Infrastructure as a complex adaptive system

Ed Oughton & Pete Tyler

2013
Abstract

National infrastructure systems are increasingly being challenged to deliver what societies require of them. This includes the need to address the de-commissioning of ageing infrastructure assets, rapid demographic change and a variety of environmental pressures, while critically maintaining economic competitiveness. It is thus crucial to understand more about the economic forces that influence both the supply and demand for infrastructure, and the nature of the economic interdependencies that exist between different components and sectors of the infrastructure stock. However, much of our understanding of infrastructure economics is based on models that adopt assumptions that seem to be rather at odds with the reality. This article argues that we need to reframe our thinking of infrastructure systems and this can be achieved by utilising concepts drawn from complex adaptive systems (CAS) theory. It can help recognise the interdependencies that exist between supply and demand, between infrastructure sectors, and how the agents of national infrastructure systems tend to adapt and co-evolve over time. This perspective is illustrated with a case-study example of Information Communications Technologies (ICT) infrastructure, where this complex adaptive lens is found to be particularly amenable for thinking about this sector. Further research should use this lens to examine the workings of other infrastructure sectors, as well as move away from theory to critically review the quantitative methodologies utilised in the field. This should include their potential applicability to understanding infrastructure systems, with the ultimate goal of improving decision making and the formulation of national infrastructure strategies.

Key Words: Infrastructure Systems, Complex Adaptive Systems, Information & Communications Technologies (ICT)

1. Introduction

Infrastructure in developed countries is becoming increasingly challenged as a result of escalating climate related hazards, ageing populations and the fact that many components of current national infrastructure systems are themselves ageing and coming to the end of their lifespan (OECD, 2006). Indeed, change is very evident in the extent to which energy and Information Communications Technologies (ICT) now underpin the functionality of almost all infrastructures; hence we have transitioned to a position where infrastructure sectors are increasingly interconnected and interdependent (Rinaldi et al. 2001; Stefanini, 2008; Rahman, 2011a; 2011b; Hémon and Benoît, 2012). Although on a day to day basis it may seem that infrastructure hardly changes, when one examines a national infrastructure system over numerous decades it has evolved to have a significant impact on settlement patterns, how society functions, the structure of an economy and importantly the economic competitiveness of cities, regions and nations. A thorough understanding of how the long
and rigid lifespan of infrastructure can affect economic outcomes is crucial, particularly when it comes to understanding why economic activity occurs where it does and what scope there is for change. Without it there is the risk of different economic geographies being locked into less productive and unsustainable growth trajectories.

With the exception of transportation planning, perhaps one of the most significant gaps in the existing evidence base has been a relatively limited understanding of the ways in which changes in the supply of infrastructure may modify the demand for it and lead to other long-term changes in the economy and society as a whole. In other words what is lacking is a better understanding of the positive economic feedbacks that may exist between supply and demand, and thus how agents of infrastructure systems may adapt and co-evolve over time, including the economics that lie behind these changes.

Substantial policy attention has been given to considering how concepts drawn from behavioural economics (Wilkinson & Klaes, 2012; Kahneman, 2012) may improve our understanding of individual behaviour. However, while behavioural economic approaches provide more insight into decisions and preferences in the short-term than the prevailing neoclassical economic discourse, it does not provide a novel lens for considering the evolution and trajectory of entire national infrastructure systems over long-term periods such as decades and centuries (Grubb et al. 2013). Moreover, the rigid application of the assumptions shared by neoclassical economics and general systems theory have been widely criticised (Nelson & Winter, 1982; Dosi et al. 1988; Savioiti & Metcalf, 1990; Nelson, 1995; Boschma & Lambooy, 1999; Boschma & Frenken, 2006), and therefore we need new ways to understand how national infrastructure systems and their agents adapt and change through time. A complex adaptive system (CAS) approach may be able to assist with this, as well as with understanding how to plan, finance, manage and regulate infrastructure systems. In this article we investigate this possibility by building on the work of others who have examined this problem (Rinaldi et al. 2001; Herder & Verwater-Lukszo, 2006; Nikolic & Dijkema, 2010; Lei et al. 2010; Dam et al. 2013).

Section (2) considers the characteristics of a national infrastructure system. Building on work of Arthur (1999), Delorme (2010) and Lei et al. (2010), Section (3) then outlines the basis of a CAS approach and how it can be distinguished conceptually from general systems theory. In Section (4), an examination of how the properties of a CAS approach can be used to understand more about the workings of national infrastructure systems is carried out. Section (5) then focuses on the specific example of ICT infrastructure as being of particularly amenable to a CAS approach. The article concludes by suggesting where further research might be focused.

2. What is Infrastructure?

High-quality national infrastructure systems comprised of the energy, transport, water, waste and ICT sectors are essential for economic growth, development and prosperity, and
for maintaining and improving the wellbeing of all members of society (World Economic Forum, 2012; OECD, 2006; Straub, 2011). National infrastructure systems enable goods and services to be traded between places, public health to be maintained and individuals to socially interact across space. The impact a nation’s infrastructure can have on a range of economic, social and environmental factors should not be underestimated.

