for network infrastructure industries. This was mainly done directly by government departments but sometimes indirectly (e.g. via concession contracts).

Over the period 1930-80, the main infrastructure industries operated almost everywhere as vertically and horizontally integrated monopolies - national, regional or local, frequently state-owned, particularly after 1945. Hence, from the late 1940s to the 1980s, the main UK infrastructure industries (telecoms, electricity, natural gas, and railways) all operated as state-owned monopolies. In other countries, they were either state-owned (e.g. France and Italy) or heavily regulated and directed (e.g. the US and some of West Germany) in the private sector.

That era was followed after 1980 by unbundling and competition with telecoms, electricity and natural gas as the pathfinder exemplars. However, since 2005, we have seen more activist government policies and regulation both in network energy industries and ICT, focusing on affordability. The last is a topic discussed in another paper[3].

2.1 18th and 19th Century Railway Franchises

Railways were the first modern infrastructure industry and economic regulation as we understand it today began with them. In both the UK and the USA, rail franchises were awarded by legislatures and these franchise contracts specified the terms of operation.

Modern UK franchises started from the eighteenth century model used for canals which operated under 21-year renewable franchises originally devised for toll roads[4]. These set canal toll rates and imposed a dividend payout ceiling. However, 21 years was not long enough for a rail franchise, particularly as the train company had to provide rolling stock, signalling facilities and make very large investments in the track and associated facilities. Hence, both in the UK and the US, the initial railway franchises became permanent so long as the private operator who owned them was still actively operating in the railway industry[5].

---

3 See Stern (2017).

4 These UK franchises were established by enactment via a private act of parliament promoted by the sponsors of the facility.

5 For more information on this and other aspects of UK railway history see Stern (2003) and its references. The best single guide is Foster (1992). For nineteenth century US railway history, see Kanazawa and Noll (1994).

2.2 From Competition to Monopoly in Infrastructure Industries 1840-1930

The early years of the UK railway industry demonstrated intense competition. During the 1840s UK 'railway mania' period, companies established railways that competed over the most popular routes. In addition, railways competed with canal and road transport, something that became progressively more important after 1900. However, railways always had considerable local monopoly power and after 1850 there was considerable consolidation towards regional monopoly rail companies. This culminated in the 'grand amalgamation' of 1923 of the remaining British railway companies into four major regional monopoly companies.

US railway experience from 1840 onwards follows a similar pattern with companies gaining legislative approval for franchises and generating intra and inter-State competition. However, this on-rail competition had also largely disappeared by 1900 as less successful companies merged with others or closed.

It is worth noting that this pattern of initial competition in the market followed by consolidation and regional monopoly was to a considerable extent replicated in electricity and town gas. The late nineteenth century saw short-lived competing town gas supply and distribution companies in London and electricity franchises in the US. In all of these industries, economies of scale and classic network effects promoted merger and eventual monopoly[6].

Particularly in the UK, vertically integrated monopolies became progressively more important from 1890 onwards – sometimes privately owned (railways and some energy companies) sometimes municipally owned (many energy companies, water supply and sewerage) and sometimes owned by the national government (postal services and telecoms)[7].

For telephony, before 1912 there were many UK local telephone companies, public and private. However, in 1912, the Post Office took over the privately owned National Telephone Company creating a single state-owned entity, with only the municipally owned Kingston upon Hull telephone company as an exception[8]. The US also had competing phone companies from 1894 until into the 1920s when the Bell Systems monopoly became near complete and a single natural monopoly US telecommunications company was established, a position which lasted until 1982 [9].

2.3 Economists and Natural Monopoly

Before 1980, it was common to see discussions of infrastructure industries such as railways, electricity and telecommunications treating them as 'natural monopolies'. Hence, Alfred Kahn's classic 1970 two-volume survey of economic regulation had a

---

6 See Turvey (1995) and Kahn (1970).

7 In the 1950s, there were over a thousand local monopoly water undertakings, a number that fell to 198 by the early 1970s, when there were still over 1,300 sewage disposal authorities. See Vickers and Yarrow (1988) pp 389-90.

8 See http://home.bt.com/news/bt-life/history-of-bt/a-short-history-of-telecommunications-in-the-uk-11363870786446

9 See Vickers and Yarrow (1988), pp 202-3 and Wallsten (2005).

major chapter on natural monopolies and their role in telecommunications and natural gas transport. Natural monopoly was typically justified by appeals to economies of scale and scope at the whole-industry level.

In fact, Kahn was critical of classing infrastructure industries as natural monopolies, but he was unusual among mid-20th century regulatory specialists of the time. Kahn consistently sought to find ways of bringing in competition where possible, but the natural monopoly perspective was clearly powerful for him and the intellectual perspective against which was writing. Regulated vertically and horizontally integrated infrastructure industries were the policy norm from around 1900 to at least the 1970s leaving little or no scope for any type of competition.

Kahn was a major pioneer in arguing that infrastructure industries had natural monopoly *segments* (e.g. unavoidable use physical networks) but that they were not natural monopoly *industries*.

For electricity, as for post-1850s railways, consolidation and monopoly arose because the scale of electricity companies was small and their efficiency low. Also, competition drove down the quality of electricity supply. Network economies of scale and externalities along with the growth in the optimal scale of plant, rising investment costs further encouraged horizontal and vertical integration. For electricity, gas and telecoms, the rapid growth in usage from the 1920s led to a growing need to serve customers on demand which also promoted regional and national monopolies.

These pressures changed as computerisation allowed more sophisticated management and control systems and the potential efficiency from unbundling became greater after 1980. Hence, AT&T (the US monopoly telecom company) was mandatorily unbundled in 1982 and PURPA (the Public Utilities Regulatory Policy Act) which introduced competition in electricity generation was passed in 1978. These examples were followed in Australia, Chile, the UK and many other countries during and after the 1980s.

3. The Rise of Economic Regulation from 1840 - Why do we need Economic Regulation for Infrastructure Industries?

Economic regulation was a product of and for the railway era. There were calls for regulation of the railways from the 1840s as they became regional and national monopolies. In addition, the railways became the dominant supplier of freight transport services and an increasingly important passenger transport mode. However, as discussed later, the model of regulation developed for 19th century UK and US private sector railways became the basic regulatory model for telecoms, electricity, natural gas, water and other infrastructure industries.

Why was this? Monopoly exploitation of customers (particularly small businesses and farmers) was clearly one factor, but so was the role of very long-lived fixed networks. As railways became an increasingly 'essential' factor for economic success so the demand for affordability in railway tariffs grew.

There was much discussion of railway regulation and its purposes from 1840[10]. However, the classic statement in modern economic literature of the reasons why we have economic regulation of infrastructure industries comes from Levy and Spiller (1994). This was developed in the context of infrastructure investment in developing countries but provides a good perspective for discussing the rise of economic regulation in the UK and comparable countries in the century after 1840.

Levy and Spiller provide the following three reasons for the rise of economic regulation of infrastructure industries. These are as follows:

1) Regulated infrastructure industries are *highly capital intensive*, with long-lived and often *sunk* assets
    - This implies private investors are at risk of losing value of assets by governmental/regulatory open or hidden confiscation.

2) The industries typically have *considerable economies of scale* and often *economies of scope*
    - This is most obvious with 'unavoidable use' monopoly networks as in railways, electricity and water but may apply to whole industries in small countries.

3) The services supplied are (a) *consumed by and necessary to the welfare of all households*; and (b) *provide critically important inputs for all firms and industries*.

Most attention in recent years has been given to the first of these Levy and Spiller criteria. Large numbers of small customers facing a monopoly is always likely to generate a call for regulation – and it has done so in practice since railways in the mid nineteenth century. Preventing large, powerful infrastructure companies from exploiting their consumers remains a classic regulatory objective even where monopoly has been replaced by oligopoly, as in ICT markets and energy supply markets. In 19th century discussions about railways, this criterion was regularly related to an affordability variant of the third criterion to justify regulation of railway prices.

The second Levy and Spiller criterion has become very important in 20th century regulatory economics but was not discussed in the context of mid-19th century railway regulation. The development of these concepts in practical terms did not take place until after the development of marginalism in economics after the 1870s.

In recent economic discussions, much less attention has been given to the third of Levy and Spiller's criteria. As discussed in Sections 4, 6 and 7, the third criterion – goods sold to all households and essential for welfare plus critically important inputs to all businesses – became crucial from 1840 and has, if anything, become more important since then.

