Physical and institutional challenges of low-carbon infrastructure transitions: constraints and potential solutions

Katy Ellen Roelich

Submitted in accordance with the requirement for the degree of
Doctor of Philosophy

The University of Leeds

School of Earth and Environment

July 2014

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work in Chapter *5* of the thesis has appeared in publication as follows:

Roelich, K., D. Dawson, P. Purnell, J. Busch, C. Knoeri, J.K. Steinberger (2014) Assessing the material criticality of infrastructure transitions: a case of low-carbon electricity systems. *Applied Energy.* 123, 378-386. Special Issue on Supply Chains and Energy Security

I was responsible for devising indicators, undertook the case study analysis and wrote the whole paper. Phil Purnell devised Supply Disruption Potential normalisation. Christof Knoeri made significant contributions to the development of the production-requirements ratio and detailed comments on the paper. Ruairi Revell, David Dawson and Jonathan Busch collected data and commented on the final paper. Julia Steinberger commented on indicator development and provided detailed comments on the paper.

Some of the work in Chapter 8 of the thesis has appeared in publication as follows:

Roelich, K., C. Knoeri, J.K. Steinberger, L. Varga P. T. Blythe, D. Butler, R. Gupta, G. Harrison, C. Martin, P Purnell (2014) Towards resource-efficient, service-oriented integrated infrastructure. Accepted in Technological Forecasting and Social Change.

I was sole author of sections 4.1 and 4.3, co-author of sections 1, 4.2, 5 and 6 and edited the whole paper.

Some of the work in Chapter 8 of the thesis has appeared in publication as follows:

Roelich, K., Steinberger, J.K. and Knoeri, C. (2013) Efficient and service oriented infrastructure operation: the role of the Energy Efficiency Directive in driving change towards Multi Utility Service Companies. In: eceee Summer Study proceedings. European Council for an Energy Efficient Economy Summer Study, 3 - 8 June 2013.

I was responsible for undertaking all analysis and wrote the whole paper. Julia Steinberger and Christof Knoeri undertook detailed reviews of the paper.

The work in Chapter 9 of the thesis has appeared in publication as follows:

Roelich, K., T. Schmieder, C. Knoeri, J.K. Steinberger (2013) The role of governance in accelerating transition towards more integrated, service-oriented infrastructure operation. Presented at IGov workshop 'Theorising change for a sustainable economy'. April 2013
I was responsible for undertaking all analysis and wrote the whole paper. Tina Schmieder contributed to data collection, identification of case study characteristic and commented on conceptual framework and draft paper. Julia Steinberger and Christof Knoeri undertook detailed reviews of the paper.

Acknowledgements

There are many, many people whose invaluable support and inspiration has made this thesis not only bearable but, dare I say enjoyable. I would like to give sincere thanks to:

John Barrett, who has been a constant source of inspiration since I first came up with the idea of doing a PhD and convinced me it was a Good Idea. Then he gave me a job, which put a spanner in the works, but as a result the PhD took a rather different but better path, I think. His enthusiastic support of outcome-focussed research has had a significant influence over the direction of my research and my future aspirations.

Julia Steinberger, who helped me work out what this PhD was from the seemingly incoherent mass of research I had done. Without her healthy dose of realism who knows how or when this thesis would have happened.

Frin Bale, who has provided continual motivation, unwavering support, wise advice and plenty of cakes that kept me going through the hard spots. And she was right. I can totally do this.

Tina Schmieder, who was my co-interviewer on many of the interviews and transcribed the majority of them. She also provided invaluable input to the first iteration of case study analysis.

Tom Curtis, who first planted the idea of doing a PhD many years ago and has since intervened with sage advice at crucial points throughout my career and this PhD.

Peter Taylor, who voluntarily and uncomplainingly reviewed my most difficult chapters.

Phil Purnell, champion of the utilitarian writing style, who developed the method for normalisation of Supply Disruption Potential used in Chapter 5.

The sure-infrastructure team, who acted as sounding boards throughout this research and specifically Christof Knoeri for helping with the development of the production-requirements ratio and detailed reviews of many of the papers contributing to this thesis, David Dawson, Ruairi Revell and Jonathan Busch for essential data gathering and reviewing the paper contributing to chapter 5.

The participants of all the case studies, who gave their time generously and shared their enthusiasm and frustration willingly.

EPSRC, who let me into a sandpit and subsequently funded two projects, which supported this research.

Family and friends for their support and patience and particularly to Rhona Smith, whose sound advice in the latter stages of this thesis kept me sane and to my Mum and Dad, who were so excited that I was finally doing a PhD that it kept me excited all the way through.

And last but certainly not least, Ben, who has been through every step of this with me, picked me up when I hit a hard spot and cheered me on when I hit a good spot. Without him I would have hit the buffers a long time ago.

Abstract

There is growing recognition that the United Kingdom's ageing infrastructure systems are unable to deliver the radical reductions in greenhouse gas emissions necessary to avoid dangerous climate change. As such, there is an imperative to transform our infrastructure systems towards alternatives that deliver services reliably but within environmental limits. There are significant challenges to achieving this transformation but most current analysis focusses on the technical and economic challenges of infrastructure transition. This thesis examines two under-studied challenges to low-carbon infrastructure transition: one principally physical; the constraints posed by the disruption in supply of critical materials embedded in low-carbon energy technologies; and one institutional; the constraints to alternative modes of infrastructure operation from current policy and regulation in water and energy infrastructure. It aims to not only characterise these constraints but also to identify policy responses to alleviate constraints.

The two constraints differ greatly in character and contrasting methods were used to analyse the nature and scale of each constraint. Material criticality constraints were examined using a quantitative, indicator-based method developed in this thesis to dynamically assess the risk of critical material disruption to low-carbon electricity generation. Policy and regulatory constraints were analysed using theory building from case study analysis to identify the mechanisms by which development of alternative modes of operation were constrained by policy and regulation.

Despite the differing scale and nature of the constraints, there are some striking similarities in the potential policy responses to constraints. The results of both analyses emphasize the importance of diversity in the future infrastructure system, the need for a more targeted approach to policy and stress the need for integrated action across policy areas. The dual focus on understanding and responding to constraints forced a balance between dealing with complexity and enabling action. This highlighted the importance of adaptive policy, which takes action in the face of uncertainty but is able to modify its course as system understanding develops.

Table of Contents

List of figures and tables

List of Acronyms and Abbreviations

AC	Alternating current
BGS	British Geological Survey
BIS	Department for Business, Innovation and Skills
CCC	Committee on Climate Change
CCS	Carbon Capture and Storage
CESP	Community Energy Saving Programme
CIC	Community Interest Company
COP	Conference of Parties
CST	Council for Science and Technology
DC	Direct current
DECC	Department for Energy and Climate Change
Defra	Department for the Environment, Food and Rural Affairs
EC	European Commission
ECO	Energy Company Obligation
ECR	Environmental Country Risk
EPI	Environmental Performance Index
ESCo	Energy Service Company
EU	European Union
EV	Electric Vehicles
FiTs	Feed in Tariffs
GHG	Greenhouse Gas
HHI	Herfindahl-Hirschman Index
HNDU	Heat Network Delivery Unit
IEA	International Energy Agency
IPPC	Intergovernmental Panel on Climate Change
LCA	Lifecycle Analysis
MFA	Material Flow Accounting
MLP	Multi-Level Perspective
NDPB	Non-departmental Public Body
NGO	Non-Governmental Organisations
OECD	Organisation for Economic Co-operation and Development
Ofgem	Office of Gas and Electricity Markets
Ofwat	Water Services Regulatory Authority

PMG	Permanent magnets
PMPP	Pro-Market Policy Paradigm
PV	Photovoltaic
R&D	Research and Development
RCEP	Royal Commission on Environmental Pollution
REDD	Reducing Emissions from Deforestation and Forest Degradation
REE	Rare Earth Elements
RSAP	Resource Security Action Plan
RSP	Regulatory State Paradigm
RWA	Regional Water Authority
SA	Social Accounting
SES	Socio-Ecological Systems
SNM	Strategic Niche Management
SSA	Supplier Services Agreement
TAR	Third Assessment Report (of the IPPC)
TCRE	Transient climate response to cumulative carbon emissions
TM	Transitions Management
TPLS	Third Party Licenced Supplier
UKERC	UK Energy Research Centre
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
WGI	Worldwide Governance Indicator
WSL	Water Supply Licencing
ZCB	Zero Carbon Britain

1 Introduction: the challenge of low-carbon infrastructure transitions in the UK

Avoiding dangerous climate change requires radical reductions in greenhouse gas emissions (IPCC 2014). Infrastructure operation (for example the provision of energy or the treatment of water and waste water) is one of the biggest sources of anthropogenic greenhouse gas (GHG) emissions in the UK and must contribute to the necessary emissions reductions (HM Government 2011a). However, there is increasing agreement that the UK's ageing infrastructure assets (*"the physical assets underpinning the UK's networks for transport, energy generation and distribution, electronic communications, solid waste management, water distribution and waste water treatment*" (ICE 2009)) are unable to reduce emissions generated during operation sufficiently to become 'low carbon' and respond to this pressing need (CST 2009a; Hall, Henriques, Hickford, et al. 2012). As such, there is an imperative to transform our infrastructure systems towards alternatives that deliver services reliably but within environmental limits (Rockström et al. 2009). At the same time, infrastructure has a crucial role in supporting economic growth and improving quality of life (Infrastructure UK 2010) and must continue to do this, all in the face of a growing population and increasingly scarce resources. There are significant challenges to developing a coherent response to these competing drivers and achieving the necessary transformative change.

When defining the infrastructure system, this thesis considers not just the physical assets, such as water treatment works, power plants and electricity networks, but also the processes, actors and institutions necessary to co-ordinate the operation and maintenance of these assets and deliver economic growth and improve quality of life. To this end, the definition developed by the iBuild consortium provides most satisfactory: "*infrastructure [is] the artefacts and processes of the inter-related systems that enable the movement of resources in order to provide the services that mediate (and ideally enhance) security, health, economic growth and quality of life at a range of scales"* (Dawson 2013).

This thesis aims to inform the debate around low-carbon infrastructure transition by improving our understanding of the challenge of transformative change and in particular of how infrastructure system change is constrained. Furthermore, it aims to use this understanding to determine how policy makers and regulators might respond to alleviate these constraints and accelerate change. Infrastructure systems are inherently complex and this thesis aims to strikes a balance between recognising this complexity and fully characterising it. It is important to understand complexity to a sufficient degree to avoid unintended consequences or ineffective responses but this analysis should be as efficient as possible and not at the cost of action.

The thesis focuses on two contrasting challenges to low-carbon infrastructure transition; one principally physical and one institutional. The physical challenge is related to the availability of natural resources required to support energy system transition. Limits to the availability of resources are often conceived as only a driver of transition but there is increasing awareness that disruption in availability of particular resources, such as so-called critical materials, might constrain low-carbon infrastructure transition (US Department of Energy 2011; Moss et al. 2011). The institutional challenge is related to the suitability of institutional systems and actors to deliver the necessary change. A particular focus is placed on the constraint from current approaches to infrastructure policy and regulation. These constraints are chosen because they are under-studied in the literature but also because of the dichotomy between physical and institutional systems and in particular the approaches used for their analysis.

This thesis focusses on the transition necessary in two infrastructure systems; energy and water (including water supply and sewerage), two extremely complex infrastructure sectors, in terms of technical, institutional and social organisation. The decision to study two sectors was taken because of the many interdependencies between these sectors; for example, the electricity needed to pump water through the water network, the need for both water and electricity or gas to provide thermal comfort in the home (Roelich et al. 2014). The nature of one system might enable or constrain change in the other. For example; the tight regulation of investment in the water sector limits its potential contribution to low-carbon energy generation (Hall, Henriques, Hickford, et al. 2012). The exploitation of this interconnectedness to deliver cost and resource efficiencies is restricted by separate regulation of infrastructure sectors (Roelich et al. 2014), which could affect the scale and speed of a low-carbon infrastructure transition (Hall, Henriques, Hickford, et al. 2012). It is considered that this constraint could be significant and, therefore two infrastructure systems are analysed to address the policy and regulatory constraints caused by separate regulation of infrastructure systems.

The remainder of this chapter briefly introduces the principal drivers of a low-carbon infrastructure transition, namely the pressing need for climate change mitigation and the inability of current infrastructure to deliver this mitigation. It goes on to discuss the nature of a low-carbon transition and highlights the limited focus on challenges that might constrain this transition. It concludes with the research questions that motivate this thesis and address the research gaps identified in the following sections.

1.1 Climate change and the low-carbon agenda

1.1.1 Global climate change

The Intergovernmental Panel on Climate Change (IPCC) has shown that warming of the climate system is unequivocal and that this has resulted in reduced quantities on snow and ice, sea level rise and an increase in extreme weather events since 1850 (IPCC 2013). This warming is created by a net positive uptake of energy from positive radiative forcing, the largest contribution of which is an increase in atmospheric concentration of GHGs. The majority of these GHGs are anthropogenic and the IPCC concludes that it is extremely likely that human influence is the dominant cause of the observed warming since the mid-20th century (IPCC 2013). Continued emissions of GHGs will result in further warming and changes to all aspects of the climate system.

In recognition of this, there is a global objective of limiting the increase in global average surface temperatures to below 2^{o}C, with the aim of increasing the chances of avoiding dangerous climate change (IPCC 2007).The radiative forcing which drives climate system warming is defined by cumulative emissions; even if GHG emissions stopped now, most aspects of climate change will persist for many centuries (IPCC 2013). The principal challenge is to reduce the maximum cumulative emissions so that warming stays below harmful levels.

1.1.2 Global response to climate change

The IPCC is categorical about the urgency and scale of emissions reductions necessary to avoid irreversible climate change, yet a global agreement which reflects the magnitude of the challenge remains elusive. The global climate change regime is founded on two legally binding treaties; the 1992 United Nations Framework Convention on Climate Change (UNFCCC) (UNFCCC 1992) and the 1997 Kyoto Protocol (UNFCCC 1998). The Kyoto protocol includes individual legally binding targets for signatories listed in Annex I amounting to a collective reduction of GHG emissions by 5% from 1990 levels by 2012 and a reduction of 18% by 2020. Recent negotiations by the Conference of Parties (COP), the main political decision making body associated with the UNFCCC, have attempted to define a post-Kyoto regime, but with little success to date.

Perhaps the most significant agreement post Kyoto is the Copenhagen Accord, which in 2009 committed signatories to "…hold the increase in global average temperature below 2°C above pre-industrial levels…" (UNFCCC 2011). Signatories to the subsequent Cancun Agreement include countries that are not subject to legally binding targets under the Kyoto Protocol. Some of these countries have made pledges to address emissions; however, these pledges are voluntary and are based on emissions intensity or minor deviations from business as usual, not on absolute emissions reductions. A recent United Nations Environment Programme (UNEP) report showed that the emission commitments pledged under the Copenhagen Accord/Cancun Agreement and the second period of the Kyoto Protocol amount to only around 60% of what is needed to avoid a greater than 2°C rise in global surface temperatures (UNEP 2010).

1.1.3 The UK's response to global climate change

In parallel to the ongoing global negotiations, the UK Government set itself binding emissions reductions targets with the aim of reducing UK GHG emissions to 80% of 1990 levels by 2050 (HM Government 2008a). A target of 60% had been adopted in the 2003 Energy White Paper (DTI 2003) which was based on an earlier report by the Royal Commission on Environmental Pollution (RCEP) (RCEP 2000). The RCEP target was based on a global quota of emissions in 2050 necessary to avoid an atmospheric concentration of greater than 550pm, which was deemed at the time to be sufficient to avoid dangerous climate change. This UK's share of this quota was based on the principle of contraction and convergence and allocated on a per capita basis (RCEP 2000).

A subsequent review in the first report of the Committee on Climate Change (CCC) recommended tightening the target to 80% as a result of developments in climate science, which showed that the danger of significant climate change was greater than previously assessed (CCC 2008). The revised target, which was still based on an acceptable global quota and was allocated to the UK on a per-capita basis, was adopted as a legal obligation in the Climate Change Act 2008 (HM Government 2008a). The UK's emissions reductions target has been heralded as ambitious and world-leading; however, there is some debate as to whether a quota-based target is sufficient in light of the fact that it is the *cumulative* quantity of emissions which defines climate change. It is possible that this budget is not stringent enough and that more ambitious action is required (Anderson & Bows, 2012).

The Department for Energy and Climate Change (DECC) was formed in 2008 to develop policy necessary to achieve this target and the CCC was formed in the same year to monitor and report on progress. This instigated a rush of policies with the aim of decarbonising the UK economy, and a high level strategy in the form of the Low Carbon Transition Plan (HM

Government 2009). The subsequent Carbon Plan provided a more detailed plan including definition of a series of emissions budgets in five year periods (HM Government 2011a).

The UK is on target to outperform the first two budgets, but this achievement is principally due to a slow-down in economic growth and a steep increase in fuel prices. Reductions due to low-carbon measures were less than 1 per cent in 2011 but would need to increase to 3 per cent annually to meet future budgets (CCC 2013). This presents a real challenge if economic growth returns to pre-recession rates and if fuel prices level out.

1.2 Infrastructure and the low-carbon agenda

1.2.1 Infrastructure's contribution to GHG emissions

One of the biggest sources of anthropogenic GHG emissions is the operation of infrastructure; more than half of the UK's territorial GHG emissions originate from energy supply and transport (see figure 1.1). The water industry contributes around 1 per cent[1]; however, its contribution is growing as a result of increasing demand for water and more stringent quality obligations on water and wastewater treatment that require increasingly energy-intensive treatment technologies (Water UK 2009). Furthermore, water management itself is an important issue, particularly as climate change already set in motion affects the distribution of water causing increased flooding and drought (IPCC 2013).

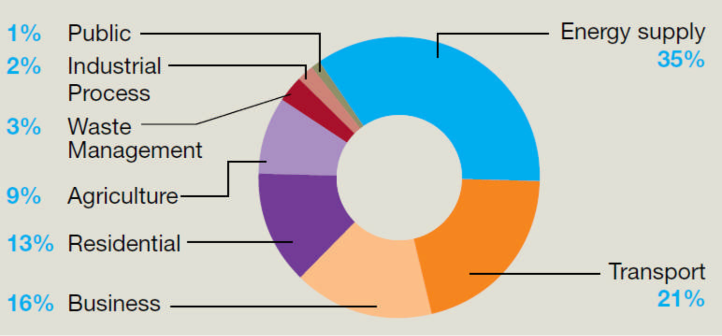


Figure 1.1: Total UK Greenhouse Gas Emissions by sector, 2011 (excluding Land Use, Land Use Change and Forestry) Source: UK GHG Inventory (UNFCCC coverage) (Source: DECC 2013b)

Infrastructure, as one of the principal sources of GHG emissions, must make a significant contribution to the necessary acceleration of action. Decarbonisation of the UK's energy infrastructure will be one of the primary mechanisms for achieving the UK's emissions

---

[1] This is not reported by DECC as an individual sector, but is distributed between waste management and industrial processes.

reduction target, which places infrastructure at the centre of the low-carbon agenda (HM Government 2011a).