However, defining infrastructure is difficult as the term is used to refer to a variety of objects and technological artefacts. Conceptual and measurement problems proliferate in the economic literatures with prominent concerns being how to define and obtain adequate measures of what constitutes the infrastructure stock (identified prominently by Gramlich, 1994). Consequently, before defining what constitutes infrastructure, it is more appropriate to discuss its functions and attributes. Elsewhere these have been treated as separate approaches (e.g. Torrisi, 2009), but here an overview will be provided in order to synthesise a working definition in the specific context of this research.

Infrastructure is diverse because many different assets carry out diverse activities, and as an enabling device it helps us fulfil certain economic, social or environmental means, we can rightfully deduce that this diverse group of assets provides a range of different services to intermediate and end users.

**2.1 The Key Characteristics of a National Infrastructure System**

The different functions of infrastructure are process based and carry out sequential operations in a co-ordinated way. A range of technologies embodied in infrastructural assets are controlled so that they can carry out one or multiple functions, in a chain of operations. Increasingly, infrastructural assets require inputs from other infrastructure systems to function. Single or multiple processes are then co-ordinated to provide *infrastructure services*. Often this takes place in a hierarchical manner at a range of spatial scales, and consequently these can be seen as a range of nested processes. Two of the main processes which infrastructural assets can undertake are the *transformation or preservation* of different material and immaterial entities. These processes can be carried out to a range of materials and objects or to intangible forms of capital such as information. In addition, infrastructural assets are also able to *transmit* and distribute these material or intangible forms of capital across space.

Given this overview we can define infrastructure as the *co-ordinated operation and management of a group of physical assets to perform a range of processes, thereby providing infrastructure services to users.*

Evidently, infrastructure can be many things to many different individuals and institutions – infrastructure is often a catch-all term for production and consumption enabling commodities and assets. In economic terms infrastructure can function as an asset, service or market, and can carry out processes which transform, preserve or transmit. Asset planning times are frequently elongated, often exceeding a decade or sometimes even more. Infrastructure services can have particularly low substitutability but potentially large
complementarities with other factors of production (Baldwin & Dixon, 2008). Infrastructure assets are spatially-fixed durable commodities with very long life spans, and hence they can be categorised under Courant & Deardorff’s (1992) definition as ‘lumpy’ production factors. These systems are frequently very large in scale and consequently, particularly with regard to the aforementioned factors, they can become susceptible to ‘lock-in’. Infrastructure investments, particularly once reinforced by increasing returns, can lock infrastructure systems on path dependent economic or environmental trajectories which are incredible hard to break away from due to the substantial financial hurdles involved with path divergence.

An important economic characteristic of infrastructure is that it forms a network (Erdős & Rényi, 1960). This is well recognised in contemporary infrastructure analysis (Atkins et al. 2009; Tranos, 2013; Lam et al. 2013) where changes in network characteristics, from efficiency, to scale, structure and resilience are all part of the evolution of a network. From an economic perspective, this is where the narrative of increasing returns plays a considerable role in understanding the topology of infrastructure networks and, for example, why the degree of connectivity can differ greatly between the nodes of an infrastructure network.

Yet, infrastructure networks do not just appear overnight, they are intrinsically woven through the intra-urban fabric of the built environment over decades, as well as forming an economic lattice of links between cities and regions (à la Christaller, 1933). Their growth is incremental, and therefore can have a significant effect on the shaping of the built environment. Indeed, the built environment in turn shapes the way in which individuals and firms routinely interact within the spatial economy. Consequently, as a network the presence of network externalities means that user’s of infrastructure services can in some circumstances gain increased benefit as more users begin to use certain services. For example, communications externalities in the spatial economy have long impacted on producers, supply chain intermediates and consumers (Capello, 1994; Geenhuizen et al. 2005). As a result of these network properties and increasing sectoral, geographic interdependencies (Robert & Morabito, 2010), markets underpin both the supply of and the demand for services in national infrastructure systems.

Infrastructure is deeply intertwined in all economic and societal functions. It mediates the way we create new value, how we move across space, the way we interact and communicate. In particular, some types of infrastructure can have wide impacts on the economy affecting practically all economic activities (Helpman, 1998; Lipsey et al. 2005). Transport, energy and ICT are good examples as they can have the most dramatic economic effects on productivity, location and innovation. Recent analysis indicates that infrastructure stocks can positively impact on long-run output by somewhere between the range of 0.07 and 0.10 (Calderón et al. 2011). Although a variety of papers use different infrastructure stock definitions and econometric techniques, recent studies indicate that there are generally positive economic effects from infrastructure investment (although they can vary by infrastructure sector), even if they are relatively humble (Démurger, 2001; Crescenzi and
Firms who rely on infrastructure services are attracted to highly connected places with resilient and reliable infrastructure networks, particularly if they can also provide extra capacity for growth (Holl, 2004). The ability to transport raw goods and finished products for manufacturers, and to transport people to disparate clients for firms in the tertiary sector, is essential. Hence, the cost and quality of infrastructure services are an influencing factor in firm location (Rietveld, 1989; Vickerman, 1991; Vickerman et al. 1999; Anderson, 2012). Spatially bound external resources and capabilities, including the fixed capital stock of infrastructure, have an influence on productivity, strategy and the options available to firms. Microeconomic benefits eventually have an aggregate economic impact (Karlsson et al. 2007).