At the heart of these concerns is that these commodities (plus water and sewerage supply) are generally seen as essential to life, health and welfare by citizen consumers

[10] See Foster (1992) and Stern (2005) and the references cited there.

as well as politicians and governments of all parties. It is clear that these issues have since 1840 always been the spur for the introduction of economic regulation of infrastructure industries. It is this (and associated affordability issues) that more than anything else has driven the demand for the economic regulation of infrastructure industries.

In what follows, I first discuss the economic impact of the main network infrastructure industries on industry and on household living standards before setting out in more detail the rise and development of economic regulation for railways, telecoms, electricity and water up to the 1970s.

4. The Economic Impact of Infrastructure Industries from 1840-1970

This is a classic topic and one much studied by economists and economic historians.

There has long been an interest in quantifying the impact of railways on growth and some major economic history studies, including the classic 1964 Fogel study for the US and Crafts (2004) for the UK. These studies have been followed by lots of subsequent research on the impact on GDP and on productivity growth of electricity and of other technologies, including ICT. Indeed, when in the 1990s economists wanted to project how computers and (what became) ICT would develop and impact on economic growth, they turned to experience with railways and steam and, in particular to the experience of the development and impact of the electricity supply industry. [11]

Almost all of this work has focused on the impact of these technologies on productivity and GDP growth. This is also the ostensible main focus of Gordon's recent researches and his major 2016 book. However, one of the most interesting facets of Gordon's recent studies has been his explorations of the impact of these industries on the *lives and welfare of households and people*.

As is well-known, GDP measures output and expenditure – not welfare[12]. Gordon has brought together a lot of useful evidence of the impact of railways, electricity as well as water and sewerage on the quality of life of American citizens. The demand for regulation has been driven by the wider benefits from industries that, after an initial competitive surge, operated as monopolistic network-based industries providing 'necessity' goods to large numbers of urban and later rural businesses and households.

The resulting direct benefits and externalities from the revolutionary new network industries have always been perceived as very important for peoples' health and welfare as well as for economic development. Hence the introduction of economic regulation to prevent monopoly abuse and also to ensure 'affordability' became an increasingly important political economy driving force in the rich economies when these industries became widespread in their use.

---

11 See, for instance, David (1991) and Crafts (2013).

12 See Coyle (2014) for a full discussion.

This combination has consistently been the main focus for the introduction of price regulation of infrastructure industries since the first calls for railway regulation in the US and UK in the 1840s.

4.1 Impacts on Productivity and Output

The first and most obvious impact on economic welfare of the rise in the main infrastructure industries has been their impact on GDP and productivity growth rates. This has been much studied for the US, the UK and many other countries. The discussion below presents a summary of the current consensus.

Regarding the impact of steam and the railways on productivity and growth, Crafts (2004) suggests that steam (steam engines an steam ships as well as railways) added around 0.2% per year to UK labour productivity growth from 1830-50, 0.4% per year 1850-70 and 0.3% per year from 1870-1910[13].

For electricity supply, Crafts estimates that electricity added 0.4% per year to US labour productivity from 1899-1919 and just under 1% per year from 1919-29. The main boost to productivity growth from electricity in the US appears to have been completed by 1930.

In both railways and electricity, the increase in labour productivity was lagged by several decades from the initial inventions. Over 100 years separated James Watt's invention of the steam engine to the period when steam had its greatest impact on productivity and growth. For electricity, that period was considerably shorter – about 40 years.

The delay between innovation, widespread incorporation and impacts on growth has been much noticed and has had a major effect on the debates about the productivity impact of ICT, where few economists found any sizeable impact before 2000 – and sometimes after. A lot of this debate is in response to Solow's 1987 dictum "You can see computers everywhere except in the productivity statistics". Since we are still far from the end of the digital revolution, it is still extremely difficult to say what the ultimate impact of ICT on productivity and growth will be. Indeed, the problems of incorporating digital products into GDP has proved to be very difficult and has led to many questions about what is actually measured – or should be measured - in GDP statistics[14].

Solow's quote is from 10-15 years before broadband usage developed. However, Crafts (2014) is clear that productivity growth from ICT has already been significantly higher than it was from steam and railways or from electricity over a comparable period.

13 See Crafts (2014) Slide 7

14 See Coyle (2014), Bean (2015) and many others.

4.2 Wider Welfare Benefits from the Growth of Infrastructure Industries

The impact of infrastructure industries on GDP and productivity growth is very important, but for the purposes of this paper, I want to focus more on the wider welfare benefits to households and citizens.

Regarding railways, Crafts (2014) notes the wider economic benefits arising from their geographical expansion – in particular, the growth of passenger travel and commuting. Railways allowed the growth of short and long distance rail commuting and the expansion of major urban centres such as London, Manchester, New York, Chicago and Paris. In addition, the modern tourist industry was in very large part the creation of the railways – both short day/weekend excursions and longer duration holidays.

This phenomenon was very important and was well-recognised at the time e.g. in the exchange of poems between Wordsworth and George Heald about the impact of the newly arrived Kendal to Windermere railway in 1847. This railway opened up the Lake District to mass tourism. Only some of the benefits to the urban workers from their excursions will appear in GDP as lower costs. Many of the subjective benefits will not appear in GDP (e.g. the pleasure given to large numbers of urban workers from visiting, seeing and smelling Wordsworth's celebrated Lake District spring daffodils)[15]. The same applies to the health and well-being benefits from living in lower pollution areas and commuting to work by train.

Commuter railways also gave rise to major wider non-market consumer benefits from the growth of passenger railways, as well as the output and productivity gains from the growth of (primarily freight) railways discussed above. Commuter railway services began in the 1840s and expanded rapidly with underground services from 1859. This allowed a seven-fold expansion in London population after 1850, which resulted in major output effects from urban agglomeration. To this can be added the welfare benefits, firstly, from those able to move from London slums to rather less polluted inner or outer London suburbs while working in London; as well as, secondly, the increased leisure time from shorter journeys to work. The same phenomenon was happening to a lesser extent in other major UK cities[16].

The changes to economic welfare were even greater from the spread of electricity and also from the development of network-based water and sewerage systems. Gordon's 2016 book describes this for the US with some extremely telling statistics. The same story would very largely apply in the UK and other OECD countries, albeit some years later than the US.

The story can be largely summarised as the arrival of major new commodities, originally only affordable to (and purchased by) the most prosperous families which then rapidly spread to urban working class families as the necessary networks were rolled out. This happened quite quickly as major and profitable new business

---

15 See https://en.wikipedia.org/wiki/Kendal_and_Windermere_Railway for the fascinating exchange of poems between Wordsworth and George Heald, a contracting engineer.

16 See Ball and Sutherland (2001). See also Robinson (2011) http://www.bbc.co.uk/history/british/victorians/london_modern_babylon_01.shtml

opportunities were perceived from expanding supply. There was high demand for these income elastic goods where the costs of supply could fall substantially with mass roll-out. These commodities then increasingly became seen as 'essential' for well-being and, with falling prices, urban access became near-universal over a 20-30 year period. Rural areas lagged behind in coverage and it typically required explicit government expenditure or subsidy to extend it to all citizens. This is the story that Gordon (2016) sets out for the USA.

This story started with railways and passenger travel. However, it also applies well to electricity and to water and sewerage and also applies well to telephony (particularly the spread of mobile coverage), and to the roll-out of broadband and other digital services[17].

For the US, Gordon's recent economic history for the 1865-2000 period sets out in detail the main findings of the take-off of post-1870 impact infrastructure industries – electricity, water supply and sewerage and others – and how these impacted on household and family welfare.

A summary of Gordon's main findings is as follows:

4.2.1 Electricity[18]

(i) In 1900, 3% of US citizens had electric lighting, but this rose to 79% by 1940;

(ii) By 1940, electric lighting was ubiquitous in urban areas but 66% of farms and 81% of 'rural' farms still relied on kerosene for their lighting;

(iii) The price of lighting fell by 6% per year over the 1902-29 period (81% over the 27 year period), which greatly assisted rapid take-up as did the quality of electric light as shown by the very large falls in the cost per lumen over this period;

(iv) Electric lighting resulted in a much lower fire risk than candles, kerosene and town gas. It also greatly improved atmospheric quality – and the incidence of associated health problems;

(v) The use of standard electrical appliances rose very sharply after 1920. From near-zero in 1920, by 1940, 40% of US households had a refrigerator and/or washing machine and 80% had a radio. By 1940, 80% of US households also had an electric iron and a vacuum cleaner;

(vi) The spread of electricity was also crucial to the growth of central heating. From zero in 1900, 40% of US households had

[17] See Gordon (2016) for the US. For UK electricity see Hannah (1979) and Stern (2005) Section 4. For telecoms and ICT see Ofcom Strategic Reviews 2005 and 2015-6.