Emissions reduction of the magnitude required to achieve UK government targets will require rapid, systemic change to physical infrastructure, especially energy systems. This will require a step-change in both the scale and rate of the roll out of low-carbon energy supply technologies (HM Government 2011a). However, this needs to be accompanied by a transformation in end-use technologies to dramatically reduce demand for energy. Demand reduction is critical if the UK is to achieve the necessary scale of emissions reductions and minimise the cost of technology replacement (Wilson et al. 2012). A similar transformation of water provision and demand needs to occur to further reduce emissions and avoid seasonal water scarcity (HM Government 2011b).

1.2.2 The nature of infrastructure in the UK

The definition of infrastructure used in this thesis is helpful because it not only defines the physical systems (the means) but also highlights the purpose of infrastructure (the end). It reinforces the fact that infrastructure underpins the functioning of society and the economy. In mature economies, such as the UK, infrastructure investment has been shown to have a positive effect on productivity of the economy and employment (Zegeye 2000; Égert et al. 2009). It could be argued that infrastructure plays an even more important role in developing countries; access to energy is considered to be essential to achieving all of the Millennium Development Goals (Wilkinson et al. 2007) and access to water and sanitation is now recognised by the United Nations as a human right (United Nations General Assembly 2010).

Despite its importance to the economy and wellbeing, infrastructure in the UK is ageing; a considerable proportion of the infrastructure stock was built in the 19th Century (Infrastructure UK 2010). A combination of obsolescence and deterioration means that much of the UK's infrastructure is close to systemic failure. However, reinforcing or replacing the current physical infrastructure like for like will not suffice if the UK is to contribute to decarbonisation at the necessary rate and scale (Hall, Henriques, Hickford, et al. 2012). Furthermore, infrastructure must continue to provide reliably the services essential to human development, wellbeing and economic growth throughout any reconfiguration. This thesis focuses on climate change mitigation as a driver for infrastructure transformation but recognises that there are competing drivers which make the task of transforming our physical infrastructure more complex.

It is important to note that the influence of infrastructure on society is not just related to the physical assets; physical infrastructure also shapes the institutional and social organisation of a society, through a historical process of change and evolution described as “co-evolution” (Foxon, 2011). This implies that changing physical infrastructure necessarily involves larger social and institutional shifts as well as technical improvements, which makes system change even more challenging. Physical transformation needs to be accompanied by change in the way that infrastructure is regulated, used and operated.

1.2.3 Infrastructure operation

In the UK, there has been a shift from state controlled infrastructure systems towards liberalisation, private provision and competition (Hall et al 2012). Privatisation in the 1980s was driven by a strong ideology within the Conservative government of the time that state control of infrastructure was inefficient and undesirable (Roelich et al. 2014). It was considered that ‘the market’ was the best place to make decisions on infrastructure investment and operation.

Despite regulation to disband monopolies and create a ‘free’ infrastructure market, the market is dominated by a small number of large, international utility companies, for example 98% of household gas and electricity is supplied by only six energy companies (Ofgem 2013a). Whilst there has been investment in recent years, this is ad hoc and mainly replaces or in some cases renews existing infrastructure; very little is spent on modernising infrastructure (CST 2009a). The operation of both energy and water infrastructure is characterised by a throughput-based model, whereby greater profit is made by increasing the number of units of utility product sold (for example kWh electricity), or by increasing the marginal cost efficiency of producing a unit of energy. This mode of operation, in combination with the prioritisation of short-term income, discourages demand reduction (which reduces throughput) and investment in low-carbon technologies (which increase marginal costs in the short-term) (Roelich et al. 2014).

1.3 Low-carbon infrastructure transition in the UK

The remaining budget for GHG emissions necessary to avoid dangerous climate change is rapidly depleting but to date, action to reduce emissions has been incremental and slow (Anderson and Bows 2012). There is a narrowing window of opportunity and a rapid and wholesale transformation of infrastructure systems is needed to reduce the UK’s contribution to global emissions (Hall et al 2012). The throughput-based modes of operation that dominate the water and energy sector are not likely to deliver demand management on the necessary scale and are also resistant to the scale of technology renewal required (Steinberger et al.

2009). Therefore; technology transformation must be accompanied by a transformation in the operation of infrastructure to incentivise demand reduction and encourage low-carbon technology uptake. These two strategies for transforming UK infrastructure are described in more detail below.

1.3.1 Low-carbon technologies

Technology change is needed to decarbonise the UK's energy supply and to reduce the quantity of energy and water needed to deliver the ultimate needs of the end user, such as thermal comfort, which is often referred to as the energy or infrastructure service. The UK's current energy system is dominated by high emissions technologies, such as coal and gas power stations and internal combustion engine driven vehicles. There is an urgent need to replace these technologies with low-carbon technologies, including renewable energy supply (such as wind turbines, solar panels) and demand-side technologies (such as hybrid and electric vehicles, insulation and energy efficient devices) which reduce demand for energy (HM Government 2011a; Ekins et al. 2013). The scale of change in technology necessary is unprecedented and is illustrated for the UK in Figure 1.2.

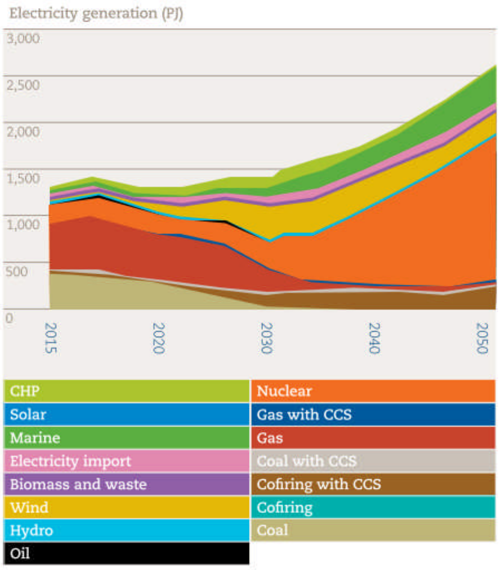


Figure 1.2: Electricity generation technology mix 2015-2050 (Data from: AEA Technology 2011; Figure sourced from: Ekins et al. 2013)

This scenario was produced in support of the government's Carbon Plan and demonstrates a change in the character of energy supply from centralised, fossil fuel-based generation to more intermittent and decentralised supply. This has implications for the transmission and distribution networks linking generation technologies to end users. The scenario is underpinned by technology and behaviour change at the end user-level that resulted in a decrease in average demand of 15-20% (AEA Technology 2011). This scale of reduction is unprecedented but is essential to achieve decarbonisation in the modelling.

1.3.2 Alternative modes of operation

Some of the greatest challenges associated with achieving the scale of technology change or demand management required to deliver emissions reductions are associated with the mis-match between the dominant mode of operation and the aims of technology change and demand management. There is a need for alternatives modes of operation to overcome these challenges (Saunders et al. 2012; IBM 2010; IEA 2010). This may include operation by different organisations, whose aims are broader than short-term profit maximisation and prioritise low-carbon transformations (Seyfang & Haxeltine 2012; G. Walker 2011; Bale et al. 2012). Or it might include business models whereby income is derived by alternative mechanisms such as selling energy saving or infrastructure services (Sorrell 2005; Hannon et al. 2013). These two dimensions of alternative modes of operation are illustrated in figure 1.3 and described in more detail below.

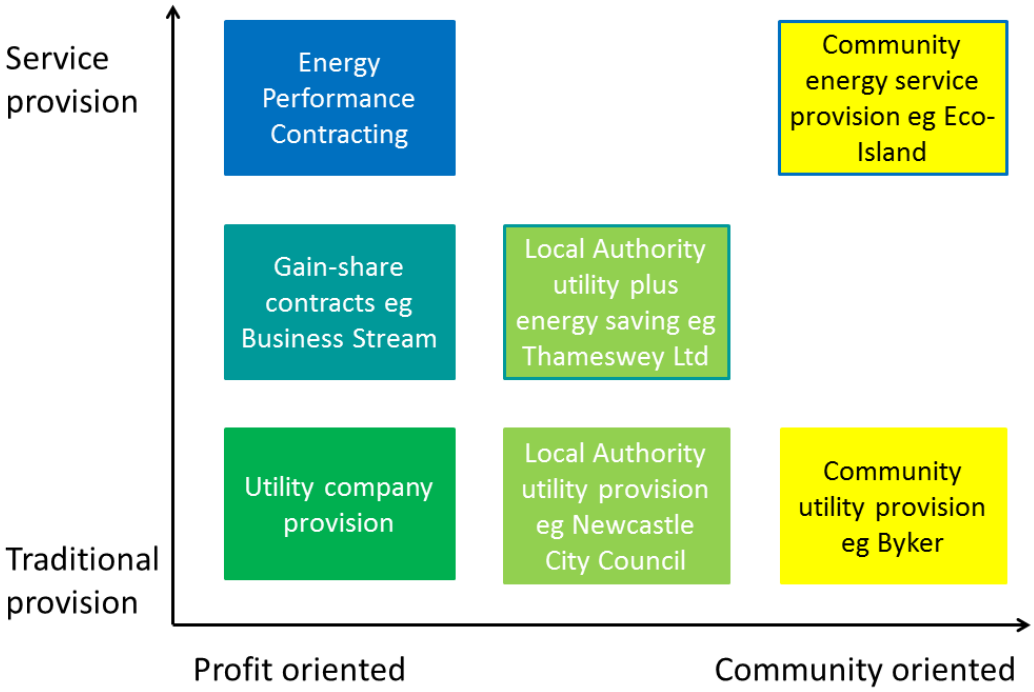


Figure 1.3: Alternative modes of operation of infrastructure

There are numerous recent examples of infrastructure provision by organisations that are not the historically dominant large, international utility companies. This includes commercial organisations, such as technology providers and independent energy generators, who are diversifying their operations (Hannon et al. 2013); local authorities, who are increasingly developing and operating local infrastructure systems (Hawkey et al. 2013; Bale et al. 2012); and individuals and communities, who are increasingly taking control of their infrastructure provision (Seyfang & Haxeltine 2012; Walker & Devine-Wright 2008). The motivations of these organisations vary dramatically but include a broader set of aims and drivers, such as GHG mitigation, attracting inward investment, self-sufficiency and fuel poverty reduction.

In parallel to increasing diversity in providers, there is an increasing diversity in the mechanisms of provision and the means by which profit is generated. Rather than revenue being coupled with sales of energy or water, there are increasing examples of organisations generating revenue from guaranteeing savings or from providing infrastructure services. This includes modes of operation such as; Energy Supply Contracting, where local providers supply useful energy streams such as heating, cooling and electricity; revenue generation from demand management, rather than supply, particularly in the energy sector but increasingly in the water sector (Business Stream 2013); and performance contracting, where the contract is for an agreed level of service, such as space heating, space lighting or motive power (Sorrell 2007; Steinberger et al. 2009).

1.4 Challenges to infrastructure transition: constraints

The necessary transition in both infrastructure technology and mode of operation represents a step change from the status quo. This systemic transformation will be far from straightforward; in particular, because the direction of transition is necessarily predetermined (towards sustainability) and the timescales are urgent (Shove & Walker 2007). There are no precedents for a transition of this nature and we are likely to face significant challenges which will affect desirable change in technological, environmental, institutional and economic systems. The focus of this thesis on physical and institutional challenges to this transition has been adopted to redress the techno-economic bias of current policy, and to some extent research, on infrastructure transition. This techno-economic perspective frames infrastructure transition challenges in relation to the cost and effectiveness of particular technologies and thus ignores some potentially crucial issues that could affect transition dynamics.

Within these broad challenges there are many specific constraints, which could curtail the necessary scale or speed of transition. These constraints include; cost of technology (HM Government 2011a; Ekins et al. 2013); lack of agency (Shove & Walker 2007; Smith et al. 2005); political ideology (Kern 2011; Mitchell & Woodman 2010); lack of agreement on goals or vision (Stirling 2009); resource availability (Andersson 2001a; US Department of Energy 2011); unsupportive and fragmented regulation (CST 2009b; Mitchell 2006); lack of community engagement or empowerment (Brown & Farrelly 2009); vested interests (Moe 2010); and rigid organisational cultures (Farrelly & Brown 2011).

It is interesting to note that the majority of these constraints are socio-institutional; they derive from a lack of support from society, organisations or government agencies or a lack of appropriate institutions to govern change. Despite this, the majority of scenarios of infrastructure transition consider only the cost and technology constraints of transitions (Ekins et al. 2013; HM Government 2011a). This thesis explores two significant, but under-studied, constraints to one of the aspects of alternatives to the current infrastructure system described in section 1.3:

- The constraint to technology change posed by the disruption in supply of critical materials embedded in low-carbon technologies. Critical materials are those which are subject to disruption in supply and difficult to substitute (European Commission 2010b). A number of key renewable energy technologies rely on critical metals; for example, neodymium in wind turbines and indium in thin film solar panels (US Department of Energy 2011; Moss et al. 2013). There has been a great deal of interest in identifying which materials are critical but less in determining the effects of supply disruption on scenarios of technology roll-out.
- The constraint to change in the mode of infrastructure operation from current policy and regulation in the UK. The current regulatory system has evolved around the dominant, centralised, throughput-based mode of operation and actively excludes alternative providers and modes of operation (Mitchell & Woodman 2010). There has been a great deal of interest in the barriers that policy and regulation creates for technological change but less in creating systems of governance which support a diversity of operators and modes of operation.

1.5 Challenges to infrastructure transition: Potential responses

Infrastructure is a complex, interconnected system of technology embedded in society and the environment, interacting with public and private institutions (Roelich et al. 2014). Furthermore, the mechanisms that constrain system change act in multiple parts of the system so isolated policy responses are unlikely to be effective. Identifying appropriate instruments to intervene and effect desirable change can be difficult and the response to these interventions is uncertain. Systemic action is required that cuts across policy sectors, infrastructure systems and public and private institutional boundaries (Leach et al. 2010). This needs to be supported by an understanding of uncertainty and the complex nature of system change (Foxon 2011).

Recent research on infrastructure is dominated by that which aims to advance understanding of infrastructure systems and, in particular the interdependencies between systems. Increasingly sophisticated models of infrastructure systems are being created to simulate the complexity of real systems to support this aim (Mortimer 2012). This is supported by work to characterise and reduce the uncertainty associated with system simulation (Usher & Strachan 2013; Hughes et al. 2012). However, there must be a trade-off between the detailed characterisation of complex systems and the imperative to initiate change as a matter of urgency. The balance between system understanding and action is fundamental to this thesis. Therefore, the analysis undertaken in this thesis includes a specific focus on how policy makers and regulators might respond to mitigate the effects of the constraints and accelerate a low-carbon transition. This is to ensure that policy relevance is not secondary to understanding transition and to force the balance between complexity and action.

1.6 Research questions

In response to the gaps identified in the literature presented, the principal research question of this thesis is: *how can constraints to low-carbon infrastructure transition be characterised with sufficient detail to enable action to mitigate disruption?*

A series of sub-research questions have been devised to structure the investigation of this research question for the selected constraints:

RQ1. How can we conceptualise the constraints of critical material supply disruption and infrastructure policy and regulation to low-carbon infrastructure transitions?

The transparent conceptualisation of constraints is essential in framing analysis and producing defendable outcomes. It can also have a significant influence over the methods used for analysis and on the type of knowledge that this analysis produces. Furthermore, there is a dichotomy in the way that physical and institutional constraints are conceptualised and this thesis explores the effect that this dichotomy has on the usefulness of resulting analysis.

RQ2. What methods are most appropriate to analyse constraints from critical material supply disruption and infrastructure policy and regulation to low-carbon infrastructure transitions?

The two constraints differ greatly in character and thus require different approaches to analysis of the nature and scale of each constraint. This thesis will consider approaches and methods that might be most effective in exploring and enumerating the two constraints.

RQ3. What is the nature and scale of the constraints from critical material supply disruption and infrastructure policy and regulation?

The characterisation of constraints, using different methods, provides different information about the mechanisms by which low-carbon transitions could be constrained and the extent and severity of these mechanisms of constraint.

RQ4. How can we use insights and knowledge from constraint analysis to identify potential policy responses?

One of the core purposes of this thesis is to identify how policy makers might respond to the constraints identified by analysis responding to RQ3. The depth and variety of responses will vary depending on the depth of understanding of the system.

RQ5. How can constraint analysis be used to improve our understanding of low-carbon infrastructure transitions and the role of governance in accelerating this transition?

This overarching research question aims to identify how we can best balance understanding low-carbon transitions with understanding constraints and responding to them.

1.7 Outline of thesis

This thesis is structured to address the research questions in the following way (shown in figure 1.4). Chapter 2 clarifies the definition of low-carbon, introduces low-carbon infrastructure scenarios in the UK then goes on to describe how low-carbon transitions have been analysed in the literature and how governance might help accelerate future transitions to support policy analysis. It provides context to the ensuing chapters and clarifies the contribution of this thesis to a number of literatures. Chapter 3 describes how the two constraints, which are the subject of this thesis, have been conceptualised in the literature and, as such, provides context for the response to RQ1 in the subsequent chapter. Chapter 4 addresses both RQ1 and RQ2 and sets out the scope and conceptualisation of the two constraints. It goes on to describe the methods used to analyse the individual constraints and the limitations of these approaches.

After chapter 4 the thesis divides and presents results for each constraint in turn: Chapter 5 presents the results of analysis of constraints to low-carbon transition from critical material supply disruption (addressing RQ3 for this constraint) and chapter 6 presents potential responses (addressing RQ4 for material criticality constraints). Chapter 7 provides additional description of the mainstream mode of operation and the characteristics of case studies to provide context for subsequent analysis. Chapter 8 presents the results of analysis of constraints to alternative modes of operation transition from current infrastructure policy and regulation (addressing RQ3 for this constraint). The final results chapter, 9, discusses potential responses to remove policy and regulation constraints, responding to RQ4 for this constraint. It goes on to describe a more systemic change in policy and regulation to enable broader transition.

The two strands of the thesis come back together in chapter 10, which discusses the relative merits of both approaches and how they can be used to improve our understanding of low-carbon transitions and the role of governance, which contributes to both RQ2 and RQ5. The thesis concludes with a summary of the main findings and recommendations for future work.