Whether an infrastructure service is supplied to a desire level can influence the ability for activities and sectors to develop and flourish at specific points in space (McCann, 2013). Furthermore, quality transport links impact on labour market access causing a highly accessible place to achieve enhanced job matching between firms that require specialist labour and individuals who require employment (Reggiani et al. 2011). Not only can this improve labour supply access for firms, but it can also improve access to producer services, customer demand and knowledge providers.

Increased accessibility opens up new markets as well as expanding existing ones, helping firms achieve greater economies of scale (Lakshmanan and Anderson, 2007; Lakshmanan, 2011). An efficient infrastructure system is also able to mediate space to a greater effect, increasing the benefits of spatial agglomeration which arise from the concentration of production, labour markets and consumers (Banister & Berechman, 2001).

Human capital such as knowledge is a vital production function input for firms and infrastructure systems are able to facilitate augmented transfer of information (Vinciguerra, et al. 2011). This is also especially pertinent given that economies are dependent on new innovations to create new value (Cooke, 2012). This transfer can result from greater face-to-face interaction, which amplifies un-coded tacit knowledge transfer (Glaeser et al. 1992; 2011). Moreover, greater capacity for exchanging data and other types of coded information electronically can aid this transfer over increasingly large geographical areas.

Infrastructure assets can be private goods, public goods or have properties that make assets fall somewhere in between. Some infrastructure components have the properties of being excludable and rivalrous. They are, in terms of welfare economics, classed as private goods because an increase in consumption by one person reduces the availability of the resource for another (Acocella, 1998). Users can also be excluded from using them. Alternatively, Coase’s (1974) lighthouse is the classic example of an infrastructure component being a pure public good. It is both non-exclusionary and non-rivalry, and hence no market can exist for it to function privately because there is a market failure. The use of this infrastructure
component by another user increases the cost of the service provided by (almost) nothing. Infrastructure that constitutes a public good is a special type of externality.

National infrastructure systems usually have mixed ownership and operation arrangements, where governance frameworks frequently span private firms and individuals, public institutions and third-sector organisations. Many infrastructure sectors now have a unique range of private and public actors and institutions involved with the planning, delivery, operation, ownership and regulation of infrastructure activities (Buhr, 2003). This reflects the historical legacy of many infrastructure systems, which were once publicly owned and centralised systems. Past governance arrangements significantly shape the current character, structure and operation of national infrastructure systems and continue to have profound hysteretic impacts in the future. As an example, Table 1 details the diverse attributes of the UK national infrastructure system.

**Table 1. Characteristics of the UK National Infrastructure System (Adapted from Hall et al. 2012)**

<table>
<thead>
<tr>
<th></th>
<th>Energy</th>
<th>Transport</th>
<th>Water</th>
<th>Waste</th>
<th>ICT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scale</td>
<td>Scale</td>
<td>Scale</td>
<td>Scale</td>
<td>Scale</td>
</tr>
<tr>
<td></td>
<td>National</td>
<td>Regional</td>
<td>Regional</td>
<td>Regional</td>
<td>National</td>
</tr>
<tr>
<td></td>
<td>International</td>
<td>National</td>
<td>International</td>
<td>Regional</td>
<td>International</td>
</tr>
<tr>
<td></td>
<td>Ownership</td>
<td>Ownership</td>
<td>Ownership</td>
<td>Ownership</td>
<td>Ownership</td>
</tr>
<tr>
<td></td>
<td>Private</td>
<td>Mixed (by mode)</td>
<td>Mixed (by region)</td>
<td>Mixed (by region)</td>
<td>Mixed (public responsibility with private operation</td>
</tr>
<tr>
<td></td>
<td>Governance and Regulation</td>
<td>Governance and Regulation</td>
<td>Governance and Regulation</td>
<td>Governance and Regulation</td>
<td>Governance and Regulation</td>
</tr>
<tr>
<td></td>
<td>Varies e.g. electricity has unregulated market prices but regulated network charges</td>
<td>Varies e.g. rail has regulated efficiency targets; roads and government planned with some private provision</td>
<td>For England and Wales, price and investment regulated by Ofwat, drinking water quality regulation by DWI, environmental regulation by EA.</td>
<td>Local Authority run. Environmental regulation by EA/DEFRA in England and Wales, SEPA in Scotland</td>
<td>Competition regulation by Ofcom</td>
</tr>
<tr>
<td></td>
<td>Issues</td>
<td>Issues</td>
<td>Issues</td>
<td>Issues</td>
<td>Issues</td>
</tr>
<tr>
<td></td>
<td>Security of supply</td>
<td>Congestion</td>
<td>Demand management</td>
<td>Energy costs</td>
<td>Waste minimisation and recycling targets</td>
</tr>
<tr>
<td></td>
<td>GHG emissions</td>
<td>High speed rail</td>
<td>Climate change</td>
<td>Environmental regulation</td>
<td>Resource</td>
</tr>
<tr>
<td></td>
<td>Airport capacity</td>
<td></td>
<td></td>
<td>Technological innovation</td>
<td></td>
</tr>
</tbody>
</table>
Like all national infrastructure systems, the actors involved in Table 1 operate over a multitude of spatial scales including the local, regional, national and international levels. Each actor is motivated by a completely different blend of drivers and constraints. This makes it a very challenging task for those trying to manoeuvre each system to provide economically efficient, spatially equitable and environmentally sustainable outcomes.

