WINIR 2017 Symposium

Christoph Klein
Universidad Rey Juan Carlos, Madrid

A categorisation of infrastructure: how old economic ideas defy the complex reality of markets

Infrastructure is on top of the public agenda around the world. In times of sluggish growth and blunt monetary tools, governments embrace large infrastructure programmes to help boost their ailing economies. Like Trump's apparent intention to invest $1 trillion in American infrastructure, all these plans suggest a too simplistic understanding of the complex nature of long-term capital projects: spend on infrastructure now and reap economic development in the form of productivity increases and higher taxes down the line. First, an overview of the prevailing popular and academic discourse will reveal that public infrastructure policy approaches are grounded in monocausal empiricism, static notions of market failure and a flawed understanding of historical funding processes. We will then take a closer look at those capital goods commonly understood to form part of the category of 'infrastructure' and compare the findings with the views usually expressed by policymakers and their economic advisors. A thorough analysis of the nature of infrastructure assets, their emergence and transformation over time, and the functioning of competitive markets proves that prevalent simplistic notions of economies of scale and monopolistic markets defy the complex reality of market dynamics, making these ideas a dangerous starting point for policy action.

INDEX

1. Introduction

The adequate provision and maintenance of infrastructure is widely considered a *sine qua non* of economic development. Services rendered by infrastructure facilitate the division of labour. Road and rail decrease transport costs and ensure mobility; telephony and data services greatly increase the coordinative abilities of an economic system; water and sewer systems reduce health hazards; and the provision of a reliable and cost-effective energy supply is an indispensable requirement for any society to thrive. Indicators such as road density, telecommunication penetration levels, the number of people living off the energy grid and the availability of transport all speak a clear language. Industrialised countries are blessed with more, better and cheaper services while poor countries substantially lag behind in all categories. The positive correlation between the wealth of economies and their respective infrastructure levels is more than robust. Accordingly, politicians and economists typically identify infrastructure as a prerequisite for economic development and growth. In poor countries, efforts to support infrastructure investment are seen as a blueprint to catch up with richer peers. In industrialised countries, increased investments in public infrastructure also appear attractive in the current environment of zero interest rates and sluggish growth. What, however, is commonly ignored in the public discourse is how infrastructure assets historically came into existence and why certain countries have well-maintained power grids and water supply networks – even in the remotest corners of the countryside – while in many other countries a large proportion of the population still lives off the power grid and fetches water from waterholes.

Despite the enormous importance of infrastructure goods for our daily lives, the academic discussion is surprisingly shallow – if the discussion can be said to comply with minimal scientific standards at all. This might be a consequence of the politically charged nature of infrastructure, but does not merit that research is driven – or captured – by political agendas. Infrastructure assets are among the most complex goods. They span vast geographical areas, require huge upfront investments, thereby tying up significant portions of countries' savings stocks for decades or even centuries, and are among the most innovative areas of economic life with a proven track record of having improved the quality of life for billions of people. This paper aims to shed light on core aspects of infrastructure assets and their funding-related and technological intricacies, which need to be understood to guide a well-informed academic discourse and are of fundamental importance for finance practitioners and political representatives.

We will start the discussion by giving an overview of how infrastructure is regarded in public discussions, what economic benefits political (and other) stakeholders assign to infrastructure – and infrastructure investment in particular – and how these viewpoints shape the political agenda. It follows a closer look at those capital goods commonly understood to form part of the category of 'infrastructure' and what difficulties arise to consider infrastructure as a clear-cut asset class. A thorough analysis of the nature of phenomena such as economies of scale, vertical disintegration of supply chains, infrastructure functions versus services and sequencing makes it clear that the emergence and development of infrastructure networks are far more complex than typically described in the public and scientific discourse. We will show why simplistic assumptions of 'old' economic theories defy the complex reality of markets in time and space and, thus, provide for a dangerous basis for policy action.

2. Infrastructure in the popular and academic discourse

Since the late 1980s, infrastructure has made it to the top of political agendas and economic discourse around the world. Not a day goes by without an article being published in one of the major magazines or newspapers urging the state to spend more money on infrastructure. No policy framework, no national development plan, no reform programme lacks an ambitious tackling of the so-called 'infrastructure problem', a term either referring to the general lack of infrastructure in developing countries or its insufficiency and deterioration in industrial countries. In 2014, Nobel Prize-awarded economist Paul Krugman urged increased public spending on infrastructure as a way to fight recession given that 'the aftermath of the bursting bubble was a very good time to invest in infrastructure'.[1] The American Society for Civil Engineers (ASCE), gave the US a D+ in its 2015 annual Infrastructure Report Card, stating that $3.6 trillion would be required by 2020 to upgrade US infrastructure.[2] An article in *The Economist*, titled 'Build Some Bridges and Roads, Mrs Merkel', advised the German government to invest money in infrastructure and not to worry about balancing the budget.[3] McKinsey, a management consultancy, stated in an infrastructure report published in 2016:'Too many countries have been underinvesting for decades, a trend that threatens to constrain growth.'[4] Former chief economist of the World Bank and US Treasury Secretary Larry Summers raised attention concerning the US's inability to fix its decrepit infrastructure in 2016.[5]

When following the public discourse, the conclusion is that, in general, more infrastructure is needed, that infrastructure investment almost automatically results in economic growth and that it is the task of the government to make sure that the required levels of infrastructure goods and services come into being. In their enthusiasm to embrace infrastructure as a panacea for economic growth, limited attention is paid to a prudent economic, technological and historical analysis of infrastructure assets. Crucial topics such as the allocation of project risks, funding patterns, restrictive regulatory frameworks or public tax spending and indebtedness levels are rarely examined with the appropriate due diligence given the enormous costs associated with infrastructure projects.

From a historical perspective, it has been widely forgotten that the majority of infrastructure assets, which were created during and provided for the basis of industrialisation in the nineteenth century, were financed and owned by private actors. Fuelled largely by private capital, but ultimately driven by changes in society, rapid urbanisation and expanding markets, the infrastructure boom was particularly virulent in towns. Municipalities and public agencies would grant concessions, but infrastructure was financed and run by private companies. For example, the telegraph and telecommunication markets in the United States were highly competitive, with hundreds of players at the turn of century. The creation of the US telecommunication monopoly in 1913 was the result of pork barrel politics, after competition had reduced AT&T's return on investment from 46 % to 8

---

[1] Paul Krugman in *The New York Times*, April 7 2016.
[2] www.infrastructurereportcard.org: 'Every family, every community, and every business needs infrastructure to thrive'.
[3] *The Economist*, October 18 2014.
[4] McKinsey Global Institute (2016): *Bridging Global Infrastructure Gaps*.
[5] Larry Summers and Rachel Lipson in *The Boston Globe*, May 25 2016. One of the few voices that admitted the possibility of an incompetent and bureaucratic government as a reason for delays in building and repairing infrastructure.

% within a decade after the patent protection expired.[6] The rise of the omnipotent nation state, and command-and-control policy structures following World Wars I and II and as a reaction to the Communist threat, also led to the nationalisation or the creation of natural monopolies in all industries regarded to be of strategic importance. Until today, large parts of infrastructure are either owned by the state or tightly regulated. Energy and telecommunication markets have only recently been deregulated in industrialised countries.