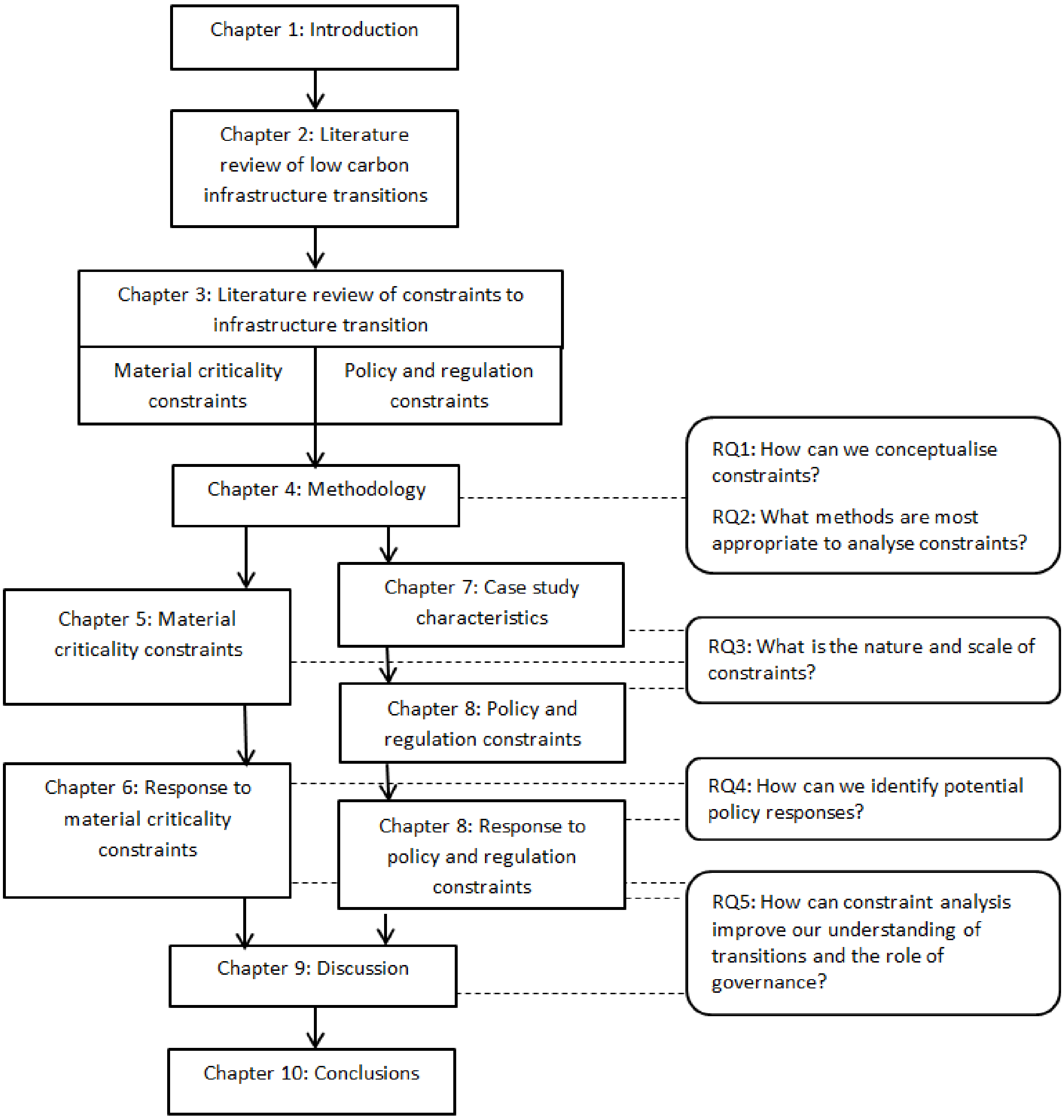


Figure 1.4: Thesis structure and connection to research questions

2 Low-carbon infrastructure transitions in the UK

Chapter 1 identified that a systemic transformation of infrastructure was needed in the UK to contribute to its low-carbon goals but that there were significant constraints to this transformation. This thesis aims to investigate how these constraints can be characterised to the extent necessary to enable mitigation action. In order to do this it is first necessary to analyse the transition that could be constrained. Therefore; this chapter reviews approaches used by researchers and policy makers to understand past and future transitions, which will inform the methods used in this thesis. An important focus of this thesis is on the response to constraints so this chapter also reviews approaches to governance of transitions to identify principles and approaches that might inform action to remove constraints. The chapter begins with a discussion of the assumptions that underpin interpretations of 'low-carbon' and a summary of recent scenarios of low-carbon infrastructure in the UK to provide context for the scale of transition necessary.

2.1 What is low carbon?

The principal driver for the low-carbon agenda is the connection between GHG emissions (of which carbon dioxide is the principal constituent in terms of volume), climate change and the associated consequences of a changing climate. Low-carbon policy and strategy has come to the fore in the last two decades as a result of increasing scientific consensus that reducing future anthropogenic GHG emissions (a low-carbon future) will reduce the scale of climate change and avoid dangerous negative consequences (IPCC 2014; IPCC 2013). The relationships between these factors are complex and it is beyond the scope of this thesis to discuss the detailed climate science that underpins them. However, there are some pertinent aspects of the GHG-climate change dynamic that are relevant to understanding the nature of low-carbon and the necessary scale of GHG emissions reductions.

The first is the consequences of climate change that could be deemed to be dangerous. In its third assessment report (TAR) the IPCC identified "reasons for concern" which are the most significant consequences of a changing climate (IPCC 2001). These include risks to biodiversity, people, markets; the risks of extreme weather; the distribution of these risks; and the risk of sudden changes. The second aspect is the scale of change in global mean temperature that would cause unacceptable risks (dangerous climate change). The so-called "burning embers diagram" was produced in the TAR to show the risks from a changing climate plotted against increases in global mean temperature (°C) after 1990 (left hand side of figure 2.1). Each column corresponds to a specific reason for concern and the colour of each column represents the increasing levels of risk at increasing global mean temperature. There is no threshold which represents dangerous climate change; the authors considered this to be a value judgement. The TAR highlighted that anything above 2°C would cause a step-change in the negative consequences of climate change and there was a broad scientific and political consensus that 2°C was the threshold above which climate change would be dangerous (Group of Eight (G8) 2009). Avoidance of more than 2°C increase in global mean temperature was adopted as the long-term goal for international negotiations in the Copenhagen Accord and subsequent Cancun Agreement (UNFCCC 2011).

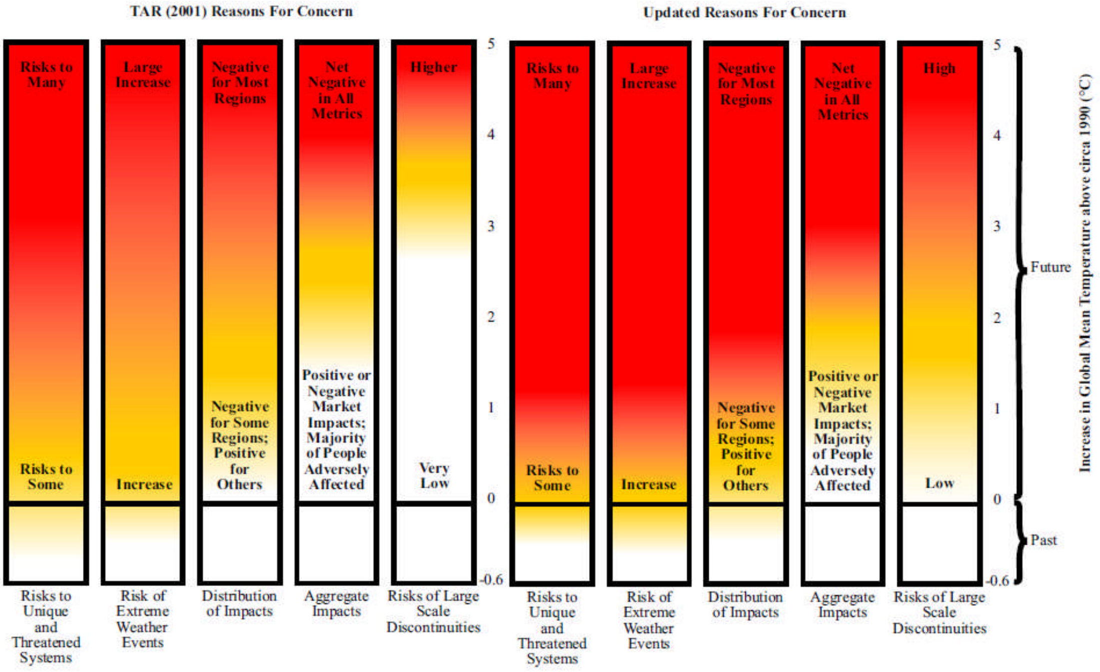


Figure 2.1: Risks from climate change, by reason for concern—2001 Third Assessment Report (TAR) compared with updated data (Source: Smith et al. 2009).

The burning embers diagram was updated in 2009, in preparation for the IPCC's fourth assessment report (right hand side of figure 2.1). The updated analysis showed that the risks of negative consequences in all reasons for concern were more severe and occurred at lower increases in global mean temperature than had been accepted previously (Smith et al. 2009). This brought into question whether a 2°C threshold was sufficiently low to avoid dangerous climate change. Despite this, it remains widely accepted as the threshold above which climate risks will be unacceptable.

Increase in global mean temperature is well connected to climate change risks but the goal of 2°C doesn't help to identify what scale of emissions reductions is necessary to avoid this increase i.e. how 'low carbon' does low carbon need to be? The rate and magnitude of global climate change is determined by radiative forcing, climate feedbacks and the storage of energy by the climate system (IPCC 2013). Radiative forcing is the change in energy fluxes (into and out of the climate system) caused by substances and processes. Emission of GHG is one of the most significant contributors to positive radiative forcing and off-sets negative (cooling) radiative forcing. This has resulted in an historic increase in uptake of energy and an associated rise in global mean temperature of 0.6^oC since the baseline of climate models (1861-1880) (IPCC 2013).

The effect of GHG-related radiative forcing and the response of the climate system to carbon emissions can be quantified as a function of cumulative carbon emissions. This is termed the transient climate response to cumulative carbon emissions (TCRE) and is defined as the global mean surface temperature change per 1000GtC emitted to the atmosphere. TCRE is likely to be in the range of 0.8°C to 2.5°C per 1000 GtC (IPCC 2013). Cumulative total emissions of CO_2 (one of the principal GHGs by volume) and global mean surface temperature response are approximately linearly related; therefore it is possible to define a cumulative emissions budget from the baseline of climate models (1861–1880). The uncertainty associated with TCRE means that particular budgets are determined to have a probability of limiting the warming caused by anthropogenic CO_2 emissions to less that 2^oC since model baselines. The probabilities and associated budgets are presented in table 2.1.

Table 2.1: Cumulative emissions budgets based on different probabilities of limiting warming to less than 2^oC (IPCC 2013)

Probability of avoiding >2^oC warming	Cumulative emissions budget (GtC)	Budget accounting for non-CO_2 forcings (GtC)	Remaining budget accounting for emissions 1861-2011 (GtC)
>33%	1560	880	349
>50%	1210	840	309
>66%	1000	800	269

These budgets are based solely on CO_2 and will be subject to the effects of radiative forcing from non-CO_2 causes. When these forcings are taken into account the budget amounts are reduced significantly to between 800 and 880 GtC. It is estimated that 531 ±85 GtC was emitted between 1861 and 2011, therefore, the remaining global budget for CO_2 emissions is between 269 and 348 GtC, depending on the acceptable risk of missing the 2^oC target. Including a 10% contribution from land-use change, we are currently emitting between 10-11

GtC annually. Even if global annual emissions remain at current levels the cumulative budget could be exceeded by 2035. If they continue to rise at an average of 2.7 per cent annually, as was the case for the decade to 2011 (Olivier et al. 2012), the budget will be exceeded considerably earlier.

Despite the scientific consensus around cumulative emissions budgets, there are no examples of regional or national targets based on this premise. Those targets that do exist are based on reduction from baseline by a fixed point in time. The trajectory between these two points (and the associated cumulative emissions) is rarely considered. However the CCC has developed five-year cumulative emissions budgets up to 2027 to monitor progress toward the UKs 2050 target (CCC 2010). Furthermore, there is little consensus about how to distribute a global emissions budget amongst regions and nations. If each country has to reduce its emissions by the same proportion this would ignore the historical responsibility of industrialised nations for the cumulative emissions released to date and also the fact that some countries have developed greater wealth and development as a result of those historical emissions (Baer et al. 2008). Therefore, some argue for a differentiated budget that recognises that the least developed countries should be allowed a greater share of the budget to enable development and that those who had already achieved high levels of development should take a greater share of the responsibility for reduction (Bows & Anderson 2008; Baer et al. 2008; Chakravarty et al. 2009).

The complexity and uncertainty associated with quantification of the necessary rate and scale of emissions reductions mean that a series of assumptions underpin any definition of low carbon. The lack of transparency and agreement over these assumptions has resulted in a great deal of contestation over the definition of low carbon.

In the UK, the focus of this thesis, the government defines low carbon as activities that contribute to an 80% reduction in the UK’s greenhouse gas emissions by 2050 (HM Government 2009). But even this simple definition has caveats. The Low Carbon Transition Plan states that getting to a low-carbon UK would involve cutting emissions, maintaining secure energy supplies, maximising economic opportunities, and protecting the most vulnerable (HM Government 2009). In common with the majority of targets the UK’s target is quota-based, rather than cumulative budget-based. The quota is based on the principle of contraction and convergence whereby the global quota of annual emissions in 2050 is shared out on a per capita basis[2] and countries must reduce (contract) their emissions to ensure convergence at this future, more equitable distribution (RCEP 2000; CCC 2008).

2.2 Low-carbon scenarios in the UK

Definition of low carbon is contentious; however further challenges are presented by the identification of activities that will most effectively move the UK towards a low-carbon future whilst meeting the other conditions of a low-carbon transition set by the UK government, including energy security and energy poverty reduction (HM Government 2009). Scenarios have been widely used in UK low-carbon policy development to inform and improve decision making in conditions of great uncertainty (Hughes & Strachan 2010). There is a wide range of scenario methodologies but low-carbon scenarios in the UK are dominated by technical feasibility studies of the energy system (Ekins et al. 2013; UKERC & ETI 2010; Skea et al. 2010). These studies demonstrate the technical feasibility of alternative energy systems in meeting energy demands at the same time as meeting a carbon reduction target. Some use energy system modelling techniques, where scenarios are the output of energy system model runs working within an end-point carbon constraint (Hughes & Strachan 2010). Some examples of these principally techno-economic scenarios are presented in sections 2.2.1 to 2.2.3. One example of a descriptive scenario for a low carbon water sector is presented in section 2.2.4, but this is far less detailed than those for energy. Examples are confined to UK scenarios in accordance with the focus of this thesis.

2.2.1 UKERC Energy 2050 Scenarios

Under the UK Energy 2050 project, the UK Energy Research Centre (UKERC) developed a set of scenarios designed to explore the different pathways towards a low-carbon energy system in the UK (Anandarajah et al. 2009). The scenarios were developed using MARKAL, an energy systems model which characterises significant technologies by a series of economic and technical parameters (for example cost, efficiencies and emission factors). The model shows how a given level of energy service demand (for example heat, power or mobility) can be met and by which energy sources. It optimises the energy system to minimise cost or maximise the sum of producer and consumer surpluses, based on the assumptions and data fed into the model and the policy constraints that have been applied to it (Ekins et al. 2013).

A series of scenarios were produced using differing assumptions about application of different policies at different times, such as targets, differing carbon floor prices, renewables obligations, emissions performance standards and feed in tariffs (see table 2.2). The scenarios

---

[2] Based on future projections of population.

present final energy demand by sector and energy generation mix by source in response to these constraints (figure 2.2).

Table 2.2: UKERC Energy 2050 scenarios (Ekins et al. 2013)

Scenario	Scenario name	Annual targets % reduction from 1990 level	Cumulative emissions GTCO$_2$ (2000-2050)	2050 emissions mtCO$_2$
REF	Base reference	-	30.03	583
CFH	Faint-hearted	15% by 2020 40% by 2050	25.67	355
CLC	Low carbon	26% by 2020 80% by 2050	22.46	237
CAM	Ambition (low carbon core)	26% by 2020 80% by 2050	20.39	118
CSAM	Super ambition	32% by 2020 90% by 2050	17.98	59

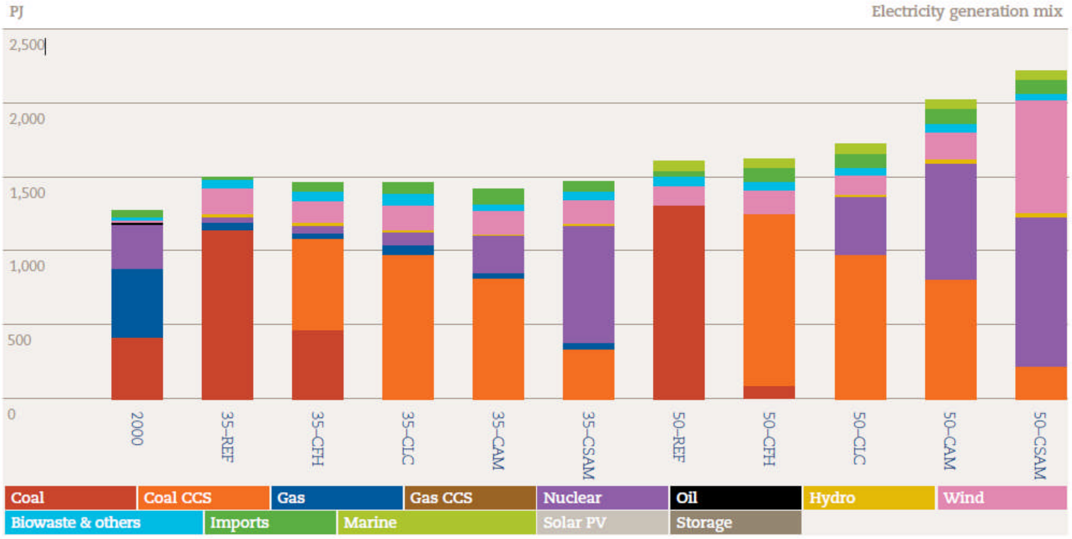


Figure 2.2: Electricity generation mix in 2035 and 2050 in the UKERC Energy 2050 scenarios (Source: Ekins et al. 2013)

Costs of the particular scenarios are presented in terms of the energy system costs and the welfare costs (reduction in the sum of producer and consumer surplus). Overall, the runs with increasingly stringent carbon reduction targets follow similar emission reduction pathways, with additional technologies and measures being required as targets become more stringent, and costs rapidly increase. Updated scenarios were produced in 2013 to reflect new policy, update assumptions about technology and resource costs (Ekins et al. 2013).

2.2.2 The UK Government's Carbon Plan

The UK Government replaced its Low Carbon Transition Plan (HM Government 2009) with the Carbon Plan (HM Government 2011a), which provided greater detail of pathways to achieve its carbon budgets. The Carbon Plan described a number of pathways to achieve a 90% reduction in carbon emissions from 1990 to 2050, including a 'Core MARKAL run' generated using an updated version of the MARKAL model. This Core pathway was used to illustrate the "*technologies likely to contribute to reducing emissions, and the most cost effective timing for their deployment*"(HM Government, 2011: 16). The Core pathway was adapted to explore alternative technology trajectories using the Department for Energy and Climate Change's 2050 Calculator[3] to produce three additional pathways (and see figure 2.3): **Higher renewables, more energy efficiency:** assumes a major reduction in the cost of renewable generation alongside innovations that facilitate a large expansion in electricity storage capacity. **Higher CCS, more bioenergy:** assumes the successful deployment of carbon capture and storage (CCS) technology supported by significant natural gas imports. It also assumes low and plentiful sustainable bioenergy resources. **Higher nuclear, less energy efficiency:** is a future that is more cautious about innovation in newer technologies.