Moreover, the national infrastructure system does not remain static. Its ownership, operation and institutional arrangements change in response to wider economic, societal or environmental factors. As Markard (2011) observes, the speed of change is highly dependent on each sector’s capital intensiveness, asset durability, environmental impact, form of organisation, degree of regulation, competition and ‘systemness’ (the amount of interdependency or complementarity shared with other sectors).

In the next section we consider the key properties of a CAS and assess how it might assist in understanding how a national infrastructure system functions. In making this assessment we draw upon the growing literature of CAS theory.

3. The Properties of a Complex Adaptive System

This section identifies the key features of a CAS and how they differ from that of a general systems approach. A complex system has a multitude of individual components and agents that are highly connected and interdependent, to the extent that ‘emergent’ behavioural phenomena occur which cannot be explained using other reductionist approaches. Complexity theory is used as a form of guiding meta-theory to understand a range of natural and social systems (Waldrop, 1993; Allen, 1997; Lansing, 2003; Bennet & Bennet, 2004; Batty, 2007).

Many authors make the distinction between systems that are simple or complicated, but not complex (Cilliers, 2002; Bulloch and Cliff, 2004; Garnsey & McGlade, 2006; Martin & Sunley, 2007; Dam et al. 2013). This is because many find it easier to begin by defining what complexity is not, before attempting to define what it is. Table (2) draws upon the work of Arthur (1999), Delorme (2010) and Lei et al. (2010) to compare and contrast the properties of general systems theory and a CAS approach. In what follows we refer to this comparative analysis.

Table 2. Properties of General Systems and Complex Adaptive Systems (Arthur, 1999; Delorme, 2010; Lei et al. 2010)
In a complex system functionality arises not only from the multitude of (often non-linear) interactions between the physical components and incumbent agents of the system, but also from how the system per se interacts with its surrounding environment. System openness is an inherent feature of a CAS.

Complex systems are seen to undergo a variety of possible states. They are in a state of flux which, often resulting from self-organising tendencies, changes the configuration of the system. It is important to make the distinction between a complex system and the subsequent concept of a complex adaptive system. The word adaptive comes from the Latin *adaptatio* which relates to the modification required to suit new conditions. Thus, to adapt is to improve over time in relation to one’s environment (whether natural, economic, social, technical, institutional or some other variant). But the key to the definition is the verb to improve. Responding mechanistically to external stimuli by simply changing is qualitatively different from adapting to gain an advantage (Nikolic & Kasmire, 2013).
Winder et al. (2005) state that not all complex systems display evolutionary dynamics as some can exhibit mechanistic dynamics. The dynamic quantitative change which results from a complex system responding to an external stimulus may display only mechanistic, responsive change. The evolutionary dynamics evident in complex adaptive systems only take place if the relationships between the parts of the system components are modified, in order to gain some form of advantageous position. Hence, the ability for a group of entities to generate a variety of responses to a changing selection environment is a defining feature of an evolutionary system. The absence of this changing, evolving behaviour may indicate that a system is only dynamically mechanistic, rather than evolutionary. Consequently, a CAS can be defined as containing a large number of agents which interact, learn and most crucially, adapt to changes in their selection environment in order to improve their future survival chances (Holland, 2006).

The main properties of a CAS can be identified as evolution, aggregate behaviour and anticipation (Holland, 1992). Evolution is a key feature of complex adaptive systems, and whilst adaptation can be described as the improvements made by entities in response to external environmental stimuli, evolution is different as it is the algorithmic process that produces these improvements (Nikolic & Kasmire, 2013). Yet the use of this metaphor as a transformative process is often ill-defined (Boschma and Martin, 2010). For a process to be described as evolutionary it must exhibit certain properties, specifically the generation of novelty endogenously, from within the system. A system is not evolutionary if change is only incorporated as some form of exogenous shock which momentarily changes the system’s ‘equilibrium’ position. Moreover, evolutionary systems are also dynamic and undergoing perpetual change, with the relationships between the key components in continual flux. Also, the process of evolution is discontinuous and irreversible. In the Darwinian sense, the generation of novel phenotypes out-compete other less adapted species through natural selection and force their exit from the ecosystem forever. In evolutionary economics, this metaphor has been widely applied in relation to the competitive market mechanism. Thus, Schumpeter first wrote about how new products, technologies, processes, industries and sectors have a ‘creative-destruction’ effect. For example, less efficient, old technologies and firms are replaced by more efficient better adapted forms. The changes are discontinuous and inherently irreversible because these processes are temporally dependent. One cannot go back in time.