It is surprising that the discussion of infrastructure as a particularly important asset category to explain economic progress is a very recent phenomenon having only entered mainstream economic discussion in the 1980s. Adam Smith (1776) was impressed by the benefits of infrastructure (without using the term 'infrastructure'): 'Good roads, canals and navigable rivers, by diminishing the expense of carriage, put the remote parts of the country more nearly upon a level with those in the neighborhood of the town. They are upon that account the greatest of all improvements. They are advantageous to the town by breaking down the monopoly of the country in its neighbourhood.' Despite this early endorsement from Smith, historical economic literature and statistics reflect only broad concepts of a country's capital stock. The mainstream capital theory puts forward the idea of some form of an aggregate of capital goods but contains limited explanatory power to explain economic progress. Classical economists, from Smith to Alfred Marshall, would refer to public works when discussing roads, bridges and water supplies without paying particular attention to the specific characteristics of infrastructure assets and who would be best suited to provide them.[7]

In the twentieth century, the picture changed. Large parts of orthodox economic theory explored questions of whether a particular asset shall be considered as a public or private good, how to ensure that positive externalities associated with public goods benefit the consumer (i.e. that the benefits are not lost due to profit-seeking behaviour of entrepreneurs) and how to regulate so-called natural monopolies. Infrastructure assets such as energy and telecommunication networks were considered national assets, which could not be in private hands due to their overriding importance to a nation's economy. Markets were distrusted, believed likely to waste consumer surplus and national resources by unnecessarily duplicating infrastructure. [8]

The widespread economic stagnation experienced in leading industrial countries in the 1970s caused by excessive regulation, high tax-rates, inefficiently run public enterprises in core areas of the economy and harmful monetary policies as a direct result of interventionist Keynesian policy approaches, resulted in a more liberalised market order commonly associated with the administrations of Reagan and Thatcher and the economists Milton Friedman and FA Hayek. The 1980s saw the first wave of liberalisation efforts in various industries, which accelerated following the collapse of the Soviet Union and socialist economic systems in years 1989 and 1990. Free market-oriented supply side economics gained traction, perhaps best exemplified by the infamous Washington Consensus, a one-

[6] Thierer (1994).
[7] Goldsmith (2014).
[8] Scholars such as Murray Rothbard (1962), Israel Kirzner (1973) and Gerald O'Driscoll (1982), who raised valid arguments against the prevailing economic theories concerning natural monopoly and public goods, were confined to a shadowy existence on the margins of economic discourse. Rather than paying heed to the marginalised Austrian voices, politicians and economists were preoccupied employing the Keynesian toolset of macroeconomic policies in an effort to fight market failures and fine-tune economic results.

size-fits-all economic reform programme prescribed by the World Bank and IMF to instil growth in developing countries trapped in levels of low income and high external debt. One element of the economic policy frameworks, also known as economic structural adjustment programmes (ESAP) and overseen by IMF or World Bank, for the first time explicitly highlighted the crucial role of infrastructure for economic development.[9] From thereon, infrastructure as a holistic concept to explain economic wealth and progress would be – more or less – systematically analysed and form part of the macroeconomic policy tool set.

The theoretical economic discussion of infrastructure has been largely limited to two areas. The first broad area pertains to empirical research seeking to establish a systematic correlation between levels of infrastructure investment and economic growth (and development). Particular attention is paid to the question of whether state-led spending on infrastructure projects is growth enhancing and in which categories the state shall be engaged in spending public money. The tackling of the latter issue leads us to the second area of enquiry, which deals with the notion of an optimal level of infrastructure from a macroeconomic perspective. When neoclassical economists discuss optimal or efficient markets in the context of infrastructure, they typically claim that markets fail in providing the optimal levels from a society's perspective. Notions of market failure, e.g. in the context of natural monopoly and public goods, are firmly embedded in orthodox economic theory and provide for the reason that governments are entrusted with the mandate to implement policies aiming at the provision of optimal macroeconomic levels of infrastructure services.

2.1. Empirical correlation between infrastructure and growth

On the first empirical line of enquiry, the literature is particularly excessive. Hundreds of papers have been published since the late 1980s. We will have to limit the extent of our overview to presenting a few highlights only. Empirical evidence is brought to bear upon two closely related questions that lie at the core of the debate on the macroeconomic effects of fiscal policy: First, is public spending productive? And second, to what extent, if at all, does government spending substitute for private investment? The first question has received considerable attention though the results are, not unexpectedly, ambiguous. Some authors[10] have found a positive empirical relationship between government spending and growth whereas others[11] have found an inverse relationship, i.e. increased public capital stocks were not the cause but the result of increased productivity and economic growth. The outcome is dependent on factors such as the sample of countries, the period under study, econometric techniques and difficulties of measurement and, most importantly, the category of spending. It appears that spending categories that promote physical capital accumulation are positively associated with growth while government consumption has negative effects.[12] In order to determine the net growth effect on the economy, proof is required that growth resulting from

[9] Often ESAPs were imposed on borrowing countries as a condition to obtain loans, debt moratoriums and debt cuts. This conditionality and the social impacts caused by 'textbook' macroeconomic tools such as instant price deregulation and abolishment of food subsidies caused massive social frictions in many countries, amongst others the famous IMF food riots. See Walton and Seddon (1994).
[10] See Rubinson (1977), Ram (1986), Aschauer (1989).
[11] See Evans and Karras (1994), Grossman (1988), Gier and Tullock (1989).
[12] See Easterly and Rebelo (1993), Barro (1988).

public investment is not outweighed by crowding-out private investment and efficiency losses.

Aschauer (1989) for instance finds that the crowding-out effect of public investment is (significantly) exceeded by a crowding-in effect associated with the role of public capital as a productive input and its complementarity with private capital. He was one of the first scholars to stress the specific role of infrastructure, such as highways, streets, water systems, and sewers, as a means to determine optimal productivity levels of public investments.[13] An article reviewing the literature on the link between infrastructure and development from 1989 onwards has criticised that theoretical approaches often fail to clearly lay down the relevant theoretical questions to be addressed and that these studies ignore regulatory frameworks, market structure and detailed institutional mechanisms.[14] Likewise, Hahn et al. arrive at the conclusion that 'only a few of the enormous bulk of studies on the output effects of infrastructure base their estimates on solid theoretical models. But to understand non-linearities and heterogeneity, we must understand the channels through which infrastructure affects economic growth. After all, government roads as such do not produce anything, and to include infrastructure or public capital as a separate input in a production function neglects the usually complex links.'[15]

As a third example, in a paper which empirically analyses the question whether public investment crowds out private investment based on data of 14 OECD countries, the authors find support for the existence of a crowding-in effect of private investment by public investment through the positive impact of infrastructure on private investment productivity. However, this is stated subject to the following proviso: 'Given the data limitations, we need to be cautious about applying these results to particular instances. It is quite possible that specific types of government consumption may help GDP growth, and the opposite might be true for some public investment projects. On the other hand, the analysis presented in the paper does not deny that particular ways of funding public spending may exert their own influence upon investment.'[16] In other words: the authors are prudent in admitting that they cannot generalise. Maybe the best assessment of the empirical literature provided has been made by Prud'homme who stated that 'what matters for economic development is infrastructure usage, but what we observe is infrastructure supply.'[17]

From a methodological point of view, all empirical studies bear the risk that the assumptions of statistical models already embody the results that are to be derived. Even if one assumes a perfect statistical model able to render unambiguous results, economists should be very hesitant to accept the outcome as a policy prescription applicable to the future. Mises pointed out that statistical correlations are purely historical data and cannot be

---

[13] See Aschauer (1989). His work is very often cited as a seminal contribution being the first to systematically assess the empirical relationship between infrastructure spending and economic productivity and growth.
[14] See Straub (2012).
[15] See de Hahn et al. (2007).
[16] See Argimón et al. (1994).
[17] See Prud'homme (2004). In another contribution reviewing the empirical literature, which was presented at an IMF seminar, de Haan et al. (2007) conclude: 'We would like to mention a few issues we believe have not been well researched. First, attempts at explaining existing differences in capital stocks are only in their infancy. Second, only a few of the enormous bulk of studies on the output effects of infrastructure base their estimates on solid theoretical models. But to understand *non-linearities* and *heterogeneity*, we must understand the channels through which infrastructure affects economic growth. After all, *government roads as such do not produce anything*, and to include infrastructure or public capital as a separate input in a production function neglects the usually complex links.' [Italics added].

used to establish laws.[18] The world is a dynamic place and each outcome has to be seen in the historical context of myriad variables and parameters, which, even if they were known, cannot be isolated in respect of cause and effect by econometric models. This is particularly true for infrastructure projects that are characterised by their long-term nature and complex interdependency with the rest of the economy.