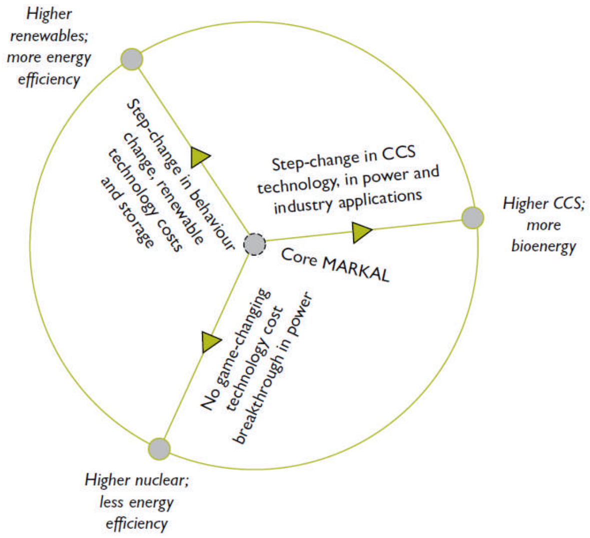


Figure 2.3: DECC's alternative low-carbon pathways (Source: HM Government 2011a)

The pathways describe energy demand by sector and energy mix by source, similar to the UKERC scenarios, but do not include any specific policy measures to achieve this mix, beyond stated ambitions up to 2020. A range of potential costs are presented for the three additional pathways, using estimates included in the 2050 Calculator, however, they are not cost optimised pathways.

---

[3] http://2050-calculator-tool.decc.gov.uk/

The impacts of the pathways on UK growth, security of supply and the environment are briefly discussed in Annex B of the Carbon Plan, but there is no quantification of effects or the potential constraints on any of the pathways.

2.2.3 Centre for Alternative Technology's Zero Carbon Britain

Zero Carbon Britain (ZCB) was developed by the civil society organisation, Centre for Alternative Technology, and describes a scenario in 2030 in which the UK has reduced GHG emissions rapidly to net zero (Centre for Alternative Technology 2013a). The scenario was developed as a communication tool that describes how rapid decarbonisation is possible with current technologies and what it would be like to live in a ZCB. The scenario does not model the transition of the energy system, it rather models the energy flows in one year in the future (Figure 2.4). This was supported by hourly modelling of energy demand and supply, which identified the need for short-term energy storage and back-up generation.

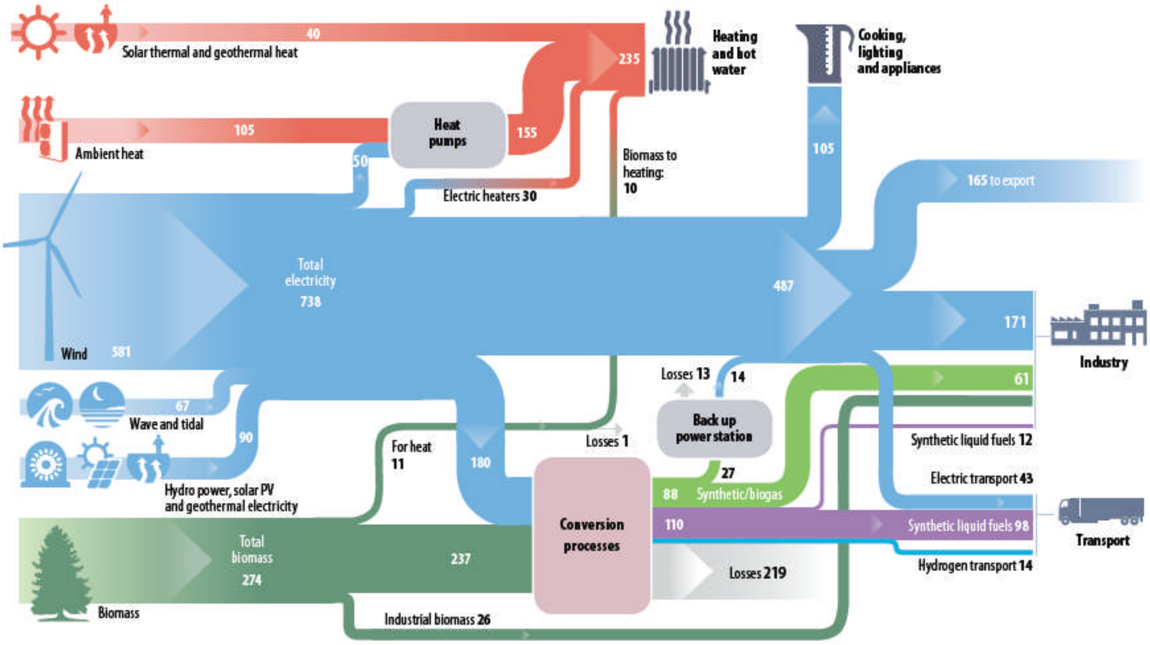


Figure 2.4: Centre for Alternative Technology Zero Carbon Britain energy flows (TWh/yr) (Source: Centre for Alternative Technology 2013)

The scenario is far more radical than the scenarios described above, in that it includes higher emissions reductions in a much shorter space of time. It relies more heavily on demand management and agricultural change than other scenarios, including a 60% reduction in energy demand and significant reductions in livestock farming. The scenario also specifically excludes certain technologies which are prevalent in other scenarios, such as nuclear and carbon capture and storage.

2.2.4 A low carbon water industry in 2050

The Environment Agency commissioned six academics to independently describe their visions for a low-carbon water industry in 2050 (Environment Agency 2009). These visions took the forms of descriptive essays, rather than quantitative analysis but displayed some common opinions about the challenges, opportunities and constraints a low-carbon water industry would face. The authors agreed that a holistic approach was essential with joint working between water companies, customers and other infrastructure providers, particularly energy. Decarbonisation of the energy sector was considered to be fundamental to a low carbon water sector but the authors recognised the potential of the water sector to contribute to low-carbon energy generation and the need for energy demand management in water and wastewater treatment. This may require decentralisation of water provision to enable more sustainable drainage systems and recovery of rain and grey water.

The visions also recognised the need to increase awareness of the value of water and assign greater consumer responsibility for its conservation. Furthermore, it may be necessary to lower the quality of water used for non-potable applications, such as clothes washing, to enable decentralisation and water re-use. This requires a great deal more engagement of end users and an associated cultural change in attitudes to water consumption.

New business models will be required to ensure a focus on long-term strategies that encourage sustainable design and operation. Water tariffs will have to be increased to support investment in new technologies and operational procedures. As a result, regulatory models will have to change to allow the recovery of low-carbon investment and to enable partnerships with other infrastructure sectors.

2.2.5 Limitations of low-carbon scenarios

These scenarios tend to provide a great deal of technological detail and in some cases assessment of economic impacts in terms of welfare losses or additional energy system costs. However, the treatment of socio-political dynamics is usually very limited, beyond simulation of potential (economic) policy (Hughes & Strachan 2010). Some scenarios include a high-level exploration of the potential effect of social objections to energy technologies through the inclusion of ‘hurdle rates’ for different technologies. These are intended to take account of market barriers to the adoption of some technologies (e.g. energy conservation technologies) by increasing their capital cost as perceived by models (Ekins et al. 2013).

There tends to be limited discussion of the motivation and agency of actors in scenarios, which can overlook important dynamics and limit policy tractability (Hughes 2013; Foxon 2012). In the majority of current scenarios for the UK the social system, and in particular policy, is represented as external to the scenarios and rather simplistically as proactive forcing of a strategy by one actor (the policy maker) (Hughes 2013). This masks the complex co-ordination and alignment of activities and viewpoints of multiple actors required to deliver change in reality. Increasing the focus of scenarios on actors would emphasise the scale of the challenge, whilst also supporting a more detailed understanding of how potential pathways could be shaped by the actions of particular actors in the system, increasing the policy relevance of scenarios. This would also help to identify challenges for particular actors, rather than abstract problems or barriers. However, the integration of qualitative and quantitative analysis remains challenging and relies on expert judgement (Hughes 2013; Longhurst & Chilvers 2012).

Few of the scenarios consider the wider environmental implications of technological change, beyond the qualitative assessment of environmental impact in DECC’s scenarios. In reality, the production of technologies and construction of connecting infrastructure places a significant burden on the environment in terms of consumption of resources and creation of pollution and emissions.

2.3 Low-carbon infrastructure transitions

The low-carbon scenarios described above all show that a rapid and significant change in energy technologies is needed to move towards a low-carbon UK. However, technologies are embedded in a complicated system of connecting networks, institutions, the environment and users. In this sense, infrastructure systems are ‘complex adaptive systems’ (Beinhocker 2006). Institutions in this thesis are defined as “*the humanly devised constraints that structure human interaction. They are made up of formal constraints (rules, laws, constitutions), informal constraints (norms of behavior, conventions, and self-imposed codes of conduct), and their enforcement characteristics. Together they define the incentive structure of societies and specifically economies. Institutions and the technology employed determine the transaction and transformation costs that add up to the costs of production”* (North 1993).

Interactions within complex infrastructure systems and between different systems mean that they are unstable and unpredictable. The outcome of intervention into the system is difficult to forecast; feedback loops make it difficult to distinguish cause from effect; there are significant time and spatial lags; and relationships are non-linear (Costanza et al. 1993; Liu et al. 2007). Small historic events can favour the initial adoption of a particular, often inferior, technology or mode of operation. An often cited example of this phenomenon of path dependency is “the war of currents” (e.g. McNichol 2006): Historically power was generated at large, centralised plants and transferred over increasingly large distances to end-users, requiring high voltages during transmission (to limit potential transfer losses) but low voltages at the end-user (for safety reasons). More efficient voltage transformers for alternating current (AC) compared to direct current (DC) transformers favoured AC for power transmission, and led to a lock-in to a highly standardised and centralised AC power system.

Once established, systemic interactions between technologies, infrastructures, institutions and users can increase the returns to adoption of that technology or mode of operation and constrain the development and adoption of a superior one (Unruh 2000; Arthur 1989). Path dependency can lock in inefficient technologies and behaviours and limit the potential of intervention to create disruptive change (Unruh 2002).

These characteristics of infrastructure systems present challenges for the speed and scale of change required; focus on technology change alone is unlikely to be effective. Instead a large scale transformation is needed in the way that technologies, institutions, organisations, and social and economic subsystems interact to fulfil societal function (Geels 2011; Kemp & van Lente 2011; Hargreaves, Longhurst, et al. 2013). Furthermore, this transition is goal oriented (towards a low-carbon infrastructure system), making it different to historic transitions which were emergent (Geels 2011). This presents many challenges, principally because the concept of low carbon, is ambiguous and contested so there is disagreement over both the pathway and end-point of low-carbon infrastructure (Stirling 2009). Furthermore, the goal of ‘low carbon’ represents a public good, because millions of actors affect the global atmosphere and they all benefit from reduced GHG emissions (Ostrom 2012). The problem is that individuals benefit whether they contribute to change or not and there is also a significant lag between action and effect. As a result individual action does not offer obvious or immediate individual benefits (Stirling 2009). These factors explain why low-carbon infrastructure transitions are difficult and provide further justification for more research in this area.

2.4 Understanding low-carbon transitions

The emerging field of sustainability transition studies provides a number of theoretical approaches to understanding the processes through which low-carbon transition could (or could not) come about. This section reviews approaches that might contribute to the understanding of transitions including; Innovation Systems Theory, Multi-Level Perspective, Co-evolutionary Thinking and Socio-Ecological Transitions. The balance of focus on the interaction of social, technical and environmental systems varies between the approaches, which has implications for the scope and relevance of analysis to infrastructure transitions. A critical review of these approaches will help to shape the methods used for analysis of constraints in this thesis.

2.4.1 Innovation Systems Theory

Innovations systems theory argues that innovation occurs predominantly in response to drivers and barriers coming from the wider innovation system (Foxon, Gross, et al. 2005; Hekkert et al. 2007; Jacobsson & Bergek 2011). While a particular technology can be relatively mature, their deployment in a distinct physical, social and institutional context presents new challenges requiring innovative organisational, contractual and commercial solutions. An excellent example of this is the deployment of district heating in the UK. The technical components are used widely in Scandinavia but the physical setting of the UK, along with restriction in the planning and regulatory systems have limited roll out of a relatively established technology (Hawkey 2012).

The innovation system is the combination of all institutional and economic structures that affect the nature and speed of technological change in society, including networks with other firms and suppliers and policy and regulatory frameworks (Hekkert et al. 2007). To this extent, the theory emphasises the co-evolutionary character of change processes; technologies evolve within particular social and economic contexts, which are in turn shaped by the technologies that are produced and used. Innovation is uncertain, dynamic, systemic and cumulative, not a linear process (Grubler 1998). A number of innovation system approaches have been developed to analyse innovation at different scales, including national innovation systems (OECD 1999), technology innovation systems, (Carlsson & Stankiewicz 1991), and sectoral systems of innovation (Malerba 2001).

The analysis of innovation systems is undertaken with the aim of identifying weaknesses in particular system elements that might be hampering the development of technologies (Jacobsson & Bergek 2011). It has been extensively used by policymakers to isolate bottlenecks in technology development processes (Foxon, Gross, et al. 2005; Hawkey 2012). This approach can reveal weaknesses associated with; experimentation by entrepreneurs; marketing of new technologies; mobilisation of resources; insufficient knowledge development and diffusion; gaining social, institutional and legislative acceptance; and the direction of search (Smith et al. 2010; van den Bergh et al. 2011; Lachman 2013). However there are few instances of the application of innovation systems analysis to alternative modes of infrastructure operation.

Traditional methods of Innovation System Analysis have not been without criticism (Hekkert et al., 2007; Smith et al., 2010; Geels, 2006, 2011):

- Cultural and demand-side aspects tend to be marginalised;
- Interactions between the innovation and the mainstream technology system are not addressed;
- Analysis focuses more on the social structure of systems and associated weaknesses; rather than system dynamics; and
- More focus is placed on large actors, such as institutions and firms, and tends to neglect smaller one, such as grass roots movements and individuals.

The ‘functions of innovation systems’ approach has been proposed to overcome the limitations of quasi-static analysis and the exclusion of actions of entrepreneurs in favour of institutions (Hekkert et al. 2007). This approach aims to undertake more dynamic analysis of the activities and processes that are important for well-functioning innovation systems (Hekkert et al. 2007). By mapping these processes (functions) over time, insight in the dynamics of innovation systems is created. However, this does not address the limitations related to representation of civil society and individuals, the exclusions of cultural and demand-side aspects or the interaction with mainstream technology systems that would be needed to analyse the constraints that are the focus of this thesis. This is not to say that technology is not important when considering infrastructure transition, but that the effect of other systems on technology deployment and use should be considered.

2.4.2 Multi-Level Perspective

The Multi-Level Perspective (MLP) views transitions as non-linear processes arising from the interplay between three analytical levels: technological niches, socio-technical regimes, and landscapes (Geels 2002).

Niches (also called the micro level) are the spaces where radical innovative activity takes place and are partly and temporarily insulated from ‘normal’ selection processes in the regime. Niches tend to be more flexible and less bound by rules (Berkhout et al. 2010). Niches are predominantly perceived as supporting innovations in technology but have more recently been used to analyse social and civil society innovations (Geels 2007; Hargreaves, Hielscher, et al. 2013).

The **socio-technical regime** is the established practices and associated rules that create and reinforce an existing socio-technical system (also called the meso level). Innovation occurs at the regime level but it is characterised as incremental. The regime is described in terms of three interlinked elements: (1) a network of actors and social groups, which develops over time; (2) the set of formal and informal rules that guide the activities of actors who reproduce and maintain the elements of the socio–technical system and (3) the material and technical elements (Geels 2004). As long as regimes are stable they create a strong alignment between these three elements, locking the entire system in (Lachman 2013).

The **landscape** (macro) level is often characterised as exogenous and represents the broader political, social and cultural values and institutions that influence the levels below. The landscape represents the deep structural relationships of a society; which change very slowly. Niches and regimes have minimal impact on the landscape level, but the landscape can influence both significantly, resulting in systemic changes (Lachman 2013).

The multi-level perspective argues that transitions come about through interactions between processes at these three levels (Geels & Schot 2007):

- niche-innovations build up internal momentum, through learning processes, networks, performance improvements and cost efficiencies;
- changes at the landscape level create pressure on the regime; and
- destabilisation of the regime creates windows of opportunity for niche-innovations and the regime adjusts.

When these processes align, niches are able to breakthrough into mainstream markets where they compete with the existing regime. This dynamic is captured in figure 2.5, which also shows the influence that the regime and the landscape have on the perceptions of niche actors and the availability of support networks. However, niches are considered in MLP to be protected from the 'selection pressures' that reinforce the regime. In reality these selection pressures can constrain the momentum of niches to such an extent that they are unlikely to challenge the regime (Bolton & Foxon 2013). This is discussed in more detail in Chapter 3.