As noted by Geels (2002), evolution is a two stage process. Firstly, local competitive forces lead to the selection and retention of novelty. Competition can either take place between two novel entities (such as different energy transmission standards – AC versus DC), or it can take place between the novel and the old. Secondly, once selected a novel entity goes through a process of diffusion and unfolding which takes place across space. Those entities or behaviours that are old, inefficient and uncompetitive are forced from the system. This process of novelty, selection through competition, and diffusion is one reason why CASs exhibit perpetual dynamics and move between different states of configuration.
4. Applying The Key Concepts To an Infrastructure System

Chaudet at al. (2009:2) state that after taking into account key properties, infrastructures ‘can be considered excellent metaphors of complex systems’. But to gain a true understanding of a specific system one must undertake in-depth application. This section considers the extent to which the key features of a CAS appear to characterise the workings of a national infrastructure system. The analysis is undertaken at the agent, network and the system level.

**Agent Level Properties**

Adaptive, evolutionary behaviour

Incumbent firms, households and individuals in the national infrastructure systems exhibit *adaptive behaviour* change driven in response to changes in their selection environment. Competition within the system, to varying degrees, forces continual adaptation to changing supply and demand conditions. This is particularly true for infrastructure service providers. These behaviour changes can result from new technologies, new infrastructure assets and flows of financial investment on the supply side. On the demand side alteration in markets, commercial and residential demand, demographics and spatial change can force adaptive, evolutionary behaviour. An example of this change is evident when firms introduce new technologies in production and distribution as, in search of efficiency, they lead to changes in economic routines. Institutional change and natural environment constraints can also play an influencing role.

Diversity among agents and more realistic assumptions

The *heterogeneous attributes of the firms, households and individuals*, incumbent to the supply and demand of infrastructure services are visibly evident. National infrastructure systems comprise some of the largest man-made systems in the world and have great diversity among different types of agents, with even greater variety within these individual types. The ongoing liberalisation of infrastructure is leading to an increasingly varied number of markets for infrastructure services, mediated by regulatory bodies, and operated for different populations over increasingly large geographic areas (Chaudet at al. 2009). Particularly commercial strategies have become increasingly diverse and specialised with the hope of providing each operator with a competitive advantage in the market place.

**Network Level Properties**

Perpetual Dynamics

*Perpetual dynamic change* is especially evident in national infrastructure systems when one examines their long-term impacts. For example, the development of infrastructure is extremely incremental, so the physical network according current and expected demand undergoes a perpetual process of expansion, modification and contraction at different points in space. Moreover, the flows across the physical network are also altered in
accordance with these processes. Innovation has a particular impact on driving change in technological and institutional regimes in infrastructure systems (Weijnen & Bouwmans, 2006). Disruptive technologies and new forms of organisation combine with and result from, the adaptive, self-organising behaviour of the firms, households and individuals that produce and consume infrastructure services on a daily basis. The economy and society is thus seen to perpetually co-evolve with changes in infrastructure networks, advancements in infrastructure technology and organisation, and the new trajectories of economic growth and development.

**A Far-from-equilibrium State**

The factors that influence the demand for, and supply of, infrastructure are subject to continuous dynamic change and thus it would appear naive to adopt a theoretical perspective that assumes it is at an optimal equilibrium at some point in time. A CAS approach would appear more appropriate, as it pays more attention to the imperfections of the system.

**Openness**

National infrastructure systems are open systems characterised by inward and outward flows. They have no precise, fixed boundary between the system under investigation, other nested systems in the hierarchy, and the wider environment. While, we define ‘the national infrastructure system’ purely for practical purposes in research, it is not an isolated, ‘closed’ system. Instead it undergoes constant interaction and exchange with its economic, social and natural environment. The system requires a wide range of inputs to flow into different infrastructure sectors to enable functionality, such as energy, raw and intermediate materials, labour, information and financial capital. This openness can result in the system behaving differently to ostensibly similar shocks from its environment, at different points in time. This is because national infrastructure systems can change their structure over temporal periods, changing the system’s complexity. Homogenous responses to perturbations are highly unlikely.

**Irreversible**

The notion of irreversibility in infrastructure systems refers to the fact that time-independent decisions are rare. In fact, infrastructure providers are almost certainly taking decisions within a set of constrained capabilities, because they must work with durable, long-lived (and often ageing) infrastructure assets and networks. ‘Lock-in’ effects with infrastructure assets often prevent viable transitions to other forms of organisation and operation. For example, in the energy sector many countries have addressed emissions controls by introducing regulatory mechanisms combined with energy efficiency and renewable energy generation policies. These have inevitably impacted on the cost of energy because they attempt to move the system away from the existing ‘locked-in’ state. Moreover, there are also microeconomic irreversibility’s resulting from the adaptive
behaviours of firms, households, individuals, regulators and other institutions who inevitably explore, learn and retain information relating to their activities.