2.2. Market failures due to private provision of infrastructure

The second line of enquiry identified in the infrastructure discussion, concerning optimality in infrastructure investment, is centred on the concept of market failure. The debate includes arguments defending natural monopolies and public goods, based on esoteric concepts of lost consumer surplus and non-realised spill-overs arising due to the nature of certain infrastructure categories. Introductory courses in microeconomic theory emphasise perfect competition's blessings and monopoly's curses: monopolists are able to restrict output, set prices and maximise profits at the expense of society, which has to bear the resulting deadweight losses. Paretian allocative efficiency is not achieved and society loses as a whole. Irrespective of the highly dubious nature of any concept of aggregated utility as a social representation of wellbeing, the ideas of the strictly static neoclassical branch of economics have guided policy-makers (and their economic advisors) around the world and ignore the dynamic nature of the economic system and the signalling function of high market shares and high profits, as typically associated with monopolistic market structures. Market structures dominated by one or few profitable players attract competitors and investors; in a competitive economy, success breeds success.

Rothbard however pointed out that monopoly prices do not exist in a free market as there are only free market prices.[19] In line with Rothbard, Hoppe argues that both restriction of output and expansion are part of profit-maximising and market-price formation, and neither of these two aspects can be separated from the other to make a valid distinction between monopolistic and competitive action.[20] Furthermore, O'Driscoll made the observation that every market starts as a monopoly.[21] Compared to the atomic world of perfect competition with passive price takers, it is scale that makes technology available and affordable to the masses in the real world. Inventions, in general, need to achieve scale to penetrate markets. Yet, instead of embracing scale as the engine of civilisation and progress, textbook economics continues to paint the idealised picture of perfect competition, according to which entrepreneurs must be prevented from appropriating undue rents at the expense of society.[22] The fact that substantial amounts of funds need to be raised relative to a society's

[18] Mises (2007, p. 212) criticised that empiricism results in transforming a unique and unrepeatable historical event into a fact of the empirical natural sciences.

[19] Rothbard (1962) assumes a property-rights view. A monopoly, along with the power to sell at a non-market price, can only exist in case governments grant an exclusive privilege. Accordingly, under the assumption of a free market it is meaningless to distinguish between monopolistic and competitive prices as they are both market prices.

[20] Hoppe (1989, p.216).

[21] O'Driscoll (1982).

[22] Nobel prize-winner Samuelson (1989, p. 567) in his famous textbook *Economics*, co-authored by Nordhaus: 'Imperfect competition prevails in an industry whenever the individual sellers have some measure of control over the price of the good in that industry.' Samuelson and Nordhaus explain the degree of monopolistic market structure by way of measuring concentration of the four largest companies in a given industry versus the next

savings base limits the number of market players in a given market in many cases. Furthermore, the entrepreneurial risks associated with early-stage inventions further narrow the ability to raise financial means. Accordingly, markets for certain goods and services have a structural limit as to how many suppliers can be sustained, which is determined by initial investments and the willingness and ability of savers and investors to make funds available.

In the neoclassical world, such markets are often considered 'natural monopolies', which again raises the question of monopolistic market power and its detrimental results for the society as a whole, as discussed above. On a theoretical level, natural monopolies are firstly characterised by economies of scale, i.e. decreasing average costs due to high initial investments into infrastructure facilities and, secondly, by the fact that one large provider can offer the product more cheaply than two or more smaller ones. This simple fact has been formally explained by Baumol et al. (1982), who characterised a natural monopoly based on the concept of a sub-additive cost function. However, Baumol also introduced the idea of contestable (monopolistic) markets deemed efficient due to the threat of short-term entrants. In the absence of entry barriers and sunk costs, and firms having access to the same level of technology, a perfectly contestable market with few players might result. Though all three assumptions are not a given in the real world and the concept of efficiency hinging on the sunk costs assumption is unwarranted, the general understanding of markets as being contestable is a good starting point in that the concept contemplates the dynamic nature of markets.

We cannot review the vast literature on natural monopolies and regulatory action, but the above briefly outlined neoclassical notions have shaped regulatory frameworks of core infrastructure markets across the globe. By way of example, AT&T's president Theodore Vail referred to competition as duplicative, destructive and wasteful. He successfully put forward the idea that fixed-line telecommunication services were to be provided by one single operator as a 'universal service' to everyone, which rendered them a prime target for interventionist regulation. Accordingly, in 1913 the US fixed-line telecommunication market was divided into a long-distance and a local market subject to rigid regulatory control – to the benefit of the incumbent operator. Vail saw regulation as a way to recapture ailing monopoly profits. It was only a few years later that a Senate Commerce Committee concluded that telephone service were to be regarded as a 'natural monopoly'.[23] The economic theory of a natural monopoly was born, eventually providing for an ex-post rationalisation of state action. AT&T provides for a perfect example of what has become known a regulatory capture, a form of state-led behaviour that occurs when a regulatory agency, created to act in the public interest, instead advances the commercial or political concerns of special interest groups that dominate the industry or sector the state is charged with regulating.

When fixed-line carriers were being privatised and exposed to competition by granting competitors access to their networks at the outset of the 1990s, the first mobile networks sprung up. The tremendous expansion of mobile voice brilliantly demonstrates the futility of the natural monopoly argument. In the telecommunication sector, through technological change the same voice minute – and later the same gigabyte of data – could now be offered

---

four largest companies. Hence, highly competitive markets such as the automotive industry, tobacco or electric light bulbs are declared as monopolistic and imperfect.

[23] Thierer (1994).

via different types of networks.[24] Consumers suddenly had the choice of which service to choose and millions in developed countries rapidly opted to skip fixed-line voice services (which were only offered in few urban hubs). In most of developing countries, the introduction of mobile networks provided people with their first ever personal telecommunication service. In most low-income countries fixed-line penetration levels had failed to exceed low-digit numbers, very often not more than 1–2 % of the population.[25] It's probably fair to say that the regulatory restrictions of competition in the telecommunication sector led to a significant delay in technological advances. Public telecommunication providers were hopelessly overstaffed and managed by civil servants who had no incentive to act entrepreneurially. Overall, the technological environment in these entities was stagnant for half a century.[26] Instead, carriers were occupied with fulfilling publically imposed mandates to achieve universal access.

In other infrastructure sectors, such as energy distribution or water supply, technological progress to date has not led to equally disruptive market shake-ups, a feature which can be explained by the pertaining of physical and economic characteristics of the particular infrastructure class. However, in the case of power supply, great advances have been made by separating generation and distribution, which has instilled competition on the generation side and opened up publically protected markets previously considered natural monopolies. In many developing countries, efficient photovoltaic systems and cheaper battery storage will allow consumers and firms to substantially reduce exposure to unreliable public grids, or even to go off the grid. Water heating in low-cost housing in South African townships is largely based on cheap solar heaters. Decentralised, self-reliant power grids are on the doorstep and will replace the need for national grids for many cities and entire regions.

Without going into the detail and based on an admittedly limited coverage, serious theoretical and empirical problems become obvious if one follows the political and economic debates around the infrastructure question. Often, even a minimum level of methodological scrutiny is not achieved, which might not be surprising given political and commercial interests. What is surprising, though, is that the same holds for many scientific contributions of theoretical economists who all too often ruminate the 'old' theories of natural monopoly, public goods and positive externalities without accurately analysing the respective infrastructure category, the role of technological progress and the ability of the market to find commercial and institutional solutions to deal with the challenges the same markets present to entrepreneurs. Mainstream policy approaches still apply static economic theory resulting in overregulated markets and disregard the market as a dynamic discovery mechanism. In order to fruitfully criticise the economic orthodoxy applied in the infrastructure discussion, we have to gain a better understanding of the economic category of infrastructure.