Smith et al. (2005) understand regime change to be a function of shifting selection pressures on the regime and the coordination of resources available inside and outside the regime to adapt to these pressures. They argue that "*without at least some form of internal or external pressure… it is unlikely that substantive change to the developmental trajectory of the regime will result*" (Smith et al. 2005: 1495)

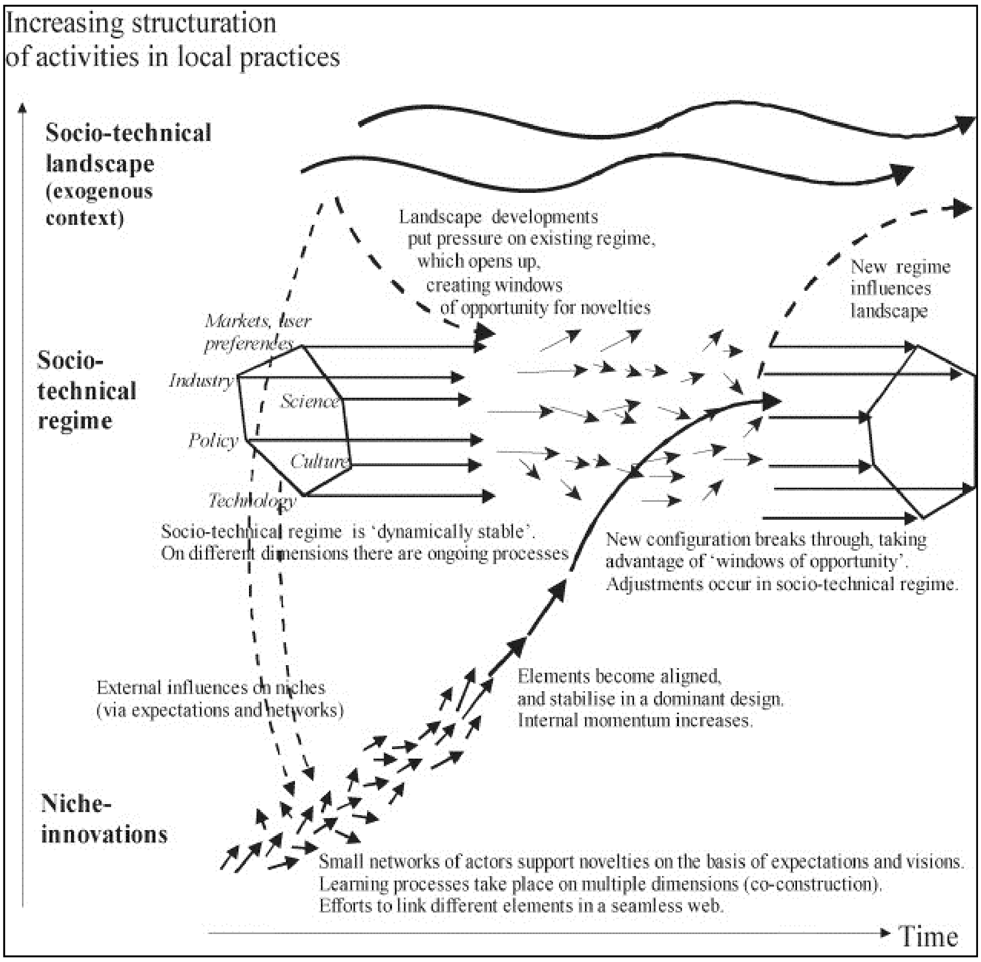


Figure 2.5: Multi-level perspective on transitions (Source: Geels & Schot 2007)

The MLP has provided a useful framework for the analysis of historical dynamics of transitions but has been criticised for a number of limitations relevant to this thesis;

- The role of the landscape, particularly in shaping selection and constraining niche activity is underplayed (Whitmarsh 2012; Kern & Mitchell 2010);
- Niches tend to be conceived as technical artefacts (Hargreaves, Longhurst, et al. 2013). There is little discussion of niches in practices (Shove & Walker 2010) or business models;
- There is a lack of granularity and structure for analysis at each level, particularly relating to business strategies, which makes comparative analysis challenging;
- Explanations are overly functionalistic and do not sufficiently account for choices of actors (Smith et al., 2005);
- Important types and agents of change are missed in much of the literature (Shove & Walker 2007), including individual user choice, development of business strategy and the role of government (Foxon 2011);
- Focus on innovation and transition (micro level) emphasises novelty and in doing so overlooks the wider system of systems that hold things in place and maintain normality (Shove 2003).

It could also be argued that social and technical systems are prioritised over the environmental system, which might underplay environmental influences over transition. To maximise its relevance to transitions in modes of infrastructure operation MLP needs to be adapted to account for relationships and agents that are under-represented in the current conception. Foxon suggests that the MLP approach could be usefully complemented by other strands of research with an explicit evolutionary framing (Foxon 2011). Here, and as discussed in the next section, recent work on co-evolutionary thinking is potentially helpful.

2.4.3 Co-evolutionary thinking

Co-evolutionary thinking has its roots in biological evolution which argues that species, or system components, have traits that change over time, rather than universal characteristics and that these traits are context specific. Random, or at least unforeseeable, introductions of components or traits occur and their survival depends on their fitness. Co-evolution is when this fitness is determined by the characteristics of other systems, with which the new introduction interacts. In this sense, co-evolution is the pattern of evolutionary change of two systems interacting. It is the "*relationships between entities which affect the evolution of entities. Entities and relationships are constantly changing, yet they constantly reflect each other…Everything is interlocked, yet everything is changing in accordance with the interlockedness*"(Norgaard, 1994: 26).

A process of technological and institutional co-evolution has been used to explain the state of 'carbon lock-in' to modern carbon-based energy systems, preventing the development and take-up of alternative low-carbon technologies (Unruh 2000; Unruh 2002; Unruh & Carrillo-Hermosilla 2006; Carrillo-Hermosilla 2006). This builds on the phenomenon of path dependency to show how systemic interactions between technologies, infrastructures, institutions and users can increase the returns to adoption of that technology and constrain the development and adoption of a superior one (Unruh 2000; Arthur 1989).

In order to examine change in a particular regime we need to look at the mutual causal influences between relevant systems including the roles of structure and agency (Kallis & Norgaard 2010). In addition to the interaction between technologies and institutions described above, Foxon (2011) argues that co-evolutionary analysis should account for business strategies (Beinhocker 2006); the practices, preferences, culture and skills of end users (Shove et al. 2008; Safarzyńska & van den Bergh 2010; Nye et al. 2010; Bijker et al. 2012); and the evolution of environmental systems (Norgaard 1994). As a result, Foxon, has identified five co-evolving systems necessary to understand industrial and economic change (Foxon 2011):

- **Ecological systems:** systems of natural flows and interactions that maintain and enhance living systems.
- **Technological systems**: systems of methods and designs for transforming matter, energy and information from one state to another in pursuit of a goal or goals.
- **Systems of institutions**: ways of structuring human interactions. This is taken to include, for example, regulatory frameworks, property rights and standard modes of business organisation.
- **Business strategies**: the means and processes by which firms organise their activities so as to fulfil their socio-economic purposes.
- **User practices**: routinised, culturally embedded patterns of behaviour relating to fulfilling human needs and wants.

Each of these five systems is treated "*as a system that evolves under its own dynamics, but in which this evolution both influences and is influenced by the dynamics in the other systems through causal interactions*" (Foxon 2011, p.2262). Foxon's framework suggests that key events in transitions may occur through technological changes, forming of institutions, revisions to business strategies or changes in user practices; explicitly addressing some of the weaknesses identified in innovation systems and the MLP. It is argued that the framework can be useful for undertaking analysis of dynamic processes (the evolution of relationships and causal interactions) that contribute at multiple levels to a transition to a low-carbon economy (Foxon 2011). As such, it is a useful framework for this thesis. The framework has been used for detailed empirical analyses of the evolution of Energy Service Companies (ESCos) in the UK (Hannon 2012); and recent developments in UK electricity distribution networks (Bolton & Foxon 2011). The authors examined historical change in interactions between systems, predominantly business models and institutions in the former and technology and institutions in the latter, to identify how these interactions might need to change in the future to enable innovation in business models or networks. It has also been used to develop actor-centric transition pathways to a UK low-carbon electricity system by 2050 (Foxon 2012). Pathways were articulated as a function of technology and institutional change and the author described how these changes might be affected by the strategies and behaviours of stakeholders (both end-users and business). Analysis using the co-evolutionary framework has tended to be somewhat limited in its treatment of interactions with the environment or natural resources. As a result, analysis of critical material constraints in this thesis might benefit from alternative approaches to analysis of technology-natural resource interactions.

2.4.4 Sociometabolic transitions

The sociometabolic transitions approach aims to understand and analyse instances of contemporary and historical radical change in relation to their interactions with the environment, in particular the appropriation of natural resources (Fischer-Kowalski & Rotmans 2009). It is based on the theoretical framework of social metabolism which defines society as a hybrid of culture (meaning and communication) and material (humans, livestock and physical infrastructure) (Fischer-Kowalski & Weisz 1999).

Based in this notion, socio-ecological systems comprise a biophysical sphere and a cultural sphere. These two spheres overlap, constituting the "*biophysical structures of society"* which is where nature and culture interact; culture is supported through the appropriation of natural resources to accumulate stocks of humans, livestock and physical infrastructure (Haberl et al. 2004). Those stocks need resources to support their development, maintenance and use so accumulation of stocks necessarily stimulates flows of goods from nature. The process of stock accumulation, maintenance and use can also produce waste and emissions which are returned to nature. This conceptualisation of stocks and flows provides a useful framework to analyse the technology-environment interaction in infrastructure transition; the accumulation of low-carbon technologies stimulate flows of critical materials from nature (Busch et al. 2014). The disruption of these flows (as a result of depletion of the natural system or of disruption of supply chains) will disrupt the accumulation of low-carbon technologies.

Sociometabolic transitions analysis uses the methodology of material and energy flow analysis to quantify the accumulation of stocks and associated material flows (Haberl et al. 2004). Material flow Accounting (MFA) is an accounting framework based on the mass balance principle that quantifies the overall material inputs to a system (the flows), material accumulation within that system (the stocks) and material outputs to other systems or the environment (also flows) (Fischer-Kowalski & Weisz 1999; Eurostat 2009; Eurostat 2001). Stocks and flows modelling has been used for a series of critical materials including Rare Earth Elements (Du & Graedel 2011b) and cobalt (Harper et al. 2012) giving insights into our reliance on these natural resources for our societal development. This thesis aims to build on this work by analysing the potential for disruption of critical material flows in the future.

The analysis of this process of stock accumulation and associated flows has identified a series of metabolic profiles (also called sociometabolic regimes (Haberl et al. 2011)) which represent three modes of subsistence: hunter-gatherers, agrarian societies and industrial society. Transitions between these socio-metabolic regimes can be analysed by tracing changes in socio-economic metabolism (Krausmann et al. 2008). Historical analysis of socioeconomic transitions has identified that transition in metabolism of resources is normally preceded by a change in energy (Fischer-Kowalski & Rotmans 2009). Transitions between these regimes fundamentally change interactions between society and nature, whereas changes and variations within each regime are gradual. Proponents of socio-metabolic transitions argue that a new socio-metabolic regime is required, which will require a fundamental re-orientation of society and the economy, not the implementation of isolated technologies (Haberl et al. 2011).

2.4.5 Reflecting on approaches to understanding low-carbon transitions

The approaches described above are underpinned by similar principles; change is a result of the interaction of different systems (for example society and technology or society and the environment); change is systemic, not linear and analysis needs to be dynamic to reflect this fact. However, there are some crucial differences in the approaches which have significant implications for the outcomes of analysis.

The representation of end users varies significantly across approaches. From almost complete exclusion in innovations system theory to being one of five core systems of analysis in co-evolutionary thinking (Hekkert et al. 2007; Foxon 2011). Exclusion or under-representation of the end-user overlooks their crucial role in selecting technologies and creating demand profiles to which supply must respond. This could encourage a focus on large-scale supply-side technologies, ignoring the potential of demand management and micro-generation to contribute to low-carbon futures.

This focus on large-scale technologies and large actors also has the potential to exclude grassroots organisations in some approaches, such as innovations system theory. If this were the case responses to potential constraints might favour large actors and actively exclude one of the key alternatives to a centralised system, creating lock out of alternative modes of infrastructure operation.

Many approaches, including innovations system theory, the MPL and socio-economic metabolism are criticised as being overly structural, which ignores the agency of actors involved in the transition (Smith et al. 2005; Shove & Walker 2007). Analysis using these approaches might overlook the power of vested interests and their desire to retain the status quo, which could result in ineffective responses to overcoming this lock-in (Moe 2010).

Interactions between innovations and the mainstream are poorly represented in all but co-evolutionary approaches to analysis (Smith et al. 2010). This could mean that the constraints on alternative modes of operation are poorly understood and they are unable to co-exist with the mainstream infrastructure system, which could limit their potential to scale up. It could also reduce the effectiveness of policy aiming to support innovation.

The majority of the approaches focus on technology change, with the exception of co-evolutionary thinking, which has been applied to alternative modes of operation, such as ESCos. A technological focus ignores the fact that the current mode of operation is unsustainable and that alternatives are required to not only respond to technology change but also to encourage the necessary change. A technological focus could reduce the effectiveness of policy responses if they do not enable or support alternative modes of operation.

Few of the approaches recognise the appropriation of environmental resources as being crucial to transition; only socioeconomic metabolism focussed on the environment centrally. It is likely that analysis of material constraints will have to draw on this approach and the associated tools to ensure that resource appropriation is considered fully.

None of the approaches explicitly address diversity; many consider the conditions necessary for an individual technology to become established. In reality, a range of technologies and modes of operation will need to co-exist in order to deliver the necessary scale of change and to create the diversity necessary to avoid future lock-in. Focusing on individual technologies or modes of operation could actively exclude technologies that need to become established in parallel.

Government is represented in the majority of approaches as policies or processes (Hughes 2013); none address the political obstacles created by deep-rooted policy paradigms (Kern & Mitchell 2010). Excluding these macro-political constraints ignores the importance of the political acceptability of potential response. Alternatively, this could limit responses to those that are politically acceptable now at the expense of potentially more effective policy which could be implemented if the policy paradigm changed in the future.

It would appear that relying on one approach would be of detriment to the proposed analysis in this thesis, therefore it will be necessary to use a combination of approaches to ensure that vital parts of the system are not omitted. There are precedents of combining transitions approaches with other social science approaches, in particular to reflect the agency of individuals (Hargreaves, Longhurst, et al. 2013)

2.5 Governing infrastructure transition

There is growing interest in the role of governance to initiate intentional transition towards more sustainable development, including low-carbon infrastructure systems (Smith et al. 2005; Kern & Smith 2008). Responses to constraints could be informed by empirical and theoretical research in this domain so some of the most relevant literature is reviewed below.

The complexity and interdependence of infrastructure systems, both technically and socio-economically, requires that any action to intervene must recognize both the historic co-evolution of the system and the path dependency that could limit the potential of disruptive change (Unruh 2002). Overcoming system lock-in requires systemic action that cuts across policy sectors, public and private institutional boundaries and state jurisdictions (Smith 2007). Understanding change, or transitions, as a co-evolutionary process "*highlights the uncertain, path-dependent and cumulative nature of systems change"* and the need to develop interventions that recognise the interconnectedness of infrastructure, with the potential for overcoming system lock-in (Foxon 2011: 2265).

Transitions governance scholars argue that while it is not possible to predict and direct transitions, it should be possible to influence ongoing transition dynamics in terms of speed and direction. Smith et al (2005) define transition dynamics as a function of:

- selection pressures - processes which exert pressure for change for example competition, regulation and cultural attitudes; and
- the capability of an organisation or individual to respond to those selection pressures - determined by availability of resources and the ability to co-ordinate resources.

This definition is useful because it highlights three principle intervention mechanisms to influence transition dynamics; alignment of selection pressures acting on the regime with a particular direction of travel; increasing resource available to respond to selection pressures; and supporting co-ordination of responses to pressures (Smith et al. 2005). Alignment of selection pressures could be achieved by improving the coherence and orientation of selection pressures or by translating pressures into a form that prompts and enables response. For example water companies in the UK are required to demonstrate action to reduce demand but are not able to generate revenue from sources other than sales of units of water, which could off-set losses made from reducing unit sales of water.

This conceptualisation supports the evolutionary concepts of variation and selection; however, it is important that diversity is maintained to enable flexibility to changes in the environment (Stirling 2010). Furthermore, policy will need to take a longer-term approach to transitions and support both the rapid diffusion of sustainable innovations in the short-term and gradual changes in existing regimes in the long-term (Verbong & Loorbach 2012). It is important that there is a link between the long-term and short-term to ensure that momentum toward large-scale and desirable change is maintained (Walker et al. 2013). This means that policy itself will need to develop adaptive capacity, amending support for innovations as the regime alters and new institutions, practices and cultures develop. The great challenge is to break old lock-ins without creating new, equally unsustainable lock-ins (Berkhout 2002).

In response to these challenges, a number of specific approaches to governance of transitions have emerged in the academic and policy literature, including Transitions Management and Strategic Niche Management. Furthermore, theories of governance have been developed to address the multiple levels of governance necessary to influence transitions in complex systems. Therefore, literature on Multi Level Governance and self-governance are briefly discussed including the principal features of these approaches, along with their limitations.

2.5.1 Transitions Management

Transitions Management (TM) is a model of co-evolutionary management of transformative change in societal systems through a process of searching, learning, and experimenting that is widely used in the Netherlands (Kern & Smith 2008; van der Loos & Loorbach 2012; Rotmans & Kemp 2008). Managing is taken to mean adjusting, adapting, and influencing rather than the command-and-control mode (Kemp & Loorbach 2006; Rotmans & Loorbach 2010). The TM approach uses positive visions of the future and explores a range of possible options and pathways to that future by carrying out a diversity of small-scale experiments. It is designed to be a participative approach which combines long term thinking with short-term action through a process of reflection and learning. The approach uses a cyclical process that alternates between three levels of activity:

- Strategic: transition arenas, comprising public-private networks of frontrunners, are created to develop visions, goals, roadmaps and milestones that overcome lock-in to existing systems effectively bypassing existing policy networks and acting outside the regime.
- Tactical: the regime is aligned with the long-term goal; allowing a transition in the desired direction. Including the regime in the transition is assumed to boost legitimacy, support and financing of the process (Rotmans & Loorbach 2009).
- Operational; experiments are used to stimulate learning and guide variation and selection.

However, in practice, it is often more complex than this. When reflecting on whether the Dutch government had governed the energy transition van der Loos and Loorbach concluded that the government had created the space for energy transition and had facilitated innovative development in society but had not brought about mainstream changes (van der Loos & Loorbach 2012). They recognised the challenge of government culture, structure and practices, concluding that the current Dutch government was an unlikely candidate to lead the process of systemic change.

Shove and Walker criticise the simplistic portrayal of transitioning as a managerial task (Shove & Walker 2007). This neglects the influences of belief systems, political interests and culture, which could obstruct or even prohibit transitions. The approach had also been criticised for the fact that it abstracts the system from its historical and contemporary environment and that transition managers exist who "*sit outside 'its'[the system's] boundaries and who can apply management tools including levers, niche-building machinery, and engineering devices from a privileged, knowledgeable, and, above all, external position'* (Shove and Walker 2007: 765). However, the purpose of transition management is to accelerate change and it is important that the need to represent system dynamics is balanced with the need to act.

2.5.2 Strategic Niche Management

The MLP literature emphasises the role of niches in transition and strategic niche management (SNM) theory has evolved from evolutionary theories on technological change (Lachman 2013). SNM explores how niches persist and evolve and how policy makers can support those processes to challenge and potentially replace existing socio-technical regimes (Hargreaves, Hielscher, et al. 2013; Raven 2012). SNM has been used extensively to investigate a wide range of innovations, including; the diffusion of grassroots innovations community energy (Seyfang & Haxeltine 2012; Hargreaves, Hielscher, et al. 2013); alternative fuel vehicles (Kwon 2012) and urban water management (Farrelly & Brown 2011).

It proposes three principal strategies to support niche development; creating protective space within which niches are shielded from regime selection processes; nurturing innovations to support their development; and empowering niches to move from the protective space and challenge the mainstream (Smith & Raven 2012). A number of analytical themes have emerged to support the understanding of niche dynamics and management of transitions (Raven 2012).

Niche development processes occur at both the local and global scale. Experiments occur in a local context, with support from local networks and generate locally applicable lessons. These are supported by networks of actors with the aim of knowledge exchange, defending niche interests and interacting with the wider world – what Hargreaves et al (2013) call intermediaries. Intermediaries help to translate local lessons into the generic context, independent rules and knowledge.