**System Level Properties**

**Complex**

The system is complex because it is comprised of diverse adaptive agents and there is distributed control throughout the system. Moreover the functionality of the system arises not only from the multitude of (often non-linear) interactions between the physical components and incumbent agents of the system, but also from how the system *per se* interacts with its *surrounding environment*.

**Emergence and limited functional decomposition**

The highly dynamic structure of national infrastructure systems is partly due to the high level of interconnection between components. Therefore, *they cannot be functionally decomposed for analysis into individual stable parts*. For example, the considerable human-technology interactions present demonstrate that the system is socio-technical in nature. Large technical infrastructure systems comprised of countless technological artefacts, are deeply intertwined with the human aspects of how firms, households and individuals produce and consume infrastructure services, and how we decide institutions should govern their operation and usage. Unlike in the natural sciences where individual natural processes can often be split off and compartmentalised, the socio-technical processes pertaining to the economic, technological, spatial, demographic, institutional and environmental aspects of national infrastructure systems need to be considered and examined in the widest sense, because they are not readily decomposable into single units (Epstein & Axtell, 1996).

**Distributed control**

National infrastructure systems in developed, free market economies do not have one individual entity in control of the system. This can be problematic for coordination. Often top-down management approaches can yield undesirable outcomes as a result. A prime example of a decentralised infrastructure system with distributed control is that of the ICT sector.

**Non-determinate and non-linear**

The national infrastructure system, and the economy and society it sustains and reacts to, is thus highly interdependent. The interdependent flows of inputs can promote or constrain change in the system. As a result of increased interconnectivity in national infrastructure systems, interactions between incumbents are not independent and positive feedback mechanisms develop which can deeply modify system dynamics (Miller & Page, 2007).

Feedback mechanisms arise when externalities in the system alter the costs and benefits which accrue from an individual’s decisions, therefore causing behavioural change. Positive feedback often has a destabilising effect on a system, while negative feedback can create an
effect which is homeostatic. Non-linear feedback mechanisms can cause the system to undergo transitions to other organisational states. Moreover, endogenously created technologies can be responsible for this type of transition, inducing qualitative change in the relationships of key system components. For example, the development of energy and transportation systems over the past century has instigated organisational and operational transitions in the national infrastructure, encouraging co-evolutionary change in tandem with the demand frontier.

The non-linearities and the degree of technological innovation inherent in the system make unpredictable outcomes occur. These emergent outcomes might not have seemed logical or possible from close examination of the actions of individuals. Complex adaptive properties mean that the system can shift to a very wide variety of possible future states. Indeed, the path dependent properties of certain components often play a part in this. There is thus a need to move away from the past ‘predict and provide’ planning approaches used for large technical systems, and recognise the importance of greater understanding of the plethora of interacting technical and social processes which can push the system into states that arise unexpectedly. Rather than focusing solely on technology and infrastructure, we should shift our thinking to also include socio-technical regimes, including the users and their behaviours in infrastructure analysis. There is perhaps no better example of the properties of a CAS than in the Information Communications Technology (ICT) infrastructure.

5. Information Communications Technology (ICT) Infrastructure as a Complex Adaptive System

In this section we select the ICT infrastructure sector to assess its workings using a CAS approach. ICT infrastructure is most suitable for this purpose because it is subject to very rapid technological change and encompasses a very wide range of activities. Thus, it is an ideal sector to test the relevance and suitability of a CAS approach over others. The underlying properties of the ICT system were assessed using an approach that contained two main elements. The first involved a detailed and extensive desk based review of the large body of existing research that has been undertaken on the ICT infrastructure sector in recent years, particularly with regard to its economics. The second required undertaking core interviews with those who could be regarded as having a good understanding of the workings of the sector and thus able to describe the system’s properties. Ten interviews were conducted with representatives spanning a major UK Internet Service Provider (ISP), computer science academics based in Cambridge, UK, consultants from a leading telecommunications consultancy and a representative from UK regulator OFCOM. These were conducted between December 2012 and February 2013.

5.1 What is ICT Infrastructure?

ICT infrastructure is comprised of the software and systems that store and process information, as well as everything from wired and increasingly wireless networks and their
components (cables, masts, satellites, etc.), to broadband, voice, data and positioning services. Here we focus specifically on the latter infrastructure components of ICT, whereby the co-ordinated operation and management of these physical components provide infrastructure services to users. This approach excludes electronic end-user connected devices, as they do not constitute part of the ICT infrastructure.

ICT infrastructure development is responsible for increased capital deepening in firms, and is an enabling device for creating innovative new products and processes which have positive economic growth and productivity effects. However, ICT does not just have demand side economic impacts, as it is a force that also mediates the supply of other infrastructure services – be it energy, transport, water or waste. As discussed previously, it fundamentally changes how other infrastructure services are operated and delivered.