---

[24] In popular, particular political discussions it is usually held that competition should be taking place on the service level, but not on the infrastructure layer, as this is deemed duplicative and wasteful. Nothing could be further from the truth as inter alia the progress in telecommunication is solely explainable by competition on all infrastructure elements of the telecommunication network.

[25] The advent of agile and highly competitive MNOs left behind dozens of bankrupt incumbent fixed-line operators which would only survive through public subsidies, in many cases by way of self-awarding mobile licenses to the state operators based on which they were able to cross-subsidise their inefficiently operating fixed-line operations.

[26] After all, man was flying to the moon in 1969 while fixed-line telephony has been static from a technology perspective for more than fifty years until the late 1980s. The building blocks of today's telecommunication world such as packet switching-based TCP/IP protocol and integrated circuit microchips were being developed when Armstrong set foot upon the moon.

3. Taxonomy of infrastructure

3.1. Definitions

In this chapter we will draw attention to the enormous difficulties in narrowing down the meaning of the term infrastructure. *The Oxford English Dictionary* defines infrastructure as 'the basic physical and organisational structures and facilities (e.g. buildings, roads, power supplies) needed for the operation of a society or enterprise: the social and economic infrastructure of a country.' This is only one of many attempts to capture the meaning of the term infrastructure. Though there is no single definition, intuitively most people associate infrastructure with immobile and long-lived physical assets such as road and rail, energy, water, sewerage and telecommunication networks that require a high upfront investment and are concurrently used by large parts of the population.[27] There is unanimity about the indispensable function that infrastructure plays in modern societies in that it facilitates and deepens the division of labour. The fact that infrastructure assumes such a crucial role in the economic system as a whole seems to justify a macroeconomic viewpoint based on GDP growth and considerations of social equality in terms of fair and equal access.

So far, scholars could not agree on a standard definition. According to Torrisi this arises from 'the need for simultaneous realisation of three analytical objectives: (i) the formulation of a concept for the term "infrastructure", (ii) the incorporation of theoretical approaches (for example, the theory of public goods), and (iii) the description of the reality of infrastructure.'[28] As outlined above, public discussion is centred on the valid idea that infrastructure facilities are important features of economic performance. We have also seen that empirical analysis has not rendered clear results, neither in terms of statistical significance nor as to the direction of causality between the emergence of infrastructure and economic growth.

According to one of the first systematic analyses on infrastructure, Jochimsen defined infrastructure as: '[ ] the sum of material, institutional and personal facilities and data which are available to economic agents and which contribute to realising the equalisation of the remuneration of comparable inputs in the case of a suitable allocation of resources, that is complete integration and maximum level of economic activities.' If we take this definition literally, we have to conclude that everything is infrastructure. Buhr supports this notion and stresses that the identification of the term 'infrastructure' with the term 'material infrastructure' is an unnecessary and misleading contradiction and would add nothing to understanding the problem.[29]

Due to the difficulties in narrowing down the meaning of infrastructure based on physical characteristics and economic attributes, Buhr suggests characterising infrastructure according to its essential functions. A function is *essential* when it initiates changes in economic variables, e.g. changes in costs of firms or changes in household utilities. As such he identifies material, personal and institutional facilities as categories of infrastructure.[30] Institutional infrastructure comprises all customary and established rules of the community

[27] Certain definitions include in the category of infrastructure education and health-related goods or even all institutions important for the functioning of society such as the legal system and the underlying law codes.
[28] Torrisi (2009).
[29] Buhr (2003).
[30] Buhr (2003) considers the provision of institutional infrastructure to be a task of the state.

as well as the facilities and procedures for guaranteeing and implementing these rules by the state.[31] Personal infrastructure is equated with human capital, which, according to Jochimsen, refers to: '[ ] the number and the qualities of people in the market economy characterised by the division of labour with reference to their capabilities to contribute to the increase of the level and the degree of integration of economic activities.'[32] Other authors add social and public infrastructure.

What becomes obvious is that the use of functional criteria to arrive at a precise definition entirely defeats the purpose. The approach waters down the intuitive notion of infrastructure and fails to delineate human capital from physical infrastructure. As an example, a prison is an institutional good resting on a society's understanding of law and justice, but at the same time material infrastructure. The institution of a norm can be codified as a law, i.e. institutional infrastructure, alternatively it can also be seen as personal infrastructure in a sense of human capital as defined by the OECD (Organisation for Economic Co-operation and Development: '[ ] the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic wellbeing.'[33] This definition of human capital as an infrastructure category is arbitrary. Lastly, human capital as a category in itself is too broad a concept to yield sufficient explanatory power for our discussion.

3.2. Characteristics of infrastructure

In the following, our understanding of infrastructure will be based on physical and economic characteristics. Physically, infrastructure facilities are a subset of tangible capital goods, which comprises all durable goods used in the production process, e.g. land, property, machines, buildings, trucks, installations etc. Economically, infrastructure networks such as roads, energy distribution or telecommunication differ from other capital goods in that the former are used 'publically' whereas the latter's use is restricted to the owners, e.g. a company buying a machine or a public agency building an administrational office. By publically we mean that the key feature of the business model termed 'infrastructure' is to serve large numbers of people and enterprises concurrently. The following core economic and technical characteristics are typically associated with infrastructure:

(i) long economic life;
(ii) immobility/indivisibility;
(iii) relatively low technology risk;
(iv) high upfront investment;
(v) large user base;
(vi) network character (public/spatial); and
(vii) ownership of infrastructure supply chain.

Important infrastructure categories typically render their services for a long time (i) given the life span of a human being. Road and rail networks, if properly maintained, can be used for

[31] See Buhr (2003).
[32] Jochimsen (1966).
[33] OECD (2001).

hundreds of years. Copper-based telecommunication landline access and energy distribution networks are generally in use for fifty years and more. Once built, core infrastructure facilities are immobile and indivisible (ii). From an economic viewpoint, it is not useful to tear apart a telecommunication network and redeploy it somewhere else. [34]

The third and fourth characteristics are closely related. The ability to raise financial capital for high upfront investment, which is required in order to roll out infrastructure, is dependent on the overall risk as perceived by capital providers. Usually, funders are willing to accept a certain degree of market risk provided over-optimistic forecasts of user numbers or market prices do not negatively impact on the financial performance when corrected for and would not lead to a financial default.[35] In respect of technology risk however, only those projects that have been tested thoroughly and are deemed sufficiently reliable in respect of their technical features will attract the high upfront investment required for a widespread infrastructure network. This does not mean that technology risk can be mitigated entirely. When the first mobile network carriers entered the market, the strong subscriber uptake quickly led to a congestion of networks. Though the resulting dropped calls and inaccessibility of networks are certainly unpleasant from a user's perspective, they did not present a commercial problem from a single operator's perspective. The rapid growth of the subscriber base (v) and strong competition in all mobile telecommunication markets have provided for the incentive to develop better spectrum utilisation, process higher bandwidths per cell and more efficient data compression methods. The high upfront investment of infrastructure requires a broad user base that is the fifth characteristic we list above.

Finally, we will discuss what can be well regarded as the preeminent two features of infrastructure: its network character (vi) and the ownership of networks (vii). Merriam Webster defines a network as an interconnected or interrelated chain, group or system. There are many other definitions most of which are related to information technology. Since the advent of the Internet, many applications and services are able to offer benefits due to their 'network effects' referring to an increased individual use-value when a network grows. The spread of ICT has helped to drastically decrease transaction costs by linking together people and firms, thereby driving dissemination of information and the division of labour.[36] Essentially, forming networks is the ability of an economy to generate structures which embody knowledge and add value. From a theoretical viewpoint, the advent of ICT is just an accelerated, particularly efficient version of a network structure. Infrastructure networks are constituted by their physical, technical and spatial characteristics that only render certain combinations of capital goods feasible. Austrian economists, in particular Lachmann (1956), refer to the complementary nature of the use of capital. Hence, the specific structure of how a combination of capital goods is arranged determines its functions and usefulness. In case of

[34] The high upfront investment in infrastructure is usually driven by the costs of civil construction works, which are much higher than for example the costs of copper cables, power wires or road surface. The book value of infrastructure assets cannot be recovered by redeploying or selling them.