Clear articulation of niche expectations and potential future benefits helps to reduce uncertainty in the innovation process and allows the mobilisation of resources. Once accepted, the expectations can be used to position the niche and influence others. Expectations are particularly powerful when shared by a number of actors, when tangible and when specific. Expectations are dynamic, vary between different actors and depend on dynamics in the wider regime and landscape (Geels & Raven 2006).

Social networks are important to niche development because they enable learning and dissemination of expectations, lessons and experiences. Networks are only successful if they have representation from actors outside the regime, especially those from policy, non-governmental organisations (NGOs) and users (Schot & Geels 2008). Learning itself is essential to niche development but learning is not constrained to technology and economic dimensions. Alignment is needed between technical learning and social learning about user preferences, culture and symbolic meaning, industry and production networks, regulations and government policy (Smith et al. 2010).

SNM has been used to develop a deep understanding of the dynamics of niche building but has not been used as a framework for policy makers or managers. It's not clear which processes promote network broadening, successful negotiation of shared expectations and destabilisation of regimes. More work is needed to understand how protection is developed, applied and withdrawn (Hommels et al. 2007). Furthermore, little attention has been paid to the winners and losers of the transition processes, particularly when resource is moved from supporting the regime to supporting niches, and the role that power plays in supporting or preventing this processes.

2.5.3 Multi-level governance

The study of multi-level governance is concerned with the changing interaction between different levels of governance and the engagement of non-governmental actors in policy development and delivery (Smith 2007). There is agreement that this dispersal is more efficient and that it is desirable for governance to be multi-level to capture variations in territorial reach of policy and to better reflect diverse preferences among citizens (Marks & Hooghe 2003). The distribution of responsibility promoted by multi-level governance is very relevant to infrastructure transition, where change can be initiated by different actors, at different scales and is likely to involve more than one type of actor (Smith 2007).

There are contrasting views of the ideal form of multi-level governance in the literature which Marks and Hooghe (2003) have formalised in two types:

Type 1 – well ordered, nested responsibilities at clearly demarcated levels, usually following territorial boundaries. Responsibilities are distributed between multi-functional institutions and networks. However, whilst clear demarcation of responsibility is beneficial, it can be undermined by a breakdown between levels as a result of the varying drivers and capabilities of actors at different levels (Leventon & Antypas 2012). Type 1 multi-level governance tends to provide a greater degree of stability and security in policy direction but has been criticised for lacking innovation and adaptation to changing conditions (Smith 2007).

Type 2 – more fluid and problem focussed, intended to be flexible, rather than durable, with intersecting membership. This presents more opportunity for innovation and adaptation. There are some examples of type 2 multi-level governance in the UK, particularly in the former Regional Development Agencies (Smith 2007). However, it presents significant challenges in communication and monitoring progress, can reduce the stability of the policy environment and reduces certainty about future policy and represents a significant change from the current top-down approach.

The problem-focussed nature of this form of type 2 multi-level governance supports engagement of a multitude of actors across territorial scales and encourages more integrated and innovative policy (Smith 2007). A type 2 approach identifies three constituent governance processes: establishing common problem framing; building resourceful actor networks; and negotiating commitment between different interests (Keeley & Scoones 2003).

One challenge of effective operationalisation of multi-level governance is that type 2 tends to be embedded in legal frameworks determined by type 1 (Marks and Hooghe 2003) and ultimately depend on type 1 authority for their effectiveness (Smith 2007). This is a particular problem in the governance of infrastructure, where the strategic regulation of infrastructure and markets are negotiated at the national government level. Sub-national and non-governmental actors have very little direct control over these issues. This can actively inhibit the adoption of decentralised solutions that arise from actors outside the traditional infrastructure marketplace (Bolton and Foxon 2013; Smith 2007).

Multi-level governance provides some useful lessons in the analysis of alternative infrastructure governance:

- The importance of developing a common framing of the problem, which requires translation between different discourses and actors (Smith & Stirling 2006).
- The need to engage private, public and civil society partners in the process of innovation, business development, community involvement, knowledge production, infrastructure provision, communication, regulation, market creation and policy (Smith 2007).
- The need for national regulation to allow innovation at different scales and from different, non-utility company actors (Bolton and Foxon 2013).

However, there are a number of limitations to the application of multi-level governance in isolation to understanding how governance might respond to constraints. Multi-level governance is somewhat passive in its approach to analysis, in that it aims to analyse the characteristics of relationships, rather than understand how these relationships will evolve and how this can be used to drive change. Furthermore it does not specifically address the interaction of actors and technology, which is fundamental to infrastructure transition (Foxon 2011). Nor does it adequately deal with the nature or structure of internal governance of local initiatives, thus we look for additional theories that might explain these phenomena.

2.5.4 Institutional dynamics

Elinor Ostrom's work on institutional diversity has provided a great deal of insight into the diverse arrangements that make it possible to conserve and efficiently utilise jointly managed resources (Ostrom 2010). This work has focussed on collective management of Social-Ecological Systems (SES); however, these systems have much in common with infrastructure systems, particularly those which demonstrate characteristics of common and public goods (Little 2005). In this section the relevance of Ostrom's work on self-governance of natural resources to infrastructure systems is discussed.

Ostrom identifies a series of variables that are suggested to affect the likelihood of resource users to self-organise in order to manage a particular SES (Ostrom 2007). The capacity to create effective governance of SES depends on the nature of these variables; particularly how they affect the costs of investing in better norms and rules and the benefits of self-organising. The variables relate to the nature of the resource system, the characteristics of the resource users, and the capacity of resource users, in terms of their autonomy and system knowledge.

These variables were developed into design principles to help understand the requirements of internal systems of institutions required to support self-governance (Ostrom 1990). These principles were updated based on 20 years of empirical evidence (Cox et al. 2010). They represent "*core factors that affect the probability of long term survival of any institution developed by users of a resource"* (Ostrom 2010: 653). The updated design principles are:

1A. *User Boundaries*: Clear and locally understood boundaries between legitimate users and nonusers are present.

1B. *Resource Boundaries*: Clear boundaries that separate a specific common-pool resource from a larger social-ecological system are present.

2A. *Congruence with Local Conditions*: Appropriation and provision rules are congruent with local social and environmental conditions.

2B. *Appropriation and Provision*: Appropriation rules are congruent with provision rules; the distribution of costs is proportional to the distribution of benefits.

3. *Collective Choice Arrangements*: Most individuals affected by a resource regime are authorized to participate in making and modifying its rules.

4A. *Monitoring Users*: Individuals who are accountable to or are the users monitor the appropriation and provision levels of the users.

4B. *Monitoring the Resource*: Individuals who are accountable to or are the users monitor the condition of the resource.

5. *Graduated Sanctions*: Sanctions for rule violations start very low but become stronger if a user repeatedly violates a rule.

6. *Conflict Resolution Mechanisms*: Rapid, low cost, local arenas exist for resolving conflicts among users or with officials.

7. *Minimal Recognition of Rights*: The rights of local users to make their own rules are recognized by the government.

8. *Nested Enterprises*: When a common-pool resource is closely connected to a larger social-ecological system, governance activities are organized in multiple nested layers.

These principles were developed to guide the design of institutions for self-governance of natural resources, there has been no research to date that attempts to illustrate their application to infrastructure.

2.6 Summary

Low carbon is a contested term with multiple interpretations and myriad assumptions underpinning any definition. The UK defines low carbon as activities that contribute to an 80% reduction in the UK's GHGs by 2050 whilst maintaining secure energy supplies, maximising economic opportunities, and protecting the most vulnerable (HM Government 2009).

Scenarios have been used extensively in the UK to identify different pathways to a low-carbon future but these are predominantly technical feasibility studies, which aim to optimise the combination of technologies rolled out to achieve low-carbon targets at minimum cost (Hughes 2013). The societal and environmental implications of these scenarios are rarely explored in any detail.

Infrastructure systems are complex and liable to become locked in to potentially unsustainable configurations (Unruh 2000). Understanding past change and the potential constraints of future change requires us to account for the interconnection of social, technical and environmental systems. This chapter has reviewed a range of approaches to support the analysis of transitions, which will help with the analysis of constraints to future transitions. These methods have been applied predominantly to technology-related transitions, rather than changes in the mode of operation of infrastructure, and have not paid a great deal of attention to the link between technology and natural resources. Therefore, these approaches will need to be adapted to make them more relevant to the aims of this thesis.

The literature relating to the governance and management of transitions provides useful insights into the mechanisms and processes that might help to remove constraints and accelerate desirable transitions. However, these approaches do little to account for the role of power in helping or hindering transition, particularly in relation to organisations with vested interests in retaining the status quo. Current approaches to governance of transitions tend to underplay the role of the landscape and the policy paradigm in resisting change (Whitmarsh 2012; Kern & Mitchell 2010); the approach to development of policy responses to constraints in this thesis will explicitly take into account this relationship.

3 Constraints to infrastructure transition in the UK

Despite recent developments in understanding transition dynamics and transition scenario development there are still real challenges when it comes to moving from visionary scenarios to practical implementation. The challenges of altering transition dynamics and the implications of change are predominantly overlooked. This is a particular problem when fundamental system change is necessary, such as the scale of infrastructure change necessary to avoid dangerous climate change. Furthermore, analysis of constraints to realising transitions is confined to technical and economic issues. Recognition of a broader range of constraints could help to understand and overcome lock-in to the current regime (Unruh 2000). This chapter describes the current work which has attempted to analyse the two constraints identified in the introduction chapter along with the principle gaps in relation to the aim of this thesis.

3.1 Critical material constraints to low-carbon energy technologies

Low-carbon technologies, such as wind turbines, solar panels and hybrid and electric vehicles rely on critical materials, such as rare earth elements, indium and tellurium (US Department of Energy 2011; Moss et al. 2011). Critical materials are commonly defined as those at risk of supply disruption and which are difficult to substitute (European Commission 2010b). The risk of disruption is amplified by the potential scale of demand for low-carbon infrastructure, which is unprecedented. It has been recognized that the deployment of low-carbon technologies is susceptible to disruption in the supply of critical metals (IEA-RETD 2012).

Studies of the implications of critical materials for low-carbon technologies fall into two broad categories:

1) Those that quantify the availability of a critical material for a particular technology in order to determine the growth potential of that technology over a particular period; and
2) Those that estimate the potential for disruption in the supply of critical materials and the impact that disruption has on the system of interest.

These two approaches to critical material constraint analysis and their application to low-carbon energy technologies are discussed in more detail below.

3.1.1 *Estimation of critical material availability to enable technology growth*

Quantitative analysis has been undertaken by a number of authors to determine the effect of critical material availability on the maximum growth potential of low-carbon technologies. Availability of critical materials tends to be defined as the production of mines, which is reported annually by the US Geological Survey (US Geological Survey 2013).However, in order to quantify availability over a period of time, future production must be forecast. This would require a series of assumptions about the potential for existing mines to increase output and for new mines to be constructed. These assumptions are affected by the quantity of a material present in the environment that could be exploited[4] but are more dependent in the short-to-medium term on restrictions to increasing mining operations. There is a great deal of uncertainty about both the ultimate quantity of material available and the potential to scale up mining which makes forecasts of production, and hence material availability, extremely difficult (Speirs, Houari, et al. 2013).

Demand for critical materials is determined by the number of technologies using a particular material and the intensity of material use in that technology. Scenarios of energy system change are used to determine the future roll out of low-carbon technologies, with many researchers using the International Energy Agency’s (IEA) global scenarios in Energy Technology Perspectives (IEA 2010). However, a range of technology scenarios are used by other researchers which can lead to highly variable estimates of technology numbers. The material intensity of each unit of technology is equally variable. It is influenced by the performance characteristics of the technology, technology design and the efficiency of production (Speirs et al. 2011). The development of these variables over time can substantially influence the outcomes of analysis.

The growth potential of low carbon technologies is usually assessed by comparing the availability of a particular material to the demand for that material from a pre-determined roll-out of low carbon technologies over a particular period of time. The maximum potential growth rate is deemed to have been constrained if availability over the period is less than forecasted demand. However, this tells us little about when this constraint might occur. This is important for low-carbon transition, because the timing of technology roll-out is as important as the final number of technologies deployed. Timing is important for a number of reasons:

---

[4] Reserves are the part of the resources that have been fully evaluated and are commercially and legally mineable. These constitute only part of the total resource, which includes material that has properties that make it suitable for specific uses and is present in sufficient concentration and quantity to be of intrinsic economic interest, but that is not yet economically viable to extract.

firstly, as discussed in chapter 2, reducing the quantity of cumulative emissions is more important than decarbonising the energy system by a particular point in time. If the deployment of low-carbon technologies is slowed in the early stages of a transition it could mean that cumulative emissions over the period exceed safe limits, even if deployment eventually catches up. Secondly, disruptions are more likely to happen in the short-term because mine production cannot adapt quickly to meet structural changes in demand patterns (Morley & Eatherley 2008). The increasing volatility in commodity prices does not provide sufficient motivation for investment in new mines (Cashin & McDermott 2002; Morley & Eatherley 2008); furthermore, even when investment decisions are taken, it can take between 9 to 25 years to get the necessary permissions and infrastructure to increase production (Moriguchi 2010). Although an equilibrium between supply and demand may be achieved in the long-term, the metals market in the short-term is more characterised by disequilibrium than equilibrium (Morley & Eatherley 2008).

Material availability analysis has been undertaken for a series of individual technologies such as electric vehicles (Andersson 2001b; Speirs, Houari, et al. 2013; Angerer et al. 2009; Alonso et al. 2012); solar panels (Fthenakis 2009; Candelise et al. 2012; Angerer et al. 2009; Andersson 2000); concentrated solar thermal power (Pihl et al. 2012); and wind turbines (Kara et al. 2010; Speirs, Houari, et al. 2013; Alonso et al. 2012). All studies compared estimated material requirements for a particular technology with estimates of availability. These studies are summarised in table 3.1.

In the majority of studies, production was assumed to remain static because of the uncertainty associated with estimating future production potential (Speirs, Houari, et al. 2013). Some studies, such as Fthenakis (2009) and Kara et al (2010) produced forecasts of future production but provided a range of future estimates to take into account the uncertainty associated with these values. Some studies used estimates of geological reserves as an alternative to annual production, to represent the total supply available for technology production. Where used, estimates of reserves were derived from data produced by the US Geological Survey (US Geological Survey 2013).

Table 3.1: Summary of studies of critical materials use in low-carbon technologies

Study	Goals	Technologies	Geographic focus	Time-scale	Materials analysed	Analytical approach
Andersson 2001	Identify if availability of metals could constrain technology rollout	Electric vehicles (EV)	Global	2050	Co, Li, Ni, V, REE[5], Cd, Pb	Determine number of EVs that could be produced from available metal resources
Alonso et al 2012	Identify future demand for REEs and assess implications for REE production	Wind turbines and EV	Global	2035	REE	Scenarios of demand from all uses. Assessment of future ratio of demand to supply
Angerer et al 2009[6]	Identify how emerging technologies will drive demand for materials	Fuel cells for EV, thin film Photo voltaic (PV), permanent magnets	Global	2030	Ga, Nd, In, Pt, Co, Se,	Estimate future demand from emerging technologies. Assessment of ration of future demand to current production
Candelise et al 2011	Explore claim that material shortage will affect technology roll out	Thin film PV	Global	2050	Te, In, Ge	Comparison of previous estimates of production with estimates of future demand using IEA forecasts of market
Fthenakis 2009[7]	Explore whether availability of materials could limit potential for photovoltaic systems	Thin film PV	Global	2100	Te, In, Ge	Scenarios of material production and calculation of annual production of PV units that can be supported by this production.
Kara et al 2010	Identify how to avoid demand and supply imbalance	Wind turbines and EV	Global analysis, UK recommendations	2030	REE	Scenarios of future supply and demand
Pihl et al 2012	Identify material restrictions of large-scale application of CSP technology	Concentrated solar power	Global	2050	Cr, V,	Future demand estimated using IEA CSP roadmap and LCA data and calculating ratio with reserve data
Speirs et al 2013	Impact of resource scarcity on achievability of CO_2 reduction targets	Wind turbines and EV	Global	2050	Li, Nd	Comparison of estimate for future demand with current production

---

[5] Rare Earth Elements

[6] Only materials and technologies relevant to this thesis included

[7] Note – one of studies assessed in Candelise et al 2011

Assumptions used to determine material intensity vary from using figures quoted from industry sources (Kara et al. 2010) to detailed analyses of technology composition and efficiency (Candelise et al. 2012; Speirs, Houari, et al. 2013). Alonso et al (2012) included demand for Rare Earth Elements (REE) from sectors other than low-carbon technologies, one of the few studies to include additional demand in comparisons. The authors developed a series of demand projection scenarios for the evolution in demand for REEs in different sectors(Alonso et al. 2012). The inclusion of all end-uses provided a more realistic reflection of future demand and the use of scenarios was an effective way to manage uncertainty.

The majority of studies of individual technologies concluded that the growing market for low-carbon technologies could have significant implications for material supply, exceeding current production levels by factors of up to 2,900% (Speirs, Houari, et al. 2013). However, there was less agreement about whether this would cause constraints. Speirs et al (2013) reviewed a number of assessments of material availability for electric vehicles, wind turbines and solar panels and argued that there was no evidence to suggest any barriers to increasing production to meet the midpoint of estimates for future demand (Speirs et al. 2011; Speirs, Houari, et al. 2013). Andersson was less optimistic and concluded that of the four photovoltaic and nine electric vehicle battery technologies assessed, all were constrained by metal availability to some extent (Andersson 2001a). Similarly, Alonso concluded that some specific uses of REEs, such as high performance magnets, would be negatively affected by constraints in some REEs (Alonso et al. 2012).

All authors recognised the significant uncertainty associated with estimates of future availability, technology roll out and material requirements per unit of technology. Speirs et al concluded that *"a thorough assessment of the long term effects of material availability … still requires a much improved understanding of the potential for, and the economic implications of, expansion in both production and recycling…"* (Speirs et al. 2013: 73). In subsequent work Speirs and colleagues have employed system dynamics modelling to better represent the complexity of resource availability (Houari et al. 2014). However, despite this more sophisticated treatment of the growth potential, economic implications were excluded from analysis. The study did not account for the effect of material price, which is *"dependent on complex supply and demand dynamics of the end use markets as well as economics of extraction and production"* (Houari et al. 2014: 142). The inclusion of price impacts on availability and demand would be a valuable addition to analysis to this kind of analysis but to date has been excluded, owing to the complexity of analysis.

There is a great deal of consensus between studies about the unprecedented scale in the increase of demand for critical materials from low-carbon technologies. One crucial difference is the inclusion of demand from sources other than low-carbon technologies. This is substantial and its exclusion could account for the optimism expressed by Speirs et al, when compare to other, recent studies. It is important that this is included in the analytical approach used in this thesis.