The human aspects of the ICT infrastructure system must be considered as intrinsic to its functionality and operation. These include the behaviours of supply side network operators, hardware and software developers, ICT specialists, as well as those demand side intermediate and end users consuming ICT goods and services for economic or social purposes. This combination of technology and social agency typifies ICT, like its other infrastructure counterparts, as a socio-technical system (Dam et al. 2012). From relatively humble beginnings in the 1950s, ICT and its infrastructure components have undergone multiple phases of transition (Bullock & Cliff, 2004), so much so that this sector has percolated practically every part of human life in the developed world. The extent to which ICT infrastructure is now used can be demonstrated from examining usage statistics. For example, in 2013 there were 2.7 billion global Internet users amounting to 40% of the world’s population, with 6.8 billion mobile phone subscriptions, standing at 96% of the global population (ITU, 2013). The expansion of the ICT infrastructure system can be credited to the growth in computing power growth, intensification in fixed and mobile devices and the combination of computers with advances in communications technologies. This has led to device integration across multiple operating platforms. Moreover, commercial, public and third-sector organisations have increasingly embedded these technologies in their operational processes and practises (Hanseth and Ciborra, 2007). This has all culminated in increased ICT infrastructure demand with user’s requiring higher and higher bandwidths.

5.2 Changes in the Sector

The ICT infrastructure system is in a state of perpetual dynamic flux. ICT operates between a variety of user needs which are continually satisfied by multiple technical solutions and combined with the uniform standards that make interoperable interchange possible (Garnsey et al. 2006). Technological change takes place at a high rate with many digital products and services being transferred quickly with negligible cost (Meijers et al. 2008:5). Hence, knowledge regarding infrastructure usage is quick to disseminate over large geographical areas, yet it provides firms with only a momentary advantage before being eroded by competitive forces. Some ICT infrastructure components have short life-cycles
and their use often become standardised across multiple types of activities over relatively short time periods.

It is unrealistic to seek to understand the workings of the ICT infrastructure system by assuming it converges to some sort of ‘equilibrium’ state in a mechanical way. New technologies which utilise new content, applications and services significantly impact on the demand for ICT infrastructure. For example, an increase in supply bandwidth does not equilibrate with demand as higher bandwidth content, applications and services become available for users. Over temporal periods, demand actually increases with augmented supply of the network infrastructure. Hence the supply frontier for infrastructure providers must adapt to this positive feedback, whereby the system’s structure and internal flows fluctuate. It is thus more sensible to consider its existence as being it in a far-from-equilibrium state.

Multiple transmission technologies are being newly implemented in the network infrastructure, as well as there being fundamental changes in the different transmission mediums available. These include the supply evolution from narrowband communications using copper to DSL technologies, or cable and fibre fixed lines, as well as wireless methods of transmission such as WiFi, WiMax and 3G and 4G. The move to fibre technologies takes advantage of the interesting characteristics of fibre-optic cable; that the capacity per cable is vast, and we might use only a fraction of it depending on the technology utilising the cable. Moreover, wireless technologies have very interesting implications for carrying out economic activities over space. This incremental change in ICT infrastructure allows the introduction of new product and process innovations in other sectors of the economy, amplifying further structural transformation. For example, when exploited locally it can enable firms to re-engineer business processes, reconfigure business networks and redefine their abilities and scope of control (Remenyi et al. 2007). Hence, changes in supply side technologies have the ability to further modify demand for content, applications and services through feedback effects.

5.3 The Role of Diversity and Innovative Behaviour

The force of innovation is so strong in the realm of ICT, that future organisational states even for relatively short time scales are inherently hard to predict. It is thus challenging to predict the technologies which will govern our behaviour in the future, especially as our own future behaviours will influence technological development in the ICT sector. Yet, this rapid, but high impact change makes the sector a very important part of the evolving economy.

Taking Fransman’s (2010) ICT ecosystem model as a guide, there is evidently an extremely diverse set of actors involved in each of the six hierarchical layers – from producing network elements, operating the physical infrastructure network, programming the TCP/IP interface, providing connectivity services, developing middleware, search and innovation platforms, delivering content, application and services, through to the vast array of final consumers who each demand a specific, niche range of these layered services. The actors in the ICT
ecosystem come together as an ensemble, based on a variety of complementary parts to enable functionality.

In the ICT ecosystem, *no one actor is in control of the infrastructural assets*, as a plethora of stakeholders hold ownership or control of different local, metropolitan or wide area networks. For example, the Internet is comprised of over 42,000 Autonomous Systems which indicatively operate independently, and are owned and managed by separate organisations. Hence, the Internet is referred to as a network of networks, with each service provider operating their own policies over their own assets. Systemic functions are distributed across incumbent components at heterogeneous scales, giving the system a high degree of distributed connectivity. While there are common standards in the form of codes and computer languages (think TCP/IP), there is no organisation centrally planning the development of the ICT ecosystem; it is a prime example of a *decentralised, distributed system with no top-down central command*.