[18] Gordon (2016), pp 114-22. He brings together results from earlier studies by Lebergott, Nordhaus and Platt.

central heating by 1940 and 70% by 1970. (This was followed by the growth of domestic air conditioning in the mid-late 20th century);

(vii) In Chicago, by 1929, 80% of households had vacuum cleaners and electric irons, 37% had a washing machine and 36% a toaster.

This list demonstrates how the spread of electricity revolutionised domestic work and the home environment. Many of these changes (e.g. refrigeration) had clear and important health consequences. In addition, the spread of reliable, low cost electricity also revolutionised cultural opportunities by enabling the development of radio, gramophones and cinema.

The US led the way on this but other countries were not massively behind. For Britain, Hannah (1979) reports that, by 1939, "almost two-thirds of homes had electric service, but many of these had only a lighting service, some had only one plug socket and only a small minority had the full complement of sockets which was to become standard in a post-war [post-1945] house". Of the houses with some kind of electric service, 77% had irons, but fewer than 40% had vacuum cleaners. Interestingly, 27% had electric fires (a standard pre-central domestic heating method in Britain for cold days)[19].

The pervasive transformation of these changes on family and individual welfare helps explain why the price and quality of electricity has always been a crucial 'affordability' issue for governments and regulators, particularly in recessionary times.

4.2.2 Water and Sewerage[20]

It is well-known how the spread of clean, running water and, even more, effective sewerage have dramatically reduced morbidity and mortality rates, particularly for children under 5.

In the UK, John Snow and Joseph Bazalgette are famous for their role in ridding London of cholera in the 1850s. In 1853-4, over 10,000 Londoners died of cholera. In 1854, Snow, in pioneering epidemiological work, identified the contaminated water source in Soho; he then had the handle of the pump at the infected source removed, following which deaths fell sharply. Snow's epidemiological studies were crucial for the sewerage developments that followed. Bazalgette was appointed Chief Engineer to London in 1856. Following the 'great stink' of 1858, he designed and supervised the building of a complete new London sewer system which was completed in 1875. Unlike other UK cities or Berlin, there were no cholera outbreaks in London after 1875 and a lesser outbreak in 1866 was confined to an area not covered by

19 See Hannah op cit p.208.
20 See Gordon op cit, pp 122-25.

Bazalgette's new sewers – a system still providing the core part of London's sewage system today[21].

For the US, Gordon presents core statistical information on the rapid spread and the huge impact of the rise of network water and sanitation systems.

The key US statistics are as follows[22]:

(i) In 1900, only 15% of US households had indoor flush toilets. This rose to 60% by 1940 and 95% by 1960;

(ii) In 1900, only 33% of US households had running water as opposed to 70% in 1940 and 90% in 1960;

(iii) In 1870, there were 244 waterworks in the US but, by 1924, the number had risen to 9,850;

(iv) In 1890, only 1.5% of US houses had filtered, running water. In urban areas, 25% of houses had filtered running water by 1910, 42% by 1925 and 90% by 1940.

The impact of these changes on human welfare was enormous. In 1900, 37% of all US deaths were from infectious diseases; whereas, by 1955, fewer than 5% of US deaths were from these diseases, with most of the fall *before* the development of effective antibiotics. Cutler and Miller (2004) estimate that nearly one-half of the total US 1900-1936 mortality reduction, nearly two-thirds of the fall in child mortality and three-quarters of the fall in infant mortality arose from the spread of clean water filtration and chlorination[23].

There was also a huge impact on the everyday lives of people, especially women. In his 2012 NBER paper Gordon writes as follows:

> "But the biggest inconvenience was the lack of running water. Every drop of water for laundry, cooking and indoor chamber pots had to be hauled in by the housewife, and wastewater hauled out. The average North Carolina housewife in 1885 had to walk 148 miles per year while carrying 35 tons of water…*There was no more important event that liberated women than the invention of running water and indoor plumbing, which happened in urban America between 1890 and 1930."* [Gordon's emphasis][24]

---

21 Wikipedia and the BBC History website have good short biographies of John Snow and Joseph Bazalgette. For additional information about cholera in London, see http://www.choleraandthethames.co.uk/

22 See Gordon op cit pp 120-125 and 217-8.

23 In the early 1980s, I read a book on the health, mortality and development which claimed that water and sewerage engineers had saved more lives than all the doctors who had ever lived. Unfortunately, I have been unable to find the reference….

24 See Gordon (2012), p.7. The North Carolina material was taken by Gordon from Strasser (1982), p.86, which discussed the results of an 1886 survey by the North Carolina Farmers' Alliance.

In Britain, the development of the water sector before 1914 was predominantly local – and messy. Millward (2005) describes it as largely municipally based with 80% of UK water supply companies being municipally owned, albeit with the support of private and central government finance. However, retail charges were kept low so that British water supply companies consistently operated at a loss, squeezed between progressively tougher public health obligations and an inability to raise prices. Increasing local tax revenues helped fund the investment requirements to meet the public health obligations with total investment in water being the largest component of local infrastructure investment from 1850-80[25].

Millward summarises the position on water in 19th century Britain as a semi-commercialised halfway house between wholly publicly provided education and health on the one hand and more commercially oriented electricity, town gas and tramways on the other. This also seems to be the case in the US and in several west European countries[26]. Keeping water and sewerage prices low to household consumers was a priority throughout – and has consistently remained so since the 1970s.

4.3 Wider Economic Benefits and Economic Needs – Evidence from Household Income and Price Elasticities

In the previous section, I discussed how infrastructure products and services – from railways to ICT - started out as luxuries only purchased by the most prosperous households before, over a relatively short period, becoming seen as essential for a reasonable quality of life.

This perspective is reflected in consumer demand theory. For instance, Engel Curves divides goods into 'luxuries' whose share in consumption rises as real household incomes rise and 'necessities' whose share falls as real incomes rose. The former have income elasticities of demand greater than 1; the latter have income elasticities less than 1.

There has been a great deal of empirical investigation of income and price elasticities for infrastructure industry products, particularly for electricity and energy. Particularly useful are the explorations of how these elasticities have changed over time. In what follows, I discuss the evidence on household demand elasticities in recent years for a range of infrastructure industries as well as on historical trends in income.

4.3.1 Estimates of Current Household Income and Price Elasticities

For electricity, the consensus is that the relevant current household income elasticities are low – significantly less than 1 – and that short-run price elasticities are also low – typically under -0.5. Long-run income and price electricity elasticities have been estimated as significantly larger than short-run elasticities.

---

25 See Millward (2005), pp 41-54 for an excellent survey.

26 See Millward op cit p.50 and p.45.

For instance, the Espey and Espey meta-regression analysis of mainly US studies estimates the long-run residential income elasticity for as 0.92 at median income and the long-run price elasticity as -0.81. This compares to their estimates of median short-run income and price elasticities of 0.15 and -0.28 respectively[27]. (The difference between short and long-run is whether or not the stock of electrical appliances used is unchanged or adjusted to the changes in income and price.)

For the UK, Meier et al (2012) report a residential income elasticity for electricity of 0.17 at the sample mean but rising with income from a (2007) income of £7,200 per year. Estimates of electricity own-price price elasticities vary but are almost always less than -1.0 and typically lower (less than -0.5) for samples that exclude households without electric heating [28]. For UK water, Waddams and Clayton (2010) suggest a residential income elasticity of 0.3 to 0.4 and an own-price elasticity of -0.4 to 0.5, but point out that there is likely to be substantial variation around these estimates. For telecoms/ICT, Cadman and Dineen (2008) estimate (cross-section) income and own-price elasticities for the number of broadband subscribers per 100 inhabitants as 0.78 and -0.43 respectively across 28 countries[29].

These residential price and income elasticities show how sharp cost and price increases can significantly reduce consumer welfare, particularly for low income groups, at least for commodities where the share of household expenditure is relatively high. This has been a major issue in electricity and natural gas policy and regulation in the UK since 2008.