[35] In long-term infrastructure finance structures there are numerous ways of dealing with loss-making projects, e.g. refinancing to allow for longer amortisation periods with the result to lower short-term debt repayments, a sculpting profile aligned to the market-uptake or additional equity injections by forcing the original sponsors to release stand-by equity, raise new equity or bring in new shareholders.

[36] See Coase (1937), Hidalgo (2015).

infrastructure assets, the specific network character is achieved by a particular combination of complementary capital goods.[37]

Roads and rail only became useful for the wider society due to their widespread, spatial interconnectedness. Power supply and telephony services are distributed through physical networks consisting of masts, ducts, towers, cables etc. A large user base is testament to the important role of infrastructure in economic life. Large infrastructure networks in turn create entrepreneurial opportunities in two ways. First, labour and capital is attracted to provide solutions in all areas of the value chain resulting in a better, more effective and less costly provision of infrastructure. Second, each infrastructure category requires products and services to make use of it. Roads succeeded because of cars, power networks fostered the development of electric light and electronic devices such as TVs, mobile telephony networks resulted in the development of smart phones, the main access to the Internet for billions of users, and the prior existence of national copper backbones and submarine cables explains the incredible success of the Internet to a large degree. Hence, entrepreneurial opportunities arising as a direct (supply side) or indirect result (demand side) from the provision of infrastructure ensure a better utilisation of existing assets and trigger the further expansion and densification of infrastructure networks.

Finally, it is important to understand the ownership structure of infrastructure networks (vii). Historically, the entire value chain of an infrastructure class such as energy grids, telecommunication networks or water infrastructure was in the hands of one legal entity, which facilitated fund raising and gave entrants a control over technology and markets. However, from the fact that a vertically integrated value chain in one entity's hand seems to be efficient at the beginning of an infrastructure life cycle does not follow at all that markets cannot host more than one or many suppliers. Initially, all infrastructure markets were characterised by competition on the infrastructure layer that ensured quick technological advancements and price competition for the benefit of consumers. This virtuous circle was disrupted for most of the twentieth century by public convictions of market failure. Strongly regulated national monopolies were created, which stifled technological innovation and new commercial approaches [38] Not only were markets deprived of competition on the technology side, but even more so innovative commercial approaches that characterise today's competitive dynamics were not coming to fruition in the first place.

Opposed to typical infrastructure providers, i.e. in the energy or water sector, the phenomenon of Internet as a world encompassing infrastructure network rests on fairly atomised ownership structures; the Internet's supply chain is hundred per cent vertically disintegrated. No party controls or owns any significant part of it. Technological advancements such as IP data transfer and smart commercial models yielded massive productivity leaps defying old economic ideas of natural monopoly. In this regard, the Internet cannot be regarded as a classic example of infrastructure, yet it rests on and drives innovation of existing infrastructure.

---

[37] See Lachmann (1956). Böhm-Bawerk emphasised as early as 1890 that capital goods are not individually or intrinsically productive. The sources of any increases in value arise as a function of the production processes capital goods make possible.

[38] Until these days it is asserted that competition should happen on the service layer, but not on the infrastructure layer as the latter is regarded as harmful by politicians – unjustifiably so given the historical emergence of infrastructure.

3.3. The passage of time and economies of scale

In terms of causality or chronology of the developmental process, infrastructure is sometimes referred to as a prerequisite or as a 'basis' to enable development. Nothing could be further from the truth. The invention of the prepaid payment mechanism – a commercial idea rather than a technological solution – solved the credit-vetting problem associated with post-paid contracts in low-income countries, providing the real breakthrough of affordable mobile telephony that moved billions of people onto mobile operators' networks. An additional average monthly revenue per user of mostly less than $5 expanded the revenue base by hundreds of billions of dollars. This and other examples exemplify that there is arguably no such thing as a one-off infrastructure roll-out based on which the direction of commercial and technological development is cast in stone. In a never-ending dynamic process consumers continuously reconfirm the user value of infrastructure. Not only the infrastructure provider, but all entrepreneurs that play a role in the value chain of a certain infrastructure category, contribute to finding better technical and commercial solutions on both the supply and demand side, e.g. from low-data 2G to 4G/LTE high-bandwidth base stations on the supply side, from simple Nokia handsets to today's Apple and Samsung smart phones. These dynamics also apply to apparently less shiny examples such as water supply and sewerage systems. The fact that water utilities provide reliable and clean water to millions of households in today's megacities is due to progress in all areas of the network: chlorination, fluoridation, sand filters, water-saving techniques in flush toilets, water pipes not causing health hazards, desalination plants to protect groundwater reserves. Decentralisation of water infrastructure has grown extensively as a viable solution including rainwater and storm water harvesting where policies aim at a more rational use and sourcing of water.

Hence, we can conclude that in a free market environment, the provision and nature of infrastructure goods and services is subject to constant change in all areas of the value chain due to better technological and commercial solutions arising on the supply and the demand side. When economists speak of economies of scale, they are generally referring to high initial investments and decreasing average cost for the end-user. This is an entirely static view of the world that disregards the dynamic nature of markets. The true nature of economies of scale arises only over time, while ever more users share a network resource. Densification of infrastructure networks, based on increasing numbers of users, is a pivotal feature triggered by the search for better solutions in all areas of the value chain. We will draw a few major lessons to facilitate the ongoing discussion and resulting conclusions.

> Lesson 1: The continuous process of innovation explains economies of scale in the provision of infrastructure goods and services and not the simple fact of decreasing average (or sunk) costs.

3.4. Competitive dynamics and vertical disintegration

The above-described market dynamics not only explain cheaper and better services to an increasing user base, but impact strongly on the supply side of the economy. On the supply side, the traditional treatment of economies of scale in textbooks confines itself to the number of players and the average cost of production, usually proceeding on to market

failure and the calls for regulatory action. In very few cases do scholars deeply investigate the underlying economic forces.

On the face of it, only a few players with powerful brands stand for many industries, simply because they sell the product to the consumer. This view, however, shields the dynamics in the supply chain. According to Ganeshan and Harrison, a supply chain is 'a network of facilities and distribution options that performs the functions of procurement of materials, transformation of these materials into intermediate and finished products, and the distribution of these finished products to customers.'[39] Economists also refer to the effect of vertical integration, understood as degree to which a firm owns its upstream suppliers and its downstream buyers, and important research has been dedicated to the so-called 'firm boundary'. If firms dispose of legal ownership of the entire value chain, they are fully integrated.

According to this approach, production is not (as the production-function model would have it) a matter of combining resources according to explicit blueprints of some sort, but a matter of human skill and experience. The organisation puts its capabilities to use in the coordination of the various activities that go into producing goods and services in the different stages of production. The firm's boundaries will depend on its capabilities; it will undertake those activities to which its capabilities are appropriate and leave other activities to the market.[40]

The success of companies in the automotive industry depends to a large degree on the effectiveness of their supply chain networks. Suppliers, partners, service providers and freight forwarders need to be linked together in such a way as to enable the efficient functioning of business processes that do not stop at company boundaries. The competition between suppliers leads to better products, distribution of competencies across dozens of supply chain partners, the dissemination of knowledge and sharing of revenues and profits. This speaks directly to Adam Smith's famous observation of division of labour.