The seamless web provided by communications providers and users alike, is merely the end product of an open but complex process which utilises and emits a variety of tangible and intangible inputs and outputs with its external economic, social and natural environments. These are fused together by the national and global investment and financial strategies of operators, and their governing financial actors, with the deployment of infrastructure and its associated services being dependent on mix of economic, historical, regulatory and physically specific factors (Rutherford, 2011). Over the past two decades, changes in the technological ensemble of ICT infrastructure have caused qualitative change in the economic system, transforming the underlying structure of many developed economies. This has furthered the economic trend toward information-centric activities, the proliferation of connected devices with considerable spatial economic and social impacts. Increased connectivity and subsequent space-time compression, has been another incremental shift in the long-term economic evolution of many developed countries whereby the frequency of economic transactions and the speed at which the economy functions have both increased.

5.4 The ICT Ecosystem

The ICT ecosystem can be reduced into individual parts for scientific enquiry, but because they are highly connected and the system’s structure is open and dynamic it provides only a limited understanding of how they fit together and function as a complex hierarchical structure. We have moved in recent decades from ICT systems which were only individually linked into small-scale networks with relatively isolated and limited elements, to large-scale networks which are highly interconnected and spatially sparse with innumerable components.

This vast expansion has been enabled by the standardisation of the interconnection mechanisms between different layers in the ecosystem. Far greater scale and heterogeneity has developed in each hierarchical level, from competition in network components and
connectivity, to different applications and services which utilise interoperable platform technologies.

The functionality of the ICT ecosystem consequently emerges from the individual actions of those incumbent elements and agents responsible for individual parts of the overall system. Although they do adhere to a variety of standards enabling integration of their network with the overall network, no one individual is directing their activities leading to collective emergence and self-organisation. In the communications industry, incumbents and new entrepreneurs undertake these activities in search of profit, the classic ‘root’ example of Schumpeterian change. Self-organising behaviour is frequently evident when a form of selective pressure is placed on a population, evident in this case with the competitive market mechanism. Consequently, this leads to metaphorical parallels between natural and socio-economic systems, aiding in our understanding of how they organise and how order emerges.

As identified by Garnsey et al. (2006:177) the selection forces shaping the advances in ICT in recent years have been consumer demand, the allocation of investment and market competition. Self-organisation has been exhibited as a governing principle in local, metropolitan and wide area ICT network infrastructures, and collective emergence has been displayed from the large number of simple elements which have cumulatively produced sophisticated outcomes. When processes of self-organisation are combined with the responsive, anticipatory decision making of incumbents, CAS theory tells us that these systems change their structure and dynamics. We have persistently seen this adaptive behaviour with network operators in relation to changes in demand. Individual technological advances including the proliferation of mobile phones, to DSL technologies, and even to platform systems for ICT devices, have caused infrastructure operators to rapidly respond by shifting to different organisational states and different market strategies.

Moreover, the dynamic, adaptive behaviour of incumbent agents makes it difficult to gain tractable control of the sector. The increase in opportunities to make profits in the ICT ‘ecosystem’ has led to rapid, increased development of disruptive innovations making it inherently non-deterministic. Indeed this is problematic for those needing to understand the future direction of the sector, particularly over the long-term. Indeed, the level of discontinuity seen in the ICT sector over the past two decades has been unprecedented.

Overall, the underlying properties of a CAS offer a promising way of understanding change in the ICT sector. The approach inherently rejects concepts of equilibrium, convergence, homogeneity and determinism in the classic sense. For those who desire an understanding of long-term dynamic change in this important infrastructure sector it emphasises the importance of diversity, adaptive evolutionary behaviour and the ability for new technologies to provoke non-linear change.
6. Conclusion

In this article we have sought to build on the core properties identified by Arthur (1999), Delorme (2010) and Lei et al. (2010), to assess whether a CAS approach might help to provide a better understanding of the forces underpinning change in national infrastructure systems. We find that a CAS approach appears to characterise the workings of infrastructure systems well. It has the potential to integrate some of the dynamic, emergent phenomena we see in reality, particularly in relation to sectors like ICT that are especially characterised by rapid change.

Further research could investigate the application of the CAS approach to a variety of other types of infrastructure that vary in their ability to adjust to the forces that operate on them. It would thus be important to understand how the implications of a CAS approach vary for infrastructures with quite different characteristics along the dimensions spelt out in Section (2) of this article. Moreover, further research might want to also move from this theoretical perspective to examine whether the methodologies employed in the field of complex adaptive systems might be of specific relevance for furthering our understanding of infrastructure systems. In this way our ability to understand the nature and determinants behind the changes observed could increase.

Acknowledgements

The authors would like to thank Professor Jim Hall, Professor Ron Martin, Dr Jean Francois-Mercure, Professor Douglas Crawford-Brown, Mr Marco Jano-Ito and other members of the Energy Systems Research Group for reviewing and discussing this paper. The research reported in this paper was part of the UK Infrastructure Transitions Research Consortium (ITRC) funded by the Engineering and Physical Sciences Research Council under Programme Grant EP/I01344X/1.

References