4.3.2 Household Energy Income and Price Elasticities since 1700

The most useful evidence for the need (and development) of infrastructure industry regulation comes from the history of income elasticities and how they have varied over time. Fouquet (2011) shows clearly how developments in energy service supplies and efficiency have met the growth of successive wants – heating and lighting in particular. The growth of demand and the fall in the cost of providing the specific services has meant that the final commodities supplied have developed from being luxuries consumed by a few into mass consumer goods purchased by almost all households. This is reflected by the trends in income and price elasticities for energy reported by Fouquet for the UK from 1800-2010.

For residential consumption, Fouquet focuses on heating and lighting. He reports that, since 1700, there has been a 220-fold increase in household consumption of energy services for heating and a 295,000-fold increase in lighting consumption. These increases were largely achieved by technical advances which greatly reduced the costs of supplying a unit of heat/light and produced affordable heating and lighting services[30].

---

27 See Espey and Espey (2004). They used data from 36 different studies.

28 See Pollitt (2010), Slide 19.

29 Cadman has in correspondence written that they were surprised by the relatively low price elasticity, which suggests that broadband access had by 2007 already become a 'must-have' for many households in richer countries. In the UK, this is corroborated by DECC statistics which show that electricity consumption for personal electronic devices (including computing) rose from 19% of total domestic electricity consumption to 32% in 2009.

30 See Fouquet (2014), p 9-10.

The pattern of adoption was that first the richest households adopted the new (and relatively expensive technology) such as gas and electricity. These replaced coal, wood or tallow candles. Then middle income families adopted the technologies as their prices fell and then lower income families adopted it as the prices continued to fall and their incomes rose.

The pattern outlined by Fouquet is reflected in the estimates of UK income and price elasticities for energy that he reports for 1800-2010. His estimates are for overlapping 50 year periods. His estimated income elasticities for domestic heating and lighting were around 2.0 in 1800 but rose to peak at around 2.3 for heating and 4.0 for lighting at around 1825 and 1875 respectively. They then declined to reach unity by around 1950. Fouquet suggests that the earlier peak for heating demand reflected that priority was given to heating and cooking over lighting in lower income households[31].

Fouquet's estimates for UK energy price elasticities show low initial values for each of the energy services (around -0.5) but increasing in absolute terms towards -1.5 for lighting in the 1870s and for domestic heating in the 1920s. However, by 1950, both of these price elasticities were around -0.5, the level at which they have remained since[32].

The issues discussed above provide the context for the discussion of the rise and development of economic regulation since the 1840s. That is the topic of the next section. The pressures for regulation (by regulatory agency or by public ownership) have typically emerged once consumption of the relevant service has become widespread and the service has come to be seen as an essential commodity.

5. The Rise of Economic Regulation in Infrastructure Industries after 1850

In this section, we first discuss what is normally covered by the term of 'economic regulation' for infrastructure industries; we then discuss briefly the economic reasons for its introduction and development, including political economy concerns.

5.1 What does economic regulation of infrastructure industries cover?

Modern economic infrastructure regulation (for networks or industries) typically includes most of the following features:

(i) Approval of proposed investment programmes;

(ii) Price or profit controls;

(iii) Quality obligations;

(iv) Efficiency targets;

---

31 See Fouquet (op cit) p. 19-20.

32 See Fouquet (op cit) p 20-22.

(v) Periodic realignment of costs and prices; and

(vi) Consumer service obligations.

Not all of these are present for all regulated infrastructure industries and several were developed primarily in the mid-late 20th century. The 19th century regulatory methods focused primarily on prices, investment and (in the UK) dividend payout limitations.

As set out in Section 2 above, economic regulation in the UK had its origins in the contractual terms included in the Parliamentary Acts which authorised the construction of canals in the 18th century. Canals and toll roads were constructed following an Act of Parliament which gave a 21 year franchise and set the toll rate that could be charged for users of the canal or road. It also imposed a dividend payout ceiling of 1%. The franchise could be extended by a renewal Act. Hence, investment and price regulation were combined in a simple, if crude, way[33].

The franchise contract model was insufficient for the railways given, firstly that railway companies owned and operated the trains as well as the track; and, secondly, the much bigger investment outlays and longer life of the investments. Hence, modern-style economic regulation developed for the railways in the US and the UK from the 1860s.

5.2 The Development of Economic Regulation Across Infrastructure Industries 1860-1970

Economic regulation of public utilities is now over 150 years old. Some of its concerns have remained constant throughout (e.g. prices and fairness for households and small businesses); while others have waxed and waned. In particular, since 1980, there has been an increasing focus on economic efficiency (productive, allocative and dynamic) and considerable hostility among many regulatory economists against its use for distributional concerns and 'fairness'[34].

In modern economic texts (e.g. Decker's 2015 text), maximising efficiency – productive, allocative and dynamic efficiency – is taken as the core objective of the economic regulation of infrastructure industries and utilities. Regulation focuses on this in monopoly industries or industries with monopoly features like infrastructure 'wires and pipes'. Incorporating externalities (including network externalities) is another classic objective of modern infrastructure economists' thinking[35]. Some (more purist) economists argue that economic regulation should not take explicit account of distributional issues – they argue that those issues should be left to tax and social security policy[36].

Efficiency aspects of infrastructure regulation have become a much greater concern for governments since the 1970s, but distributional issues in general and 'fairness' in particular have always loomed large. Thus, Decker admits that "… many, including

33 See Stern (2003), p 14-15.

34 See Stern (2016)

35 See, for instance, Decker (2015).

36 See Decker op cit, p. 34. This is also an argument made by more free-market oriented economists).

economists, would concede that, in practice, regulators, politicians and the courts do consider issues of fairness and distributive equity in applying regulatory policy"[37].

However, as will be discussed below, fairness and distributional issues have historically been at the centre of economic regulation. They may have been a relatively subsidiary objective - particularly for energy - over the 1985-2008 period, but that was because of special factors. The return of fairness, protection of vulnerable consumers and other distributional issues to the regulation agenda since 2008 reflects not just the post-Great Recession pressures but a return to the pre-1970 historical norm.

In what follows, I will discuss the 1840-1970 historical record for the main infrastructure industries.

5.3 Railways and the Rise of Regulation from 1840-1900

The role of 'fairness' and the requirement of governments that modern utilities offer their services as widely as possible is well demonstrated in the way that economic regulation was introduced and developed for the railways after 1840. This is mainly a UK-US story as Continental European countries gave a much larger role to the state and, in some cases like Belgium, to state ownership.

In other European countries, military capability and planning issues were the main focus for railway policy – including state ownership (viz. Germany where the railways were nationalised in 1870s)[38]. In this role, railways became the first modern 'critical infrastructure'. This factor first became apparent in the American Civil War of 1861-65. This is often called the first 'modern' war and railways played a major role. Their role in the planning and conduct of World War 1 was, if anything, even greater[39].

The 'critical infrastructure' issue is not discussed further in this paper but is a strong factor in the development of non-economic regulation for infrastructure industries. This is reflected in current concerns and regulatory initiatives over the security of ICT and other networks, including recent concern over the security of electricity transmission networks from digital attack[40].

5.3.1 UK Railway Regulation 1840-1956

UK policy on railway regulation was first seriously discussed in the early 1840s culminating in the 1844 Railways Act. In the US, some states sought railway regulation from the 1850s. In both cases, the driving force was widespread and (it was claimed) abusive price discrimination over the freight rates charged to businesses

---

37 Decker op cit. p.35.

38 See Millward (2005), Chapter 4.

39 See http://members.kos.net/sdgagnon/mil.html for a journalistic introduction. See also Millward, p.72.

40 See https://www.wired.com/2016/01/everything-we-know-about-ukraines-power-plant-hack/ and https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/

and (in the US) to farmers. However, in the UK, the terms and prices offered to passengers were a major feature of concern as early as 1844[41].

UK railways were built via authorising acts of parliament which operated like franchise contracts. However, the previous 21 year renewable (post-review) franchise arrangements were not suitable for railways and the franchises effectively became unlimited duration. That resulted in very considerable difficulties in finding procedures within the contracts that allowed for rate reviews. Given the continued increases in efficiency of railway technology, this resulted in a progressive divergence between costs and prices – both in general and on specific rates – and very high profits for railway companies and shareholders. In addition, there was an inability to impose common standards e.g. in accounting practices, on health and safety or on consumer standards. On freight tariffs, there was not even an obligation to publish individual rates.

The results of this were:

a) Increasing difficulties in achieving rate changes;

b) An inability to stop the padding of the asset base, or to exert any downward pressure on costs;

c) An inability to handle increased horizontal or vertical integration within railways or between railways and shipping or road transport; and

d) An inability to investigate or control undue discrimination in pricing.