These studies assessed individual technologies or, in some cases, two technologies demanding the same material. More systemic analyses have been undertaken to assess material requirements of a particular part of the energy system such as the hydrogen energy system (Kleijn & van der Voet 2010) or low-carbon power generation (Kleijn et al. 2011). These reviews included a wider range of technologies, as well as connecting infrastructure. Kleijn and van der Voet identify three areas of a hydrogen energy system that could be prevented from upscaling to a level needed for a substantial transition; rare elements in current thin-film solar cells, neodymium in direct-drive permanent-magnet wind turbines and electric motors, and platinum in electrolysis and fuel cells. They also identified a significant demand for steel (and associated chromium in stainless steel) for both wind turbines and hydrogen pipeline, highlighting the potential for constraints from materials not usually perceived as critical. Kleijn et al (2011) are one of the few to use lifecycle analysis (LCA) to quantify material requirements of technologies. This has a number of limitations including the exclusion of many new and emerging technologies. LCA data is limited to current technologies; therefore this analysis would have excluded a number of emerging technologies (for example direct drive wind turbines) and the critical materials they contain (for example neodymium). Despite this, Kleijn et al summarise that the switch to non-fossil electricity would result in higher demand for a number of metals.

3.1.2 Estimation of potential and impact of supply disruption

An alternative perspective to quantifying availability of critical materials is that access to critical materials cannot be described solely as a function of long-term geological availability. In part, this is a result of the significant uncertainty associated with estimating future production or remaining reserves (Tilton & Lagos 2007; European Commission 2010b). However, there is increasing recognition that access to critical materials could be affected by a series of political, geographical and environmental factors as well as geological availability (Graedel et al. 2012; Morley & Eatherley 2008; US Department of Energy 2011; European Commission 2010b). This places the focus of analysis more squarely on the short-term dynamics of supply, rather than on long-term equilibrium, which is more relevant to the analysis of low-carbon transitions (as discussed above). Furthermore, recent studies have tried to develop a more sophisticated understanding of the implications of a disruption in the supply of critical materials, depending on the system of interest (Graedel et al. 2012).

As a result, a more comprehensive, but mostly qualitative, approach to criticality assessment has emerged, which describes constraints from material criticality in terms of the *potential for supply disruption* of a particular material, and *the impact of* this disruption on the system of interest; an approach that is analogous to risk assessment. These assessments have not yet reached a common definition of criticality, besides these two dimensions, and the conceptualization of the dimensions themselves varies significantly between assessments (Erdmann & Graedel 2011). The treatment of these two dimensions in recent studies is described below.

3.1.2.1 Supply Disruption

Supply disruption is conceived to result from a range of factors including political, regulatory and environmental factors limiting the ability to increase production; market risks associated with concentration of production; co-dependence on other markets, where the material of interest is a by-product of another material; and the risk of increasing material prices affecting the cost of emerging technologies (US Department of Energy 2011; Erdmann & Graedel 2011). Analysis of supply disruption is predominantly static, although some have done static assessments of different time periods (US Department of Energy 2011; Graedel et al. 2012).

The majority of studies considered the concentration of mineral deposits in a small number of countries to be a potential source of supply disruption. Shorter term studies have used sources of imports as a measure of the current distribution of supply (National Research Council 2008; US Department of Energy 2011), while others have used the distribution of global production as a measure of the short to mid-term supply basis (European Commission 2010b; Graedel et al. 2012; Moss et al. 2011). Some studies weight country shares of production by political risk using the World Bank's "Worldwide Governance Indicator" (European Commission 2010b) others included political, social and environmental factors as specific risk factors (Moss et al. 2011; US Department of Energy 2011; Graedel et al. 2012) emphasizing that some countries have an increased risk of disruption as a result of these factors (Le Billon 2008). The supply disruption aspect of assessments often includes additional factors such as substitution or recycling, to represent the ability to alleviate disruption through reduction in demand for primary material. This is one of the principal causes of difference between assessments, since these factors are considered to be an aspect of exposure in other assessments (Erdmann & Graedel 2011).

3.1.2.2 Impact of supply disruption

In most studies, there is some measure of the impact of supply disruption on the system of interest (European Commission 2010b; Graedel et al. 2012; US Department of Energy 2011; Mason et al. 2011). The concepts of exposure and vulnerability are used interchangeably to represent this impact and justification for the chosen concepts is seldom provided (Erdmann & Graedel 2011). To some extent this is dependent on the nature of the system of interest but the choice implies a fundamental difference in the conceptualisation of criticality. The use of exposure is consistent with the conceptualisation of criticality as risk, which is the product of the potential and severity of an event. An example of the conceptualisation of criticality as a risk is the European Commission (European Commission 2010b), which assesses impact as exposure using the relative economic contribution of the sector using the material of interest (in terms of Gross Value Added).

Others include the ability of the system to respond to disruption or its adaptive capacity in the conceptualisation of impact. In a material criticality context, such responses include substitutability or recycling of materials to reduce primary demand (Reck & Graedel 2012). For example, Graedel et al (2012) use a combination of the importance of the material of interest and the ability of the system to respond to disruption, which is better aligned with the concept of vulnerability (UNISDR 2009; Balica et al. 2012). It is important that the concept of impact is clearly defined and justified to improve the transparency and comparability of methodologies.

3.1.2.3 Results of qualitative studies

The results of two recent, qualitative assessments of material criticality of low-carbon energy technologies are presented in table 3.2. The boundary between critical, near critical and not critical is subjective and qualitative.

These studies have provided useful insights about the constraints critical material might place on individual technologies or systems, within the bounds of uncertainty of available data. However, assessments which have attempted to represent a more complex view of criticality, taking into account a more system view of the drivers and impacts of critical material supply disruption, have been predominantly qualitative in nature and have relied heavily on expert opinion. Critical material constraints are time-dependent, since the contributing technical, socioeconomic and environmental factors all vary over time. Despite this, no previous studies conducted a fully dynamic criticality analyses. Some have done static assessments of different time periods (US Department of Energy 2011; Graedel et al. 2012), or analysed stock and flows of materials over time (Alonso et al. 2008; Du & Graedel 2011b). There is a need for a more quantitative method of criticality analysis that incorporates some of the dynamic aspects of critical material constraints, which will be explored further in this thesis.

Table 3.2: Criticality of materials used in low-carbon energy technologies; summary of recent study results (US Department of Energy 2011; Moss et al. 2013)

	US Department of Energy		**Moss et al[8]**	
	Short-term	**Medium-term**		
Critical	Dysprosium Europium Neodymium Terbium Yttrium	Dysprosium Europium Neodymium Terbium Yttrium	Dysprosium Europium Neodymium Terbium Yttrium Graphite Germanium	Graphite Tellurium Gallium Indium Rhenium Hafnium Platinum
Near Critical	Cerium Indium Lanthanum Tellurium	Lithium Tellurium	Lanthanum Cerium Samarium Gadolinium Cobalt	Tantalum Niobium Vanadium Tin Chromium
Not critical	Cobalt Gallium Lithium Manganese Nickel Paseodymium Samarium	Cerium Cobalt Gallium Indium Lanthanum Manganese Nickel Praseodymium Samarium	Lithium Molybdenum Selenium Silver Nickel	Lead Gold Cadmium Copper

3.1.3 Relevance of approaches to this thesis

Both approaches to criticality analysis have aspects of relevance to this thesis. However the timing of any disruptions is particularly important to assessment of material constraints to low-carbon infrastructure transition because the rate of roll out of technologies is as important as the total number rolled out. The approach to analysis used must be able to identify potential constraints in the short- and medium-terms rather than on aggregate over the period of analysis. Comparison of availability and demand over a period of time is not sufficiently detailed to achieve this aim so will not be used in this thesis. Furthermore, short-term disruption is a function of more than just the potential rate of production; therefore, analysis of disruption in this thesis needs to take into account wider range of determining factors (European Commission 2010b). It is recognised that commodity price is an important

---

[8] Grouping is slightly different between reports. Critical includes 'high' and 'medium-high' criticality, near criticality is 'medium' criticality, not critical is 'medium-low' and 'low' criticality as defined by Moss et al.

determinant of future production; however the dynamics of this relationship are extremely complicated and uncertain and are beyond the scope of this thesis.

Many previous studies have excluded demand from sources other than low-carbon technologies from analysis, which omits an important dimension of criticality; that of competition between low-carbon technologies and other end uses for access to critical materials. It is important that the approach developed in this thesis includes demand from non-low-carbon technology sources to represent this competition. Finally, the conceptualisation of impact of supply disruption must be clearly articulated and transparent to improve comparability between studies (Erdmann & Graedel 2011).

3.2 Constraints from infrastructure policy and regulation

The following sections review literature relevant to the constraints that governance places on low-carbon infrastructure transitions. In line with the scope of this thesis it focuses on the constraints of specific policy and regulatory instruments as well at the overarching policy paradigm within which these instruments are implemented. To use the language of the multi-level perspective policy and regulation would be at the regime level and policy paradigm at the landscape level. The review draws from literature from international studies but applies the insights to the UK, the focus of this thesis.

3.2.1 Infrastructure policy and regulation

Policy and regulation plays an important role in driving change in infrastructure systems but it can also constrain change (Foxon 2011; Paavola & Adger 2005; Vatn 2005; Bolton & Foxon 2013). The current configuration and approach to policy and regulation in the UK presents a series of constraints to alternative modes of operation. The current regulatory regime has co-evolved with the centralised operation of infrastructure by large, international companies. This has created a regulatory regime designed around the requirements and capabilities of these organisations, which can actively exclude new actors (Mitchell & Woodman 2010). For example, the electricity retail and wholesale markets were introduced to liberalise the energy system and introduce competition into a monopoly market. However, the market designs themselves are oriented to suit large energy companies who have the capabilities and resources to manage complex forecasting and accounting systems. These systems can present unassailable barriers for smaller organisations, which face excessively high transaction costs, require expensive and complicated software systems and demand specific expertise to interact with the market (Mitchell 2014). In practice, this leaves little room for diversity and innovation (CST 2009b).

Competition is viewed as central to the duties of UK economic regulators[9] because it is seen as enabling the best outcomes for society (as opposed to the interests of a monopoly company) (BIS 2011). It is considered to be vital to deliver unit price reduction, which is generally used as a proxy for protection of customer interests. This view is based on the presumption that domestic customers are monolithic, rational agents, who benefit only from current low unit prices (to the exclusion of long term stability, or paying a lower overall cost including efficiency measures, for example by paying more per unit, for far fewer units). Marginal efficiency and cost-reflectiveness are prioritized over social equity in pricing (Bakker 2005). This also excludes the interests of future customers gained by avoiding the effects of dangerous climate change (Mitchell & Woodman 2010).

Rules and incentives mimic, and are integrated with, the market place as far as possible to encourage market choices to lead to the right answer, rather that favouring a particular technology or outcome (Mitchell & Woodman 2010). However, the focus on consumer protection through (short-term) unit price reduction, and using competitive markets to foster choices, creates a bias against long-term decisions, such as investment in low-carbon technologies. Cost reductions and stakeholder accountability tend to outweigh long-term goals in strategy development. A technology-blind market is also less supportive of new, smaller, diverse entrants, which might undertake new and innovative activities (Mitchell & Woodman 2010).

The focus on competitiveness and cost efficiency per unit delivered creates an environment which favours large incumbent companies who have access to finance and economics of scale. This undervalues the contribution of small and/or local companies and can exclude new entrants from developing new supply or management possibilities (Mitchell & Woodman 2010). For example, wholesale market design and an emphasis on switching in the retail market aim to reduce the power of monopoly companies but can actually 'lock-out' potentially beneficial alternatives, such as development of district energy systems at the urban scale (Bolton and Foxon 2013). This is exacerbated by the fact that business strategies have been shaped by the development of rules and incentives of markets and networks. Any changes to regulation will directly affect bottom line of these businesses and there is likely to be a great deal of resistance to change (Moe 2010). Using the terminology of the multi-level perspective, this reinforces the regime of large, centralised infrastructure companies and locks out alternative niches (Unruh 2000; Unruh 2002). Mitchell describes this as an Iron Band (Mitchell 2010).

---

[9] Including the Office of Gas and Electricity Markets (Ofgem) and the Water Services Regulatory Authority (Ofwat).

The independence of the regulators is ensured through decision making processes based on quantitative analysis using economic and technical criteria (Mitchell & Woodman 2010). However, this excludes other criteria such as diversity of market participants and technologies; development of individual, local and regional benefits, skills and inputs; value added between sectors; innovation; and delivery of sustainable energy. This exacerbates the lock-out of alternative modes of operation, which may compare unfavourably in short-term economic criteria, but offer far greater environmental and social benefits, as well as longer term economic stability.

It has been suggested that the narrow definition of economic efficiency used in regulation may actually deter resource conservation (Bakker 2005; Lovins & Lovins 1991; Lovins 1996). Efficiency in this sense is defined as price per unit of water or energy supplied to consumers. Price controls and regulation of other forms of income (such as in water sector) constrain ability of infrastructure companies to make profit other than through increasing sales of units. Therefore demand reduction on a wide scale could damage profits acting as an incentive against radical action (Mitchell 2014).

The UK Government has not been unable to introduce true competition and cost-reflective pricing (Bakker 2005; Thomas 2006). Competition is only imposed when customers switch regularly, but many domestic customers will not do this (Ofgem 2011). Cost-reflective pricing requires that externalities, such as water quality and climate change adaptation and mitigation, which are currently excluded from the narrow definition of cost efficiency, are internalised (Bakker 2005). This has reduced the effectiveness of market mechanisms and required a significant degree of subsequent regulation to improve water quality standards, support renewable technology deployment, encourage demand management and control pricing.

Infrastructure systems are highly interconnected both through their physical layout and through their interaction with end users (Rinaldi et al. 2001; Roelich et al. 2014a; Hall et al 2012). However, governance systems have not evolved uniformly across infrastructure systems and rarely take such interconnectedness into account. Governance continues to operate in sector-specific silos – synergies and interdependencies are largely ignored (Hall et al 2012).

3.2.2 Policy paradigms

Policy is generally considered to sit within the regime level of MLP; however it is influenced by a series of factors at the landscape level (Whitmarsh 2012). A key influence in the landscape is political ideology, which is under-represented and underexplored in much of the transition literature (Kern & Mitchell 2010). Hall describes the policymaking process as a function of "*the overarching goals that guide policy in a particular field, the techniques or policy instruments used to attain those goal, and the precise setting of these instruments"* (Hall 1993: 278). For instance, if the goal of the policy is to deliver sustainable energy, the chosen instrument might be a feed-in-tariff for renewable energy technologies and the setting would be the level at which tariffs were set. Hall identifies that these processes take place within a policy paradigm, which is "a *framework of ideas and standards that specifies not only the goals of policy and the kind of instruments that can be used to attain them, but also the very nature of the problems they are meant to be addressing"* (Hall 1993: 279). The paradigm of a government can prevent the implementation of appropriate policy to accelerate change or result in policy that constrains beneficial innovations and change.

Changes in policy can happen at three distinct levels; a **first order** change in policy is a change in the setting of instruments, without a corresponding change in the overall policy goal or instruments. A **second order** change is where both the policy instruments and their setting are altered but overall policy goals remain the same. A **third order** change is the simultaneous change in the hierarchy of goals, the policy instrument and the instrument setting. Third order change requires a corresponding change in policy paradigm and occurs relatively rarely (Hall 1993).

A recent example of third order change is the radical shift in ideas underpinning economic policy in the UK from a Keynesian mode of policy making to one based on monetarist economic theory (Hall 1993). The resulting change in paradigm altered the infrastructure policy discourse from one of state control towards liberalisation, private provision and competition (Hall et al 2012). There is a case for government intervention into a privatised infrastructure system over and above that which would be expected into a traditional economic sector. The services that infrastructure provides are essential to social and economic development and delivery of these services can have significant environmental impacts. Regulation is needed to correct the market and system failures that would arise in a purely privatised infrastructure system and to manage a balance between security, sustainability and affordability. However in the new paradigm, direct government intervention was discontinued in favour of independent regulators whose decisions were intended to be separate from the wishes of central Government. Despite this desire for independence, the monetarist paradigm was fundamental to the definition of the remit of regulators in legislation[10]. The purpose of regulators is currently defined in economic terms to reduce ‘market failures’, ensure competition and protect consumers (BIS 2011). This regulatory approach, which replaced state control, is described by Moran as the Regulatory State Paradigm (RSP) (Moran 2003) or by Kern et al as the Pro-Market Policy Paradigm (PMPP) (Kern et al. 2014).

Some argue that the current paradigm has been successful in delivering economic efficiency and protecting the interests of the customers (Helm 2005). However, it is unlikely to be appropriate to respond to emerging issues, like climate change (Hall et al. 2012; Mitchell 2010). It presents a number of specific constraints to sustainable infrastructure transitions.

Helm argues that new policy concerns have emerged, including security of supply and climate change, which have led to a paradigm shift because the change was “*of sufficient magnitude to require rethinking of the role of privatization, liberalisation, and competition in achieving the new priorities, and hence a recasting of energy policy itself*” (Helm 2005: 3). Kern et al suggest that this transformation is incomplete and that, although the narrative and interpretive framework of the paradigm has altered to include climate change and energy security, elements of the belief in market ideas continue to persist (Kern et al. 2014). The increasing prevalence of climate change and energy security goals has led to new institutions (such as DECC) which was established in 2008 following the implementation of the Climate Change Act (HM Government 2008) and new instruments in the UK (such as Feed in Tariffs[11]); however, there remains a great deal of tension between narratives and goals. This is particularly the case between departments which influence the same objective, such as DECC and Ofgem; it could be said that the pro-market narrative is stronger in Ofgem and that economic regulation could be incompatible with DECC’s stronger focus on climate change. As Tutton demonstrates in relation to Ofgem “*the extent to which the market fails to deliver the required carbon reduction and security of supply goals will not be seen as a question of the market not ‘working’…Rather, the problem will be a failure to respond to objectives which are not aligned with financial incentives*” (Tutton 2009: 4).

The multiple perspectives of the current paradigm, as well as the dissonance between narratives and the incompatibility of goals, serve to confuse its influence. It is unlikely that the continued tension between perspectives will be able to deliver the necessary low-carbon

---

[10] Ofgem duties (covering gas and electricity) are defined in the Utilities Act 2000 (HM Government 2000) and the Energy Act 2004 (HM Government 2004). Ofwat duties (covering water and wastewater) are defined in the Water Industry Act 1999 (HM Government 1999) and the Water Act 2003 (HM Government 2003).

[11] Which are well established in Germany and Denmark but have been introduced in the UK only recently.

transition and it is likely that there will need to be a further shift in the interpretive framework (Kern et al 2014). This thesis will explore the constraints that the current paradigm places on alternative modes of operation to explore this point in more detail.