Some scholars such as Stigler[41] have tried to link the direction of the vertical integration to a particular phase in the growth cycle of industries. However, the attempts of finding universally applicable economic theories to explain market results have a long and harmful history in orthodox economics in that they have drawn attention away from the study of entrepreneurship. Studying how entrepreneurs change the boundaries of their own industries or organisations can help us to make dynamics of a particular industry intelligible. Rather than being primarily driven by the appropriateness of particular theoretical approaches, as they highlight parts of reality, academic focus should shift towards the problems and challenges faced by entrepreneurs, and see what we can learn, in terms of theory, by providing a structured account of the commercial and technological nature of a particular industry.[42]

Generally, it appears that at the beginning of the growth cycle industries are integrated, which is due to the fact that capabilities are monopolised to a certain degree by innovators As

---

[39] Ganeshan and Harrison (1995).
[40] See Langlois (1988).
[41] See Stigler (1951).
[42] See Jacobides and Winter (2005). This view is fully aligned with the Austrian approach to the entrepreneur. Jabobides and Winter call: '[ ] for a shift away from trying to consider the 'pure' theory of 'the' firm, and towards trying to understand the factors that go into the "theory of a firm" – i.e., the theory and belief that an entrepreneur might have *vis-à-vis* their own venture and the way in which the venture can create and capture value.'

soon as other entrepreneurs enter the same market, rivalry not only emerges on the demand side, but also on the supply side. Competitive technology providers do now have an incentive to offer the best (and cost-effective) solutions in a growing market, which leads to product differentiation. The more contested the value chain becomes, the less the former 'monopolist' be able to handle the complexity of the entire production ecosystem in-house. The supply chain begins to disintegrate as new technologies are introduced by suppliers, which will lead to better service and more cost effective production. In the mobile telecommunication industry, many mobile carriers hardly own more than the customer interface and radio base stations. Towers are outsourced into so-called tower companies, fibre connections between towers and exchanges are leased, server farms are hosted in data centres, international connectivity is sourced from submarine cable operators, and minutes and data are being sold through service providers. Finally, mobile handsets and the content such as music, video and social media are entirely outside the sphere of influence of mobile network providers. Efficient production requires the use of specialised assets that have few alternative uses. The owners of such assets are vulnerable to the appropriation of their rents by their contractual partners.[43]

As a consequence, carriers' profits have decreased while the quality and extent of services have dramatically improved over the past twenty years. If these market dynamics entail a consolidation in a particular market sector, e.g. fewer operators in the mobile telecommunication space, then this might be also the result of the fact that better-suited entrepreneurs have taken over many elements and the associated profits of the value chain. Supplier diversification within the value chain is a consequence of the dynamic nature of market economies. Irrespective of the number of market players providing the final product to the consumer, innovation and competition reduce economic market power from within. Moreover, the provision of goods and services by external companies yields a dissemination of the legal ownership base in the supply chain, which is routinely ignored when competition commissions and regulatory bodies analyse market shares of final service providers that hold the customer interface.

Most importantly, the above-examined value chain effects reveal a vital – and often overseen – additional insight to further illustrate the merits of private provision of infrastructure. In a free market based on private ownership, the value chain of the infrastructure network is as contestable as the market for final products. Creative entrepreneurs seek to provide better solutions in all areas of the networks compared to what a fully integrated infrastructure supplier would be able to produce. In a competitive market, infrastructure service providers, for instance mobile operators, are essentially forced to continuously implement better technical solutions in order to exploit (short-term) profit opportunities and/or retain their market position as the value chain partner of mobile operator A strives to provide their newly developed solutions to competing mobile operator B in the same market. Accordingly, an analysis of the number of providers of the final goods and services is unable to explain economic results in a sense of cheaper and better services or dynamics in the industry. If in contrast a particular infrastructure category is run by the state or regulated as a natural monopoly, these competitive dynamics are unable to render their beneficial effects for the society. We can draw the following lesson:

[43] See Williamson (1985).

Lesson 2: Irrespective of the number of players providing the final product in a given market, competition for the elements in the value chain leads to continuous innovation as well as to dissemination of profits and legal ownership. Regulatory focus on concentration ratios of final product suppliers masks the true nature of dynamic economic development.

3.5. Infrastructure classes and service elements

Due to the difficulties of defining and the widespread disagreement concerning what constitutes infrastructure, it follows that the classification of infrastructure differs depending on the respective authors' viewpoints. The following table is based on Goldsmith's classification of infrastructure using service categories and physical works:

Table 1: Infrastructure Classes based on Physical Works[44]

Sector / Service	Typical Physical Works
Transport	Roads; Bridges; Tunnels; Ports; Canals; Railways; Tramways; Airports
Communication	Telegraph; Telephone; Wireless; TV; Internet; Broadband
Water	Irrigation Canals; Water Supply Networks; Dams; Drainage; Flood Defences
Energy	Gas; Electricity; Oil; Nuclear; Renewables
Urban	Public Buildings; Streets; Street Lighting; Leisure Facilities
Environment	Wastewater Treatment; Waste Disposal; Green Infrastructure
Social	Universities; Schools; Hospitals; Social Housing; Prisons

Source: Own classification

These categories are probably typical of what the man on the street would regard as infrastructure. What immediately becomes obvious though is that the classification does not allow for a useful differentiation in terms of the particular network features. Roads and power lines certainly represent network elements, whereas a nuclear plant will make use of power lines but it does not make up the network character of the energy grid. In the case of social facilities, such as hospitals or prisons, they present no network in either a physical or a geographical sense. In Table 2 below, a distinction is made between the true network features, which is the distribution of services, and those elements required to operate the respective infrastructure class.

[44] Goldsmith (2014).

Table 2: Infrastructure Classes Along Network and Operational Elements

Infrastructure Sector	**Network Elements (passive infrastructure)**	**Operational Elements (active infrastructure/plants)**
Transport (1)	Roads; Bridges; Tunnels; Ports; Canals; Railways; Tramways	
Communication (2)	Wires; Cables; Ducts; Masts; Street Cabinets	Radio Modules; Data Centres; Servers; Software
Water (3)	Canals; Water Pipes; Sewerage; Irrigation Canals	Turbines (dams); Water Works; Desalination Plants
Sewerage (4)	Sewer System; Pipes	Wastewater Plants
Energy (5)	Power Wires; Masts; Ducts; Substations	Power Plants (Oil, Gas, Nuclear, Renewable)

Source: Own classification

It is easily recognisable that the presentation in Table 2 yields a better understanding of network structure as it pertains to the respective infrastructure class. In the case of streets, no operational element is required, except to make the network available to the user. This is not the case for classes 2, 3, 4 and 5. The energy network, for instance, is only used when power plants generate energy. Urban infrastructure does not represent a separate category, but in essence consists of classes 1 to 5 on a dense geographical area. The sectors environment and social fall outside of our definition of infrastructure. Garbage collection and landfill facilities do also not represent infrastructure as they do not constitute a network in a physical or spatial sense.

Finally, in Table 3 we examine those elements required by the user to utilise the service rendered by the respective infrastructure class as well as the specific services the user obtains. The categorisation applied in Table 3 emphasises the enormously important role of infrastructure and at the same time delineates the core categories 1 to 5 from the remaining categories. While the function of a particular class of infrastructure defines to what extent the respective can be used, it does not explain why it is used. Merriam-Webster defines 'function' as follows: 'The action for which a thing is specially fitted or used or for which a thing exists.' This meaning is closely related to the definition of 'purpose': 'The reason why something is used.' As an example, the power grid exists because it provides electricity. However, the reason why electricity is used is that human beings greatly value the amenities associated with the use.