From the 1844 Railways Act to 1914, various attempts were made to handle these issues with more explicit regulation by agency attempting, ultimately unsuccessfully, to address the gaps left by relying on contract law alone.

The problems grew during the second half of the nineteenth century. There was increasing agitation against the high level of railway rates and, particularly from freight users, about the plethora of rates. By 1887, Great Northern railways had 13 million separate freight rates and North Western had 20 million and there was widespread price discrimination. Rates were particularly high where railways had merged with (or had no competition from) shipping or canals. In addition, there was widespread growth of costs from excessive investment and padding of capital employed – not least because it was for consumers to demonstrate via contract procedures in the courts that rate rises were unjustified. So long as the 10% dividend payout was unbreached, companies were at perfect liberty to raise rates[42].

After 1850, the only real threat to railway companies' prices and profits was the threat of new entry from competing lines but this threat was low, particularly as the system matured. This was primarily because the incentives for excess track expansion investment meant that, in most cases, new entrants faced the threat of a rates war that

41 See Stern (2004) and Foster (1992) for UK railways. See Kanazawa and Noll (1994) for US state-level railway regulation. The material below draws heavily on these references.

42 For further details, see Foster op cit chapter 2.

they would almost certainly lose – particularly when it was so difficult to obtain price comparisons. Hence revenue sharing, price collusion and merger dominated over competition. Where competition did exist, (eg London to Liverpool from steam ships), freight rates were lower, but competition was the exception rather than the rule, particularly when the railway companies vertically integrated through the purchase of canal and dock companies.

UK businesses agitated for remedies against the monopoly abuses and in 1873 the Railways and Canal Commission was established. This, however, did little to resolve the problems because the 1873 was quite unable to resolve the regulatory problems because:

(a) there were still very limited obligations on companies over publishing rates and no common accountancy requirements or other informational obligations;

(b) the onus of proof remained on consumers to demonstrate why general or specific rate increases should not be allowed; and

(c) the commission (and its successors) only considered rates on individual lines not averages or indices of rates.

These continued problems for freight users led to further protests and lobbying, the result of which was that, after 1894, the railway companies had to demonstrate to the Commission why they should be able to raise rates. This looks much more like a modern regulatory regime – except that it still applied only to individual rates. Unfortunately, the impact was to ossify rates and to increase the incentives for collusion and merger of rail companies. Rates ossified because:

(i) it was difficult and expensive to prove the case for rate increases;

(ii) as modern road transport developed from 1900 (and particularly after 1920), own and cross-price price elasticities of demand for rail rose sharply so that railway companies were worried that revenues would not increase from rate increases, particularly on freight; and

(iii) temporary reductions in rates were effectively made impossible following a 1907 appeals case in the High Court[43].

In consequence, the 1894 act made it significantly harder to raise railway rates just at the point where railway operating costs started rising sharply rather than falling as they had done until the 1890s.

Although most of the affordability pressures on nineteenth century railway prices came from freight users, governments were clearly active on behalf of passenger customers over providing affordable fares. Drafts of the 1844 Railways Act were primarily intended to provide economic regulation of freight rates. However, it is now mainly remembered because it is the first modern example of collective action to enforce quality standards and an affordable universal service obligation (USO) to

[43] See Cain (1978), p.89.

citizens. Under the 1844 Act, railway companies were obliged to run a "Parliamentary train". This was a regularly running train with an average speed of at least 12 mph and with a ticket price for third class passengers of no more than 1 penny per mile. Operators were obliged to run this train on Sundays as well as weekdays and so the Act provided the basis for Sunday excursion trains.

The provision of passenger trains with mandatory low fares was significantly extended by the Cheap Trains Act of 1883. This law exempted all railway companies from Passenger Duty if they charged less than one penny per mile. This law significantly increased the number of cheap suburban services and inner city commuting in large metropolitan areas. Historians of London have pointed to its role in enabling working class people to move from heavily overcrowded inner city and dockland slums to 'railway suburbs' like Walthamstow and West Ham[44].

The need to reconcile private sector commercial viability of the railways with affordability via regulation ended after 1945 with the nationalisation of the railways. Since then, railway services, particularly passenger railway services, have been kept in existence by continuous government subsidy.

Curiously, the Transport Tribunal, the successor to the Railway Rates Tribunal and the Railways and Canals Commission, continued in existence after railway nationalisation and was only abolished in 1956. As such, the UK regulatory commission for railways must be, in its various guises, one of the longest-lived specialist regulatory agencies. Its final rate hearings in 1955-56 had 44 days of hearings and took over 1,000 pages of evidence[45].

5.3.2 US Railway Regulation 1850-1970[46]

Price discrimination against shippers was also at the centre of the US development of the economic regulation of railways. From the 1850s, Mid-Western grain farmers complained loudly about the high freight rates that they were charged on short-distance journeys to the nearest major transportation centre. This led, firstly, to the development of State railway regulators in the post-1865 period; and, later, to the Interstate Commerce Act of 1887.

As in the UK, railway companies were, from the 1850s, were operating under unlimited duration franchise agreements. They were very profitable in areas with little or no competition, not least because of technical progress continuing to reduce railway costs coupled with strong demand growth which was increasing capacity utilisation.

Rural communities in the Mid-Western 'Granger' states lobbied hard for price controls and anti-discrimination legislation, against which the railroad companies lobbied hard. The main success for the farmers was in Illinois which, in 1870, established the first State regulatory agency. This lasted until 1886 when the Supreme Court decided that Illinois could not regulate prices for any portion of an interstate

---

44 There is an extensive set of studies of this. For an introduction and set of references, see http://www.bbc.co.uk/history/british/victorians/london_modern_babylon_01.shtml

45 See Foster op cit, p. 59.

46 This section draws heavily on Kanazawa and Noll (1994), pp14-18.

shipment. After 1886, some states continued to regulate purely intrastate components but only when given Federal authority to do so.

Although the 1887 Interstate Commerce Act abolished most of the powers of the state regulators over railroads, it limited long-haul/short-haul differentials and established the ICC (Interstate Commerce Commission) to regulate rates. It attempted to force publicity about railroad rates and it made illegal rebates and undue discrimination. The ICC was required to ensure that railroad rates be "reasonable and just".

Initially, the ICC heard complaints against railroad companies and issued 'cease and desist' orders against unfair practices but did not initially have price control powers. However, it was given powers to set maximum prices under the 1906 Hepburn Act. This Act also gave the ICC powers to regulate passenger railway services. These powers remained in place until the relaxation of railway price controls developed in the 1970s.

Passenger railway services have been much less important in the US than in Europe, particularly as cars, coaches and airline services have developed. The main area where affordability and fairness issues remain is on commuter rail and subway services which are the responsibility of municipalities and/or states. Some of these, like the New York subway, are municipally owned. The economic (and price) regulation of these within-state commuter services will therefore be a municipal or state responsibility, usually with major political involvement and pressures for 'fairness'.

6. Telecommunications and its Regulation 1890-1970

Telecoms has been to the 20th and 21st century what railways were to the 19th. They have become crucial for economic success and, increasingly, for economic welfare as well as social participation. In consequence, telephony followed by digital communications (including the internet) has been the regulatory pathfinder as well as an increasingly important technological driver over the last 130 years.

The beginnings of telephony were modest and non-monopolistic. As discussed in Section 2, many countries (including both the UK and the US) originally had competing telecom companies, private and municipal. However, by 1920, national vertically and horizontally integrated companies were the norm. From that point on, implicit and then explicit regulation became increasingly important.

6.1 The Development of Telecom Usage in the 20th Century

Telephones were not seen as a household necessity until 1960 - or later outside the US. In 1932, the UK had 46 telephones per 1,000 population as against 143 per thousand in the US. Denmark and Sweden had respectively 93 and 98 per thousand population in 1932; however, France only had 30 per 1,000 with Italy and Spain at 12 per thousand[47].

[47] See Millward op cit Table 7.1, p. 100.

For the US, the FCC reports historical telephone penetration rates of 35% in 1920, which only exceeded 40% after 1940. In 1960, the penetration rate was 78% but it had risen to 90% by 1970.[48] For the UK, in 1970, only 35% of households had a landline telephone. This rose to 72% in 1980 and to over 90% by 1990[49].

The household necessity was not *having* a telephone but *having access* to a telephone. In the UK, this was largely met by the provision of payphones. There were located in the famous red telephone kiosks and supplemented by large numbers of payphones in hospitals, shops, public houses, railway stations etc. Rural areas had at least one payphone in each village. In the US, telephone access was heavily assisted by the use of party-lines.

Numbers are difficult to find for UK payphone history, but what is clear that the number of such phones has fallen sharply since 2000 with the explosive growth of mobile telephony. In 1999, BT had over 140,000 telephone kiosks; but, by 2013, the number had fallen to 67,000 and to 47,000 by 2016. Of the BT payphone kiosks left in operation, many are email and/or internet kiosks as well as (or instead of being) phone kiosks. The average number of calls per phone box in 2013 was around one call per day with many rural phone boxes being used less than once per month[50].

6.2 Telecoms Regulation and Affordability

In the US, affordability was a major issue for telecom regulation from early on. From 1913, the AT&T monopoly was given the responsibility of developing a country-wide network with "reasonable prices of service". The 1934 Federal Communications Act explicitly prescribed the need to make an affordable basic telephone service available to all citizens. It established a Universal Service Fund by which long-distance carriers could cross-subsidise low income households and high cost areas "to ensure that all people in the United States have access to rapid, efficient, nationwide communications service with sufficient facilities at realistic charges"[51]. This model continued and evolved and was reiterated in the 1996 Telecommunications Act. Indeed, in April 2016, the universal service obligation was extended to provide a lifeline subsidy to low income Americans to include internet access.

The US telecom affordability support model has survived the unbundling of AT&T in the 1980s and the subsequent growth in telecom competition. Before that break-up, it was the classic example of US retail price setting on the basis of a cost-of-service rate of return regulatory approach. The latter may have changed since the 1980s to a more competition-oriented and forward looking regulatory process, but the affordability support model has not only been retained but has even been expanded with the rise of ICT and broadband.

In the UK, as in most European countries, telecommunications was organised from the early 20th century as part of the General Post Office (GPO), which was itself a department of government headed by a government Minister. There were initially a

---

48 See FCC (2001) 'Trends in Telephone Service', Table 17.3.

49 See Statista 2016.

50 See http://www.bbc.co.uk/news/business-22861389 and http://www.bbc.co.uk/news/magazine-35570663

51 Federal Communications Act 1934.

number of competing private sector and municipal telecom companies but the GPO took over trunk services in 1896 and the rest of the system in 1912. Only the Kingston-upon-Hull municipal service survived into the 1920s – and which still exists as an independent company. The telecoms part of the GPO remained until 1969 when it was established as a nationalised industry. However, telecoms remained along with postal services within the GPO until 1981, when BT was separated out and then privatised in 1984.

Since the UK telecommunications supply industry was part of a government department and then a nationalised industry until the 1980s, there was no need (or desire) for independent economic regulation of its activities. For household service, affordability criteria were not only important but were dominant and very important relative to efficiency criteria.

The affordability targets were achieved via a standard pattern of cross-subsidies. In the UK, the US and Western Europe, up to and including the 1980s. The pattern was that call charges subsidised household line rental; international calls subsidised national calls; and trunk calls subsidised local calls. Among different groups, business rentals and call charges subsidised households and urban areas subsidised rural ones[52]. This pattern, which emerged in the US in the 1930s, was gradually unwound in the US and in the UK and many other countries over the 1985-2005 period.

7 Electricity and its Regulation 1890-1970.

Affordability issues emerged for electricity in the US by the late 1920s and there was a growing regulatory response which began around 1910. This developed slowly and with a focus on regulated monopolies until the re-emergence of competition in the late 1970s.

For the UK and most European countries, as set out above, the spread of electricity to household consumers was relatively limited pre-1940 with limited mass use other than for lighting. Fouquet's estimates of income elasticities for domestic heating and lighting are only consistently below unity after 1945[53]. Further, the development of electricity supply was primarily local with widespread municipal ownership in the UK and Europe. Thus, explicit (let alone independent) economic regulation of electricity (or gas) supply was little used in the UK over this period. Prices were controlled either in franchise contracts or by local government acting as owners.

In this section, we set out the UK electricity retail supply history in some detail and this is followed by a discussion of US experience. In the US, unlike the UK, independent economic regulation was explicitly developed from around 1910, primarily via state regulation including rate-of-return based price caps for consumers. For the UK, the establishment of explicit agency-based economic regulation for electricity (and natural gas) in the 1980s was primarily the end-point of a long process

52 See Millward op cit, pp 250-251 and Table 13.2.

53 See Fouquet (2011), pp 25-26. After falling back after 1945, income elasticities for heating rose above unity in the late 1960s and 1970s with the growth in central heating, most of which was gas or electricity powered.

of commercialisation that had started in the late 1960s and culminated in the privatisation of the electricity and natural gas supply industries[54].

7.1 Indirect Electricity Regulation in the UK 1890-1950

Electricity in the UK was first produced and marketed for lighting. It emerged as a (high cost and low reliability) competitor to gas for domestic lighting in the 1870s. Indeed, the first parliamentary act was the 1882 Electric *Lighting* Act and many of the early electricity companies were named electricity lighting companies. Electricity was initially used almost entirely for lighting – commercial and public lighting as well as some limited sales to richer households in dense conurbations. This was the position until after 1895 with first traction (mainly urban tramways) and then, after 1900, industrial power uses becoming the main growth areas. Only in 1909 did UK industrial power sales exceed lighting sales.[55]

Between 1895 and 1913, total annual UK electricity sales grew from 38 GWh in 1895 to 180 GWh in 1900, 645 GWh in 1905 and 1,975 GWh in 1913. Between 1895 and 1913, the sales per head of population rose from 1kWh to 42 kWh. The share of industrial use also rose sharply - from zero in 1895 to 18% in 1905 and 51% in 1913[56]. This growth was very largely based on bottom-up expansion of local, vertically integrated generation and distribution companies with little or no interconnection between municipalities (unlike Germany and the US where larger regional integrated supply areas developed earlier)[57].

The UK electricity supply industry was increasingly dominated by municipally owned enterprises. Many of these municipal enterprises also provided (electrically powered) tramways. By 1900, 72% of UK electricity enterprises were municipally owned and these were concentrated in the largest towns – 44 out of the largest 50 towns had a municipal electricity service[58]. These (like their privately owned counterparts) typically combined generation, distribution and retail supply, a vertically integrated local monopoly which, in some cases, was weakly interconnected with other areas.

This pattern of ownership arose out of the 1882 Electricity Lighting Act. This act allowed companies (private and municipal) to establish electricity companies and gave them the powers to break up streets to do so. *However, a major feature of the 1882 Act was that it allowed municipalities the right to buy out privately owned franchises after 21 years at the written-down value of the capital.* The original proposal was for a franchise period of 15 years but the House of Lords amended this to 21 years. (21 years was the standard late eighteenth century franchise length for toll roads and canals as well as that initially proposed by Morrison and Gladstone for

---

54 As this is a paper primarily about retail markets and, in particular, household retail markets, there is little discussion of gas – town gas or natural gas – where for the UK it is very similar to that of electricity, especially before 1980.

55 See Hannah (1979) Chapter 1 and Table A.1. Hannah's statistics and discussion for the period up to 1914 derive primarily from Byatt (1978), which puts into book form an unpublished D. Phil thesis of 1962.

56 See Byatt op cit, pp 98.

57 NESCo (the Newcastle-upon-Tyne Electric Supply Company) was the notable exception and we discuss this below.

58 See Millward (2005), table 5.1, p.78.

railways in 1844.) In 1888, a subsequent Act increased the franchise period to 42 years.

In general, the municipally owned companies were operated on a commercial basis. Electricity prices were set in the original franchises. These were rarely a binding constraint, given subsequent technical progress, cost reductions, improved load factors and rapidly growing demand. The municipal enterprises were generally profitable and both their costs and prices were comparable to those of the private companies[59].

Municipal companies do not seem to have pursued major cross-subsidisation of local consumers, although, for 1921-22, the average revenue per unit of electricity sold by municipal companies to lighting and domestic customers was around 16% lower than for privately owned companies while the average revenue per unit sold to industrial and other customers was about 20% higher[60]. Some of this difference may, however, be due to compositional effects and regional differences.

Municipal companies also had to obtain permission (originally from the Board of Trade, after 1919 from the Electricity Commissioners) for new loan issues. However, municipal companies could borrow at low interest rates e.g. from the Public Works Loan Board. In general, it seems that the municipally owned companies typically had a significantly lower cost of capital than the private companies before 1945 and that this was frequently reflected in their charging lower prices to consumers in the 1920s and 30s than privately owned electricity companies – particularly lower prices to domestic consumers, who were also ratepayer-voters.[61].