Table 3: Infrastructure Classes Along User and Service Elements

Infrastructure Sector	User Elements (Hardware)	Service Elements
Transport (1)	Rolling Stock; Cars; Trucks	Mobility; Transport (Trade)
Communication (2)	Telephones; Computers; TVs; Fax Machines	Communication; Entertainment; Information; Knowledge; Trade
Water (3)	Ships; Watergates; In-House Piping; Water Taps; Heaters	Water; Transport
Sewerage (4)	In-House Piping; Toilets	Health; Civilised Life
Energy (5)	Electrical Appliances; Heating	Electricity

Source: Own classification

Categorising infrastructure along both user and service categories highlights that the value of the respective infrastructure classes is driven by consumer preferences and behaviour. Preferences change as entrepreneurs invent new products and exploit opportunities by changing the functional characteristics of how people can make use of infrastructure. This in turn reaffirms and strengthens the value of infrastructure. Only the concurrent emergence of cars, trains, telephones, computers, TVs, ships, water taps, heaters and myriad electrical appliances truly make the economic value of infrastructure networks intelligible. This value entirely and solely arises as a result of human needs to meet their ends. Humans act in that they travel, exchange goods and information, drink and shower, do not like to feel cold and aim to live a comfortable and safe lifestyle. Entrepreneurs, who are also consumers at the same time, seek to exploit opportunities that ensue from consumer demand. The emergence of infrastructure arises as a consequence of consumers trying to live a better life. The hardware elements required to make use of infrastructure emerge in tandem with the infrastructure network. Once consumers and entrepreneurs understand the benefits, a mutually reinforcing development process based on demand (user) and supply (infrastructure provider, supply ecosystem) kicks off the literal virtuous circle of technical and commercial innovation. At the same time, the resulting higher wealth of society increases incomes and frees up time that will be dedicated to more productive uses. In essence, the entire economic system is a network linked by consumer needs. Infrastructure is required to hold it together and oil the links. The following lesson can be drawn:

> Lesson 3: Infrastructure services arise as a consequence of economic actors meeting ends. The entrepreneurial means employed to satisfy these ends determine the physical, economic and spatial nature of the respective infrastructure category and surrounding ecosystem.

3.6. Lexicology and sequence

Usually, lexicological (or etymological) questions are treated at the outset of a scientific investigation. Based on the previous chapters, however, we shall now more easily see that the term infrastructure is too diffuse to capture the complex reality of the presented topic.

Infrastructure consists of the two words *infra* and *structure*. *The Oxford English Dictionary* defines structure as 'an arrangement and organisation of interrelated elements in a material object or system, or the object or system so organised' whereas on Google one finds two definitions. Google either refers to 'the arrangement of and relations between the parts or elements of something complex' or to 'a building or other object constructed from several parts.' The word structure derives etymologically from the Latin verb *struere* 'to build' and the noun *structura* 'structure'. The meaning of the verb is to 'construct or arrange according to a plan; give a pattern or organisation to.' We see that the verb describes a person acting – in that the actor creates something that serves a purpose (purposeful human action). Any acting is dynamic in nature, hence oriented towards the future. As the future is uncertain, acting is always entrepreneurial, which entails error and failure. The meaning of the noun 'structure' in an economic context refers to something stable that the actor can reliably use or something intangible upon which human acting can place reliance. Structures help to reduce uncertainty and complexity and serve as signposts in an uncertain world.

In Latin *infra* means below. The combined meaning of infrastructure refers to a structure that is required in order to create something else, e.g. to build the base for a house. There is a logical sequence of events implied as it's certainly not useful to start building a house without completing the groundwork. When we apply this thinking to economics, we infer that certain goods or institutions need to be created in order to produce something else. As an example, rail infrastructure has played a major role for economic development as it greatly facilitated trade and mobility. We can justifiably say that rail networks serve as an underlying basis for the economy. At this point, the public discourse around the benefits of infrastructure usually ends. However, to be able to usefully capture the true meaning and function of infrastructure, we have to make two crucial qualifications.

First, in the case of a given set of infrastructure assets, the value of the enabling role of infrastructure for the economy system – its *infra* feature – needs to be confirmed daily by consumers and entrepreneurs using the facilities. In a hundred years, we may not need trains in their current form as other transport means may have emerged. In the German Ruhr Basin, a region of heavy industries, more than 300 kilometers of rail tracks have been converted into bicycle tracks as the regional industrial structure has transformed away from coal and steel plants towards high-technology and service industries. In this case, the railway tracks, which were used for more than a hundred years and provided adequate pay to society, are assuming a different function. The case exemplifies that the functions ultimately arise from the subjective value considerations of its users. The fact that today's accelerated technology cycles render infrastructure assets uneconomic and foster the emergence of alternatives is a pervasive feature of a dynamic economic environment and needs to be factored into the planning process of infrastructure projects, particularly if the state acts as planner and owner of infrastructure projects.[45]

[45] In South Africa, the Zuma-led Government is in the process of procuring a nuclear energy plant to overcome its base load problems. At the same time, massive gas deposits have been discovered in the northern part of neighbouring Mozambique, which is already supplying gas to South Africa. Compared to the technologically mature, efficient and clean gas power stations, the extremely long building periods of nuclear plants, which exceeds ten years, will not only inhibit the development of the gas industry (and other power infrastructure) in the region, but also constrain South Africa's public budget due to the enormous costs of pursuing the nuclear option.

Second, and this is the key point in the light of the general public conviction that infrastructure is a prerequisite for economic growth, the creation of infrastructure is not a one-way track. When we look at the historical emergence of rail networks in the US or England, we must ask why and how entrepreneurs (not the state) raised the massive funds to be invested into infrastructure. Producers of agricultural goods in rural areas badly connected to trade hubs were seeking for quicker and more reliable ways of trading their commodities. This need had arisen because progress in farming techniques, such as irrigation, use of fertilisers and machinery, made the operation of large farms possible, resulting in output far exceeding subsistence levels. Farmers were able to sell the surplus production at a profit and accumulate savings. This example provides for a striking illustration of the real nature of savings, which Austrian economists never tire to stress. In Böhm-Bawerk's words: 'Productive powers are to be set free for the proposed formation of capital, and this can only be done [ ] through saving.'[46] In the US, it was the farmers themselves who raised finance and funded substantial parts of rail infrastructure.[47] But this could only happen as a result of other commercial opportunities that were exploited and in turn made the necessary funds available. Massive productivity increases in the agricultural sector preceded the emergence of railways. In today's discussions, the logic generally applied mostly works the other way around. First build, and then reap economic growth. This simplified approach, particularly advanced by governments in developing countries as part of their national development plans, ignores that infrastructure itself is based on many factors such as savings, entrepreneurial action, legal institutions conducive to long-term risk-taking or the existence of sophisticated capital markets, to name a few. We shall draw the following lesson:

> Lesson 4: In a free market environment, the emergence of infrastructure goods is always preceded by the existence of other economic opportunities that yield the returns required to accumulate savings and provide the capital base for entrepreneurs to assume risk.

Accordingly, the term infrastructure is actually misleading in two ways. First, the roll out of infrastructure requires pre-existing structures, very often complementary to other types of infrastructure. Second, once built, the emergence and continuous existence of a particular type of infrastructure is highly contingent upon the overall economic environment, particularly the invention of new technologies and dissemination of knowledge. Hence, the services rendered by infrastructure are exposed to the same principal economic laws as other goods and services. In a dynamic economic environment, infrastructure assumes as much an underlying feature for the plans of consumers and producers and the use of other technologies, as its existence hinges on exactly the same plans and technological advances. In this regard, infrastructure is as much *infra*-structure as it is *inter*-structure and *intra*-structure in terms of the role it plays in the different areas of the economic system and for different kinds of actors. Today's dynamic and complex economic systems are constituted by a never-ending push-and-pull of supply and demand, driven by actions that mutually reinforce, neutralise or inhibit each other, but eventually result in higher levels of prosperity if plans – guided by the market-based incentive mechanisms – succeed. Hence, we are able to conclude that infrastructure is not a stand-alone category with cornucopia-like powers that

[46] Most famously illustrated by Böhm-Bawerk's (1889) example of Robinson Crusoe.
[47] See Eichengreen's (1994) paper in which he draws lessons from the Railway Age, applying them to financing infrastructure in developing countries.

inevitably generates overflowing economic prosperity, but merely an integral part of the value chain of economic life, though a very important one.