Affordability issues over residential prices were not a major concern pre-1945. The interest of the supply companies was to keep domestic consumer tariffs low so as to encourage wider and deeper electrification with greater demand than a single lighting circuit. Hence, by the 1930s, domestic demand was already helping smooth the total load curve over the day, particularly in the evenings[62]. In addition, with generation being fuelled virtually entirely by UK coal, there was no risk of external fossil fuel price shocks. Hence, although in some areas there may have been some risk of monopoly exploitation of domestic consumers, this was rare and the objective need for regulatory retail price protection before 1945 was minor. Electricity only became a large-scale mass consumption good in the UK after 1950.

Before 1939, low efficiency from weak capital and management incentives in municipal owned companies was perceived as a bigger risk than monopoly abuse of customers and it was much discussed in that period[63]. Post-1945, low static and dynamic efficiency became an increasingly important policy concern in the context of monopoly state-owned electricity companies, particularly (and increasingly) from the mid-late 1960s. This was, not least, because final retail prices were set by the

---

59 See Foreman-Peck and Millward, Chapters 5 and 6.

60 See Foreman-Peck and Millward, Table 6.2.

61 Hannah (1979) pp. 214-26 has a long discussion of these issues. From it he was able to make some general observations which I have tried to summarise above. However, there was also a great deal of heterogeneity.

62 See Hannah op cit, Figure 6.2, p.201.

63 Hannah op cit provides a good discussion on pp. 225-27.

government and internal trading prices were neither based on economic allocation criteria nor used for internal transactions within the vertically integrated state-owned incumbents.

7.2 Electricity Regulation in the USA 1900-1945

Retail electricity supply in the US began, as in the UK, with franchise contracts issued by municipalities. At an early stage, there had been some competition at municipal level, especially for town gas, but vertically integrated monopolies had become the norm by 1900. The nineteenth century franchises were written as 20-50 year duration long term contracts which included retail price ceilings specified in nominal dollars[64]. As in the UK, these price ceilings increasingly became non-binding because of the rapidly improving technologies.

These municipal franchise contracts were very difficult to modify without explicit regulatory involvement. However, between 1900 and 1910, ten states passed laws that authorised municipal governments to change the contract-specified rates. In general, these laws allowed cities unilaterally to change price rates at the end of the contract period. However, it quickly became clear that this was not a long-term viable solution. In other cases (e.g. Tennessee), the state authorised a city (Memphis) to regulate gas rates subject to a legally specified minimum[65].

This municipal energy regulation system was relatively chaotic both in legal and commercial terms. Hence, not surprisingly, it was increasingly replaced by state level regulation via a public utility commission. By 1910, there were 7 public utility commissions (PUCs) - Massachusetts being the first in 1887 - and another 29 states established such commissions between 1910 and 1920. From this system, the classic US regulatory model evolved under which retail utility prices corresponded to the minimum level needed to cover all reasonably incurred costs, including the achievement of a sufficient rate of return on all relevant activities of the vertically and horizontally integrated utility. However, the regulatory objectives also included the safeguarding of affordable household retail electricity rates as well as in promoting efficiency (an increasingly important objective since the 1960s).

In general, the US municipal and state level regulatory system as it developed in the 1920-40 period allowed the co-existence of private ownership with affordable household prices[66]. It also combined this with the ability of the electricity supply companies to invest to increase coverage and quality while earning a reasonable rate of return. The UK only achieved these goals with the 1980s regulatory reforms which accompanied the electricity, telecom and natural gas privatisations[67].

It is also worth noting that the US electricity regulatory system has also consistently provided explicit support for vulnerable customers. Many of the major process reforms (and the introduction of the Federal Energy Regulatory Commission – FERC)

---

64 See Troesken (2006) pp 260-262 for a good general discussion. Much of the material in this section derives from this paper.

65 See Troesken op cit p.262.

66 There seems to have been no US equivalent to the UK rights for municipal repurchase of local franchises.

67 See Newbery (1999), pp 21-23 for a clear exposition of these processes.

took place in the 1930s i.e. at least in part reflecting affordability pressures from the Great Depression at a time when electricity was becoming more of a necessity. The US has also been the pioneer of 'lifeline rates' by which households can purchase a relatively small amount of electricity at a subsidised price.

This pattern of industrial organisation and regulation developed but remained essentially intact until the late 1970s with the renewed interest in wholesale (generation) competition from PURPA onwards. However, affordability has remained a key concern, particularly at State level.

8. Summary and Conclusions

The key points of this paper are summarised below:

1) Network infrastructure industries, from railways to ICT, have had revolutionary impacts. They have had major positive impacts on industrial growth and productivity and, even more importantly, they have transformed peoples' lives. This typically took 30-50 years, but speeding up with newer industries.

2) Apart from water and sewerage, 19th and early 20th century network infrastructure industries almost always began as competitive in the relevant market. However, from railways to telephony and ICT, they became national or regional/municipal monopolies within about 20-30 years.

3) Industrial use was dominant for around the first 20 years for railways, telephony and electricity. That gave rise to issues of monopoly abuse and affordability of service for small companies and for farmers. This was the dominant concern for mid-19th century railways and drove the growth of railway regulation from 1850-80.

4) For households, network infrastructure industry products were originally all luxuries with high income and price elasticities, but as they became widespread in use, they increasingly became 'necessity' products with low income and price elasticities. This had largely taken place in the US by 1940 but was 10-20 years later in the UK and Continental Europe

5) Pre-1900, the standard UK method of procuring network industry investment and service was by franchise contract. For water and energy outside the US, these were typically municipal or concession franchise contracts. For railways in the UK, US and some other countries, there were private franchise contracts; elsewhere they were state-owned companies.

6) Independent economic regulation began in the UK and the US from 1860 as a way of monitoring and modifying railway franchise contracts in an orderly and legally sound way. 20th century energy sector and telephone regulation in the US also followed this model, but in the UK and elsewhere, state and municipal ownership (without independent regulation) dominated until around 1980. Hence, outside the US, economic regulation (particularly of prices and investment) was a government responsibility until the 1980s.

7) Economic regulation has always had affordability as a major (if not its main) priority. The word 'affordable' occurs in most regulatory legislation. The origins of railway regulation after 1850 were primarily to protect industry and agricultural freight users from exploitation by the railway companies and later to protect commuters and others.

8) USOs to ensure lower income households could benefit from the new industries began with the UK 1844 legislation 'Sunday excursion' train with mandatory, regulated fares of less than 1d per mile (and with minimum speed and quality provisions). This was extended in 1883 to provide regulated low fares on 'workmen's' daily commuting trains.

9) Affordability issues were most important in the water and waste water industry where competition never took hold for retail markets. The development of the industry had a huge impact on mortality rates from London in the 1850s onwards and also in the US with the growth of treated water and sewerage. The mortality improvements (particularly of infants and children) were crucial. However, the industry was run in both those countries on a semi-commercial, semi public service basis and regulated with a close watch on affordability.

10) The standard pre-1980 regulatory model was developed in the US between 1910 and 1940 at state and Federal level. For telecoms, there was state and Federal regulation – the FCC was established in 1934 and the Federal Power Commission (the predecessor of FERC) in 1935. The telecom market and regulatory model included a natural monopoly telephone company (AT&T), USOs, regulation of final prices, lifeline tariffs and pervasive cross-subsidies. US regulation of electricity and gas followed a similar pattern, albeit with lags.

11) Affordability considerations (and natural monopoly arguments) largely dominated infrastructure regulation from 1930-70 for network infrastructure industries in almost all countries. After 1970, efficiency concerns (static and dynamic) investment cost and resource allocation questions became increasingly more important in the US, the UK and many other OECD countries.

From the perspective of 2017, many of the affordability concerns that dominated economic regulation from 1850-1970 have returned since 2008 after a 35 year period (1980-2005) where they were relatively dormant.

The 19th century affordability focus of utility regulation was at least as much on small industry users and farmers as on households. However, this was because it was only in the late 19th century that railways (and, in urban areas, water and sewerage) became a widespread household sector commodity and well into the 20th century before that happened for telephony and electricity. However, since 1950, affordability concerns have been increasingly on households - particularly low income and disadvantaged households.

The revival of concern about affordability since 2008 has, particularly in the UK, been primarily focused on energy (electricity and natural gas), but also on commuter railways. Continuously falling costs and expanding services have made ICT much less sensitive on affordability grounds – at least in areas where internet and mobile coverage and service quality is good, which is generally the case outside some (often sparsely populated) rural areas.

For energy, the dominant issue has been the combination of sharply fluctuating fossil fuel prices which rose very sharply after 2005 – just at the time when the UK and other OECD countries had the largest recession since the 1930s. The rise in fossil fuel and retail energy prices coincided with the tailing off of the post-1980s privatisation induced productivity gains – and the Great Recession. This made for a 'perfect storm' which meant that energy affordability became a powerful concern for the British government and energy regulator[68].

In the UK, per capita real household disposable incomes were rising steadily from 1983 until 2005-06. They were then broadly flat until 2009 but fell by 2.1% between 2009 and 2011. They only reached their previous peak by 2015. However, for low income households, the picture is much worse with a reduction of 7.1% in mean annual household income between 2009-10 and 2011-12, an income fall markedly greater than that for median households (5.8%)[69]. These reductions in UK real household incomes were unprecedented over the last 50 years, and more than anything else created the affordability furore that, at least for energy and commuter rail has continued until now (2017).

Given the major squeeze on household incomes and the growing insecurity of low-paid jobs, fluctuations in prices – and retail price spikes in particular - posed major problems in household budgeting. The absence of pre-1980 energy price 'smoothing' via average or long-run marginal cost pricing created particular problems for hard-pressed households and was a major factor in the post-2008 energy price controversies[70].

The strong correlation between major recessions and controversy over the household retail prices of infrastructure industry thought of as essential for health and welfare is to be expected – and is far from new. In the US, it is no coincidence that utility regulation was first subject to Federal regulation in the 1930s during the Great Depression when major legislation was enacted for telephony, electricity and natural gas.

A major lesson from the economic history surveyed in this paper is that we should expect continued concerns about infrastructure industry affordability to persist at a high level at least while and until household real incomes return to sustained growth. This is particularly the case for household retail energy prices and their regulation, But, household disposable income per head has been flat since 2009 and more than

68 See Stern (2014) and (2017).

69 See Phillips, IFS, 2013. Data for low and median income family incomes comes from the annual survey of households with below average incomes (HBAI).

70 See Biggar (2010) for an excellent paper on regulatory fairness and consumers' sunk costs.

50% of UK households have seen no material growth in real disposable income since 2005[71] and this is not expected to change significantly in coming years..

The decline in the importance of affordability in infrastructure industry regulation was most marked in the two post-1945 'Golden Age' periods of strong growth in GDP and real incomes – 1950-73 and 1985-2008. However, we now seem to be in a 'Leaden Age' so that, if the economic history of economic regulation provided a useful guide, affordability concerns and household retail prices are likely to remain a major concern for some years to come.

[71] See Haldane (2016), p.6.

REFERENCES

Ball, M. and Sutherland, D., (2001), *An Economic History of London, 1800-1914*, Routledge.

Bean, C., (2015), *Independent Review of UK Economic Statistics*, HMSO.

Biggar, D., (2010), 'Fairness in Public-utility Regulation: A Theory', *Agenda- A Journal of Policy Analysis and Reform*, Vol. 17 No.1.

Byatt, I.C.R (1979), *The British Electrical Industry 1875-1914*. Oxford. The Clarendon Press.

Cadman, R. and Dineen, C. (2008), 'Price and Income Elasticities of Demand for Broadband Subscriptions: A Cross-Sectional Model of OECD Countries'. *SPC Network.*

Cain, P.J. (1978), 'The British Railway Rates Problem'. *Business History*, Vol. 20, No. 1, 87-99.

Coyle, D. (2014), *GDP: A Brief but Affectionate History*, Princeton University Press

Crafts, N. (2004), 'Steam as a General Purpose Technology: A Growth Accounting Perspective', *The Economic Journal,* Vol.114.

Crafts, N. (2013), 'ICT as a GPT: A Historical Perspective'. *University of Warwick CAGE Paper.*

Cutler, D. and Miller, G., (2004), 'The Role of Public Health Improvements in Health Advances: The 20th Century United States'. *NBER Working Paper No. 10511.*

David, P. (1990), 'The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox', *American Economic Review Papers and Proceedings*, Vol. 80, No.2.

Decker, C., (2014), *Modern Economic Regulation, An Introduction to Theory and Practice*, Cambridge University Press.

Espey, J.A. and Espey, M., (2004), 'Turning on the Lights: A Meta-Analysis of Residential Electricity Demand Elasticities', *Journal of Agricultural and Applied Economics* Vol.36 No 1.

Foreman-Peck, J. and Millward, R. (1994). *Public and Private Ownership of British Industry 1820-1990*. Oxford. The Clarendon Press.

Foster, C.D., (1992), *Privatization, Public Ownership and the Regulation of Natural Monopoly*. Oxford. Blackwell.

Fouquet, R., (2014), 'Long-run demand for energy services: income and price elasticities over two hundred years', *Review of Environmental Economics and Policy*, Vol. 8 No. 2.

Gordon, R.J. (2012), 'Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds'. *NBER Working Paper No 18315.*

Gordon, R.J. (2016), *The Rise and Fall of American Growth: The US Standard of Living Since the Civil War*, Princeton University Press.

Haldane, A. G., (2016), *"Whose Recovery?"*, Speech given at Port Talbot, Wales.

Hannah, L., (1979), *Electricity before Nationalisation: A Study of the Development of the Electricity Supply Industry in Britain to 1948*. London. The Macmillan Press.

Kanazawa, M.T. and Noll, R.G. (1994), 'The Origins of State Railroad Regulation: The Illinois Constitution of 1870' in *The Regulated Economy: A Historical Approach to Political Economy*, ed. C. Goldin and G.D. Liebecap, NBER.

Levy, B. & Spiller, P.T., (1994), 'The institutional foundations of regulatory commitment: a comparative analysis of telecommunications regulation', *The Journal of Law, Economics & Organisation*, Vol. 10 pp.201-246.

Meier, H., Jamasb, T. and Orea, L. (2012), 'Necessity or Luxury Good? Household Energy Spending and Income in Britain 1991-2007', *EPRG Cambridge Working Paper.*

Millward, R., (2005), *Private and Public Enterprise in Europe: Energy, telecommunications and transport 1830-1990*, Cambridge University Press.

Newbery, D.M., (2001), *Privatization, Restructuring and Regulation of Network Utilities*, MIT Press

Phillips, D. (2013) 'Living Standards', *Institute for Fiscal Studies*, London.

Pollitt, M.G., (2010), 'The Economics of Energy Demand', *EPRG Cambridge Spring Research Seminar.*

Robinson, B. (2011), London – A Modern Babylon, http://www.bbc.co.uk/history/british/victorians/london_modern_babylon_01.shtml

Stern, J. (2005), 'Regulation and Contracts: Substitutes or Complements? Lessons from UK Historical Experience', *Journal of Policy Reform*, Vol. 6, No 4.

Stern, J. (2012), 'The Relationship Between Regulation And Contracts in Infrastructure Industries: Regulation as Ordered Renegotiation', *Regulation and Governance*, Vol 6, No.4.

Stern, J. (2014), 'The British Utility Regulation Model: Its Recent History And Future Prospects', *Utilities Policy*, December, Vol.31.

Stern, J. (2015), 'Sectoral Regulation and Competition Policy: The UK's Concurrency Arrangements – An Economic Perspective', *Journal of Competition and Law Economics*, Vol.11, No 4.

Stern, J. (2016), 'Competition and Economic Regulation in Infrastructure Industries: Lessons from Economic History and Current UK Debates', *CCRP Working Paper*, No.29.

Stern, J. (2017), 'Utility Regulation, Domestic Consumers and Affordability: Current Debates in the Light of Economic History', *EPRG Energy and Environment Seminar.*

Troesken, W. (2006), 'Regime Change and Corruption: A History of Public Utility Regulation' in *Corruption and Reform: Lessons from America's Economic History* ed. E.L.Glaeser and C.Goldin.

Turvey, R. (1995). The Sliding Scale: Price and Dividend Regulation in the Nineteenth Century Gas Industry. *NERA Topics, No. 16.*

Vickers, J. and Yarrow, G., (1988), *Privatization: An Economic Analysis*, MIT Press.

Wallsten, S., (2007), 'Telecommunications Regulation in US States: Its rise and Impacts in the Early Twentieth Century'. *AEI-Brookings Centre.*