4. Conclusion and outlook

The last six sub-chapters show that 'old' economic ideas are inconsistent with the historical and present development of infrastructure markets. Telecommunication, roads and water might share the typical characteristics of infrastructure assets – in essence nothing else than a reflection of the importance of infrastructure for the productive base of the economy – but are fundamentally different in crucial respects. There is no such thing as *a* or *the* infrastructure that somehow assumes an autonomous status in the economic value chain in a sense that more is better. Yet, it is this conviction that dominates the public discourse, which might explain why the journalistic and academic standards related to questions of infrastructure appear astonishingly superficial and imprecise. Generally, linear thinking in terms of cause and effect dominate the discussion: first build the infrastructure, and then reap the fruits of economic development. Too little attention is paid to the study of history as a learning device to better understand the emergence of infrastructure.

The hitherto presented insights of (i) the dynamic character of economies of scale due to innovation, (ii) the beneficial function of vertical disintegration as a shield against market dominance, (iii) the driving role of consumer demand for infrastructure provision and (iv) the understanding that infrastructure itself is a result of economic development are all but neglected in public discussions. When following the current treatment of infrastructure in public media, politics and economics, the strict top-down perspective of the theoretical discourse becomes apparent. Multilateral institutions, consultancies, Think Tanks and journalists outgun themselves to find ever-bigger infrastructure gaps (valued in US$) the market will fail to plug, be it in emerging or industrialised markets. They all commit a grave intellectual error by assessing infrastructure investments decoupled from innovative capabilities and the savings base of a society.

It's worthwhile to go back to economic roots. Lachmann made clear that the stock of capital must not be regarded 'as a homogeneous aggregate, but as a structural pattern.'[48] Infrastructure assets are subject to the same fundamental economic principles as established by the Austrian School, rooted in individual human action and based on subjective assessments of means and ends. Infrastructure goods are complex in nature due to their very characteristics: they are long term, costly and difficult to finance, are being used by millions of actors, determine and are being determined by the wealth of society and are greatly dependent on institutions such as the legal system and treatment of property rights. Our current infrastructure base has evolved over centuries. Driven by technological progress and changing consumer preferences – very often based on higher income levels – the combination of capital goods is in constant flux. At a given point in time, this combination might appear as a given. Dynamically, however, new (or alternative) technologies and changing consumer behaviour constantly change structural patterns.

Breakthroughs in chip technology and radio spectrum utilisation led to the success of mobile telephony tapping into a huge dormant customer base. The ensuing massive demand

[48] Lachmann (1956).

for cellular handsets, in turn, set free the development of new classes of end-user devices. Lack of or low quality fixed-line networks further fuelled demand for cellular services, leading to a quick succession of mobile radio standards from GSM (2G) to LTE (4G). Two decades ago, total telephone penetration (fixed-line plus mobile telephones) stood at 3,3 % in emerging nations and just 0,3 % in the least developed countries.[49] On a worldwide basis, in 1990 only 10 % of population had access to telecommunication, now we are approaching 100 %.[50]

We conclude that the existing base of all infrastructure categories is a result of human creativity and complex financing processes undertaken by entrepreneurs and sponsors assuming market risks. Yet, notions of market failure are dominating political and academic discussions. Despite the fantastic success of entrepreneurial creativity demonstrated in the telecommunication (and information technology) sector, thereby rendering futile orthodox economic theories of 'natural monopolies' and 'network externalities', the state still exerts a firm regulatory grip on most infrastructure sectors. If a lack of understanding and linear thinking couple with political mandates based on convictions of market failure and more-is-better approaches, then we are on a dangerous path.[51]

[49] ITU (2002).

[50] www.data.worldbank.org. Penetration levels are based on SIM card numbers. Based on individual subscribers, e.g. those owing a handset, the level is lower due to multi-SIM use. However, usually households share phones.

[51] The danger of treading old economic paths become even more obvious when analysing funding processes and risk allocation typical for infrastructure projects (both subject of a second treatise presented by the author of this paper).

5. References

Ansar A, Flyvbjerg B, Budzier A, Lunn D (2016) Does infrastructure investment lead to economic growth or economic fragility? Evidence from China, Oxford Review of Economic Policy, vol. 32, issue 3

Argimón I, González-Páramo JM, Alegre JR (1994) Does public spending crowd out private investment? Evidence from a panel of 14 OECD countries. Banco de Espana

Aschauer DA (1989) Is public expenditure productive? Journal of Monetary Economics 23

Barro RJ (1988) Government spending in a simple model of endogenous growth. National Bureau of Economic Research Working Paper no. 2588

Buhr W (2003) What is infrastructure? University of Siegen, Siegen

Coase RH (1937) The nature of the firm. Economica. The London School of Economics and Political Science, London

de Hahn J, Romp W, Sturm JE. (2007) Public capital and economic growth: key issues for Europe. International Seminar on Strengthening Public Investment and Managing Fiscal Risks from Public-Private Partnerships, Budapest

Easterly W, Rebelo S (1993) Fiscal and economic growth: an empirical investigation, 32. Journal of Monetary Economics 32

Eichengreen B (1995) Financing infrastructure in developing countries: lessons from the railway age. The World Bank Research Observer, vol. 10

Evans P, Karras G (1994) Are government activities productive? Evidence from a panel of US States. The Review of Economics and Statistics, vol. 76, issue 1

Ganeshan R, Harrison T (1995) An introduction to supply chain management. lcm.csa.iisc.ernet.in/ scm/supply_chain_intro.html. Accessed February 5 2004

Gier KG, Tullock G (1989) An empirical analysis of cross-national economic growth. Journal of Monetary Economics

Goldsmith H (2014) The Long-Run Evolution of Infrastructure Services. CESIFO working paper, no. 5073

Grossman PJ (1988) Government and economic growth: a non-linear relationship. Public Choice 56

Hidalgo C (2015) Why information grows: the evolution of order, from atoms to economies. Basic Books, New York

Hoppe HH (1989) A theory of socialism and capitalism, 2010 edition. The Ludwig von Mises Institute

ITU (2002), World Telecommunication Report 2002 – Reinventing Telecoms.

Jacobides MG, Winter SG (2005) The co-evolution of capabilities and transaction costs: explaining the institutional structure of production. Strategic Management Journal, vol. 26 (5)

Jochimsen R (1966) Theorie der Infrastruktur, Grundlagen der marktwirtschaftlichen Entwicklung, J C B Mohr (Paul Siebeck), Tübingen

Lachmann LM (1956) Capital and its structure. London School of Economics and Political Science. Bell & Sons Ltd, London

Langlois RN (1988) Economic change and the boundaries of the firm. Journal of Institutional and Theoretical Economics, vol. I44

Mises L (2007) Theory and history. The Ludwig von Mises Institute

OECD (2001) The Well-being of Nations – The Role of Human and Social Capital.

O'Driscoll GP (1982) Monopoly in theory and practice, in Kirzner (1983): 189-213.

Prud'homme R (2004) Infrastructure and development. Annual Bank Conference on Development Economics, Washington

Ram (1986) Government size and economic growth: a new framework and some evidence from cross-section and time-series data. American Economic Review

Rothbard MN (1962): Man, economy and state. Van Nostrand, Princeton

Rubinson R (1977) Dependency, government revenue, and economic growth. Studies in Comparative International Development, 12

Samuelson PA, Nordhaus WD (1989) Economics, 12th edition. McGraw-Hill, New York

Stigler GJ (1951) The division of labor is limited by the extent of the market. The Journal of Political Economy, 59 (3)

Straub S (2012) Infrastructure and development: a critical appraisal of the macro-level literature. Journal of Development Studies. Taylor & Francis (Routledge), Abingdon-on-Thames

Thierer AD (1994) Unnatural monopoly: critical moments in the development of the bell system monopoly. Cato Journal, vol. 14, issue 2

Torrisi G (2009) Public infrastructure: definition, classification and measurement issues. Munich Personal RePEc Archive mpra.ub.uni-muenchen.de/12990

von Böhm Bawerk E (1889) The positive theory of capital. Macmillan and Co., London

Williamson OE (1985) The economic institutions of capitalism. Free Press, New York