3.3 Summary

Low-carbon technologies rely on critical materials at risk of supply disruption and difficult to substitute (US Department of Energy 2011; Moss et al. 2013). There are different approaches to assessing this risk of disruption by; quantifying the availability of critical materials for a particular technology to determine growth potential over a particular time; or estimating the potential for disruption in the supply of critical materials and the impact of that disruption on the system of interest. The rate of technology roll out is as important as the total number of technologies deployed, because a slow-down in deployment in the early stages of transition could mean that cumulative emissions could exceed safe limits, even if deployment catches up in the end. Furthermore, supply disruption is recognised to be more than just a function of the potential rate of production. In the short-term in particular, geopolitical, environmental and regulatory factors limit the ability to access reserves. The method selected for analysis in this thesis should enable analysis of disruption in the short- and medium-terms and account for these additional factors contributing to disruption. Price is an important determinant of production but the dynamics of this relationship are too complicated and uncertain for the scope of this thesis.

The current policy and regulatory regime has evolved around the requirements and capabilities of the mainstream mode of infrastructure operation and as a result, can actively exclude new actors (Bolton and Foxon 2013). It is highly focussed on competition and unit cost reduction as a means of delivering outcomes for society, which can favour larger actors who can exploit economies of scale. Rules and incentives mimic the market, rather than favouring particular technologies or outcomes despite the fact that using competitive markets creates a bias against long-term decision making (Mitchell & Woodman 2010). The effect of political ideology on policy is under-represented and under-explored in transitions research (Kern and Mitchell 2010). In reality, policy is affected by the framework of ideas and standards (the policy paradigm) that specify not only the goals of policy and the kind of instruments used but also the nature of the problem they are meant to be addressing. The paradigm of governments can prevent implementation of appropriate policy to accelerate change or can result in policy that can constrain beneficial innovation and change. Therefore, the policy paradigm must be considered in analysis in this thesis.

4 Methodology

This chapter describes the rational for choosing a mixed-method approach to analysing constraints to infrastructure transition and justifies the specific approaches used for analysis of two constraints to low-carbon infrastructure transition. An important aspect of this thesis is the comparison of different approaches to the analysis of constraints. Therefore, the development of appropriate methods responds directly to two of the research questions, namely: *how can we conceptualise the ways in which the challenges of critical material supply disruption and governance present constraints to low-carbon infrastructure transitions?* and; *what methods are most appropriate to analyse constraints from critical material supply disruption and governance to low-carbon transitions?* This chapter describes two different approaches to conceptualising and analysing the constraints to low-carbon transitions. The experience of developing and applying these approaches addresses the aim of the thesis to contribute to methodological development for assessment of constraints, and will form part of the thesis discussion.

4.1 Rationale for mixed-methods approach

The purpose of the empirical part of this thesis is to develop a better understanding of the nature and scale of potential constraints to low-carbon infrastructure transition and to identify potential responses to these constraints. The two constraints analysed in this thesis are very different in nature and therefore; the most appropriate approach to analysis will differ between cases. Habermas argues that social enquiry can divided into three categories, each of which produces its own form of knowledge (Habermas 1972). These three categories are summarised in table 4.1. The division of approaches to enquiry in this way is not without criticism (Keat & Urry 1982) but it provides the means to understand how an approach, which is driven by underlying interests and is examining certain parts of social system, can result in different types of knowledge.

Table 4.1 Habermas' three categories of scientific enquiry adapted from (Blaikie 2008)

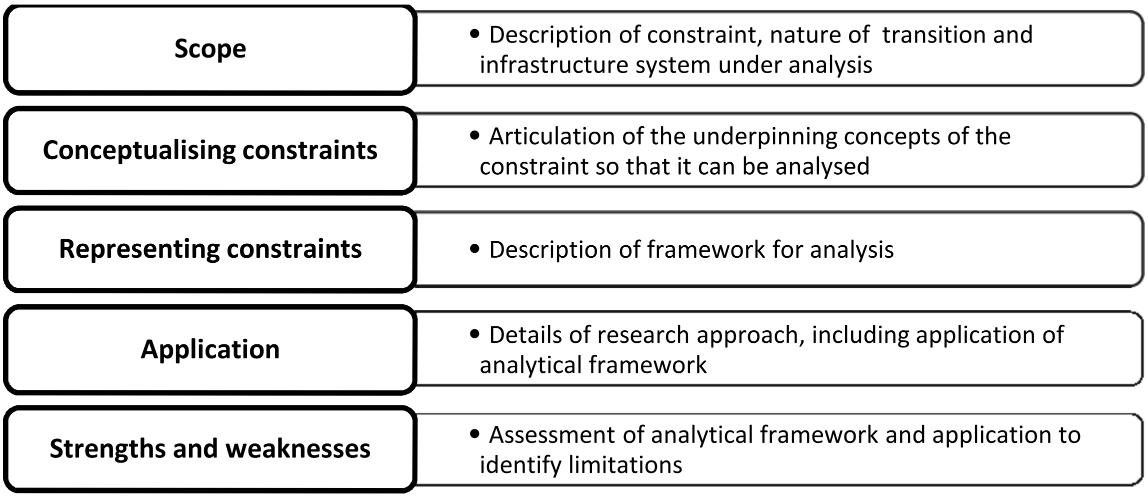
Type of Science	Underlying interests	Aspects of social existence
Empirical-analytic	Prediction and control	Ways in which individuals control and manipulate the environment
Historical-hermeneutic	Understanding	Interaction between individuals and between social groups
Critical theory	Emancipation	Power

The analysis of constraints from material criticality focuses on man's ability to extract natural resources at a sufficient rate to meet the needs of society (in this case the need for specific roll-out of low carbon technologies). As such, Habermas' categorisation would suggest that an empirical-analytical approach could be most appropriate to examine this constraint. Furthermore, the purpose of the analysis is prediction (of supply risks); therefore, this would lend itself to evaluation in terms of objective criteria, which may be quantified (Blaikie 2008). The constraint is centred on a technology-based transition with clearly defined scenarios of technology roll-out, and available numerical data relating to material production and requirements. This further supports the case for a more quantitative approach to analysis that aims to enumerate the risk of constraints and compare the relative risks of different low-carbon transition pathways.

The second constraint, of policy and regulation, will focus on a transition in operation of infrastructure, which addresses relationships between actors in different parts of the infrastructure systems. This analysis needs to provide a more complex understanding of human interaction and lends itself to a more deliberative, qualitative approach (Ezzy 2002). The analysis appropriate to this constraint would fall into Habermas' category of 'historical-hermeneutic' and requires a more conceptual and theoretical understanding of socio-technical dynamics. With this in mind, it is argued that a case study approach is most appropriate to develop and test a conceptual framework of system interactions. This allows the development of a broader and more holistic account of the complex social processes associated with infrastructure operation (Yin 2009).

An interesting point to reflect on will be the different forms of knowledge created by these contrasting approaches to analysis and the extent to which they are able to lead to the third form of knowledge. One of the core aims of this thesis is to identify responses to constraints which accelerate the process of infrastructure transition, or emancipation to use Habermas' terminology. Therefore, a key point of discussion will be the limitations of different forms of enquiry in achieving a more critical-theoretical understanding of the system of interest.

A more detailed account of the development of specific methods is presented in sections 4.2 and 4.3 for critical material constraints and policy and regulatory constraints respectively. The approach to method development is summarised in figure 4.1 and is used to ensure that the development and application of methods are systematic. This enables a more effective comparison of the contribution of different approaches.



Scope	• Description of constraint, nature of transition and infrastructure system under analysis
Conceptualising constraints	• Articulation of the underpinning concepts of the constraint so that it can be analysed
Representing constraints	• Description of framework for analysis
Application	• Details of research approach, including application of analytical framework
Strengths and weaknesses	• Assessment of analytical framework and application to identify limitations

Figure 4.1: Summary of method development

4.1 Constraints from critical material supply disruption

4.1.1 Scope of analysis

This constraint assesses the implications of critical material supply disruption for the transition to low-carbon infrastructure. It uses a transition from carbon-based to low-carbon energy generation technology as the focus of its analysis. The analysis is applied to the energy system, which presents an example of significant change in technologies which are highly reliant on critical materials (US Department of Energy 2011; Moss et al. 2011).

4.1.2 Conceptualising critical material constraints

Critical material constraints are conceptualised in this thesis as the potential for disruption in the supply of critical materials and the impact of that disruption of the system of interest, rather than a conceptualisation based solely on material availability. This is because of the significant uncertainty associated with reserve and resource estimates (Tilton & Lagos 2007; European Commission 2010b) and forecasts of future production (Speirs, R. Gross, et al. 2013). Furthermore, this choice is supported by the emerging awareness that short- and medium-term supply disruption is more likely to be problematic than long term depletion (European Commission 2010b) and serious implications of short-term reductions in technology roll-out on cumulative emissions. As described in Chapter 3, critical material supply disruption can result from limitations on expansion of production, market imbalances or governmental intervention, as well as physical scarcity (Erdmann & Graedel 2011). It is recognised that commodity price will influence the production of critical materials but this has been excluded from this analysis owing to the complex dynamics of this relationship.

This thesis will thus operationalize critical material constraints as the potential for disruption of the supply of critical materials to slow down or halt the planned installation of low-carbon technologies in the UK necessary to achieve UK emissions reductions targets.

4.1.3 Representing critical material constraints

The nature and scale of critical material constraints are analysed through evaluation of objective criteria (in this case criticality); to allow quantification of the constraints faced by different low-carbon transition pathways. A quantitative approach supports analysis of the scale of constraints and enables comparison of different pathways. A dynamic approach to analysis is used to allow quantification of critical material constraints over the duration of the pathway to identify when significant changes in criticality occur.

In accordance with recent assessments of material criticality, the objective criteria used to support analysis in this thesis incorporates measures of both supply disruption and the impact of that disruption (European Commission 2010b; US Department of Energy 2011; Morley & Eatherley 2008; National Research Council 2008; Graedel et al. 2012). Supply disruption represents the potential for the supply of the material of interest to be disrupted by a range of geological, political and environmental factors. The literature review identified that impact has been represented in two principal ways in the literature; as the exposure of the system of interest to supply disruption (making criticality consistent with risk) (European Commission 2010b); and as a combination of exposure and the ability of the system of interest to respond to that exposure (making criticality consistent with vulnerability) (Graedel et al. 2012).

Endogenising potential responses to exposure, such as substitution or recycling, would reduce the facility of the approach to identify policy intervention opportunities in support of these responses. Since one of the aims of this thesis is to identify potential policy responses, a representation of exposure as vulnerability is not preferred in this instance. In this research, the former representation of impact as exposure of the system to disruption is adopted, which is analogous to that of risk analysis. Material criticality constraints are represented as the risk faced by the low-carbon infrastructure transition from critical material supply disruption.

The analysis of risk is a well-established process and policy-makers are very familiar with risk-based decision making approaches, which increases the potential to engage policy makers (Hall et al. 2005). Risk is conventionally defined as the product of the probability of an event and the severity of harm resulting from that event (Renn 2008). In this context, exposure is represented as the effect of supply disruption on the goal of low-carbon infrastructure transition, analogous to the concept of severity, whereas the probability is the likelihood of supply disruption. Determining the risk of constraints from critical materials, rather than vulnerability, can be expected to provide more insight into the potential interventions to reduce constraints.

Typically, risk-based approaches evaluate the risks of all decision options in quantified terms and identify preferred options and mitigation activities. It is important to recognise the contrasting nature of different factors that contribute to likelihood and severity as well as the difficulty associated with quantification of these factors (Royal Society 1992). Therefore this thesis will develop an indicator set to represent the multiple dimensions of likelihood and severity necessary to quantify criticality.

The literature review identified the need for dynamic analysis of criticality to reflect the dynamic nature of the contributing technical, socioeconomic and environmental factors. A comprehensive dynamic analysis would require assessment of the interactions and feedbacks between these different factors of criticality (Knoeri et al. 2013). This would increase the complexity of analysis and would require advanced modelling, which is beyond the scope of this thesis. However, indicators will be produced as forecast time series to reflect the time-dependent nature of contributing factors. The expected relations between these indicators and related system characteristics are represented to some extent in the aggregation of individual indicators to create a composite indicator.

4.1.4 Using Indicators to assess critical material constraints

In order to quantify the objective criteria for analysis (criticality), and to represent the multiple facets of supply disruption potential and impact, this thesis will develop an indicator set that, when combined will represent criticality, which will enable assessment of the scale of constraints. Individually, indicators will represent the individual dimensions of criticality, which will provide insight into the nature of material criticality constraints.

Indicators are symbolic representations designed to communicate a property or a trend in a complex system or entity. They allow measurement of the trajectory towards a particular goal, or in the case of this thesis, the change in constraints on the goal of low-carbon energy system transition. Indicators are widely used in policy analysis because they can integrate many complex issues while providing simple signals. The European Commission based its assessment of criticality on a series of indicators used to "*evaluate some risks and the potential impact on the economy of potential supply bottlenecks or decreased availability of the raw materials*" (European Commission 2010: 23). If well designed, they can condense information for rapid assimilation, while making it possible to explore issues in greater detail (Moldan & Dahl 2007).

Most indicators are derived from quantitative measures of variables which represent the system of interest. However, some phenomena, particularly social phenomena, are best represented qualitatively using scales, barometers or colour coding. When multi-dimensional concepts cannot be captured in a single indicator, a composite indicator can be used, where individual indicators are compiled into a single index on the basis of an underlying model (OECD 2008). It is considered that the multi-dimensional representation of criticality is best quantified using a composite indicator, as long as there is transparency with regard to contributing indicators. This thesis will use an iterative process of indicator development building on Moldan and Dahl (2007) and the Organisation for Economic Co-operation and Development (OECD) (2008). This is described in table 4.2.

Indicator design necessarily involves trade-offs between technical feasibility, indicator usability and consistency of indicators with system of interest. Therefore the process of indicator development will be more iterative than implied by the linear process described in table 4.2.

Table 4.2: Summary of indicator development method

<table>
<tr><th>Step</th><th>Activity</th><th>Reason</th></tr>
<tr><td>1. Characterising system</td><td>Identify the characteristics or processes that affect supply disruption</td><td>To ensure appropriate representation of phenomenon concerned</td></tr>
<tr><td colspan="3">This characterisation is necessarily a simplification of the real (and complex) system of critical material constraints. It captures the main driving forces of supply disruption and articulate the means of exposure of the goal of low-carbon energy system transition to this disruption. The factors affecting supply disruption are complex and non-linear. Using systems science to identify interactions and processes as well as characteristics can reduce subjectivity and increase the relevance of indicators. This process also provides the basis for the combination of indicators by analysing the relationships between indicators (OECD 2008: 15).</td></tr>
<tr><td>2. Identifying data</td><td>Collating quantitative data that is able to represent important system characteristics</td><td>To ensure measurability of the indicators which represent the system</td></tr>
<tr><td colspan="3">It is important that indicators are measureable and readily quantifiable without the need for expert input. The quality of indicators inevitably depends on the underlying data that are used to compose them. Data availability is often a selection criterion to ensure a rigorous quantitative underpinning for the indicators (Moldan and Dahl 2008: 10); however, it is important to note that available data may not represent the most important aspects of the system so may limit the relevance of analysis.</td></tr>
</table>

<table>
<tr><th>Step</th><th>Activity</th><th>Reason</th></tr>
<tr><td>3. Considering scale</td><td>Determine appropriate indicator boundary and timescale</td><td>To improve policy relevance</td></tr>
<tr><td colspan="3">The timescale of indicators and system boundary must reflect the scales most commonly used in policy making. The majority of infrastructure scenarios have national scope and run to 2050, therefore this timescale is used for this study. However, technology roll out is forecast to increase dramatically in the medium term so results should be provided as a time series with appropriate time steps.</td></tr>
<tr><td>4. Defining indicators</td><td>Identify individual indicators that explain principal system characteristics</td><td>To enable trends in separate indicators to be analysed in detail if necessary</td></tr>
<tr><td colspan="3">Further simplification is required to identify variables capable of representing the significant system characteristics. The expected relations between these indicators and related system characteristics inform the next step of construction of a single aggregate indicator. At this stage, the feedback loops between indicators are excluded, but they inform the analysis of the results.</td></tr>
<tr><td>5. Construction of a single composite indicator</td><td>Develop methodology to aggregate or combine indicators into an index</td><td>To simplify the interpretation of a composite indicator, rather than identifying common trends across many separate indicators</td></tr>
<tr><td colspan="3">The understanding of linkages between indicators is used to develop an approach to combination. The combination process is transparent with sound statistical or conceptual principals, to avoid misleading policy messages (OECD 2008).</td></tr>
</table>

The specific details of the indicators developed to support analysis in this thesis are presented in Chapter 5: Material criticality constraints.

4.1.5 Strengths and weaknesses of constraints analysis using indicators

Indicators, and in particular composite indicators, have a number of strengths in the assessment of material criticality constraints. Composite indicators can summarise complex, multi-dimensional systems making data more accessible to decision makers and are easier to interpret than many, separate indicators (OECD 2008). In this way it is possible to use more information to assess a particular problem and enable users to compare complex dimensions effectively. Using an understanding of the theoretical framework within which the indicators sit to devise their combination method can better represent the interactions of indicators.

The European Commission recognises that indicator selection is not an exact science and is subject to challenges of data availability and challenges about indicator aggregation and combination (European Commission 2010: 23). Furthermore, the selection and weighting of indicators could be the subject of political dispute. Indicators may lead to inappropriate policies if dimensions of performance that are difficult to measure are ignored or composite indicators are poorly constructed or misinterpreted (OECD 2008). Moreover, lack of transparency in construction may disguise serious problems in some dimensions and increase the difficulty of identifying proper remedial action. Therefore, the construction process must be transparent and have sound conceptual principals.

The aggregation of indicators into a composite indicator represents the relationships between individual indicators to some extent, but is not able to fully capture the dynamic nature of these relationships, which would require more advance modelling. This reflects the challenge of achieving a balance between understanding the complexity of the situation and efficiently identifying effective intervention.

4.2 Constraints from policy and regulation

Constraints from policy and regulation are analysed using case study analysis and, in particular, Eisenhardt's theory building through case study analysis methodology. The method of theory building is described in detail in section 4.3.6 but first the analytical framework used in support of analysis is described, along with the individual cases used to represent alternative modes of infrastructure operation.

4.2.1 Scope of analysis

This constraint explores the challenges presented by current policy and regulation to the transition to alternative modes of infrastructure operation. Analysis is limited to the water and energy systems, which require radical transformation and are subject to a great deal of policy intervention and regulation.

4.2.2 Conceptualising governance constraints

The current policy and regulatory systems in the UK[12] have co-evolved with the physical and operational infrastructure systems. The dominance of the throughput-based mode of operation in energy and water infrastructure operation, where transactions (and associated profit) are based on units of products delivered, is reinforced by policy and regulatory systems which rely on market-based mechanisms to drive investment in infrastructure and reduce the unit cost of products (Mitchell & Woodman 2010). Systemic interactions between regulations and large, centralised incumbent infrastructure operators lock us into this mode of provision and constrain the development and adoption of alternative modes of operation (Bolton & Foxon 2013).

---

[12] The organisations, interventions and regulations that influence the interaction between infrastructure providers and end users.

The alternative modes of infrastructure operation described in chapter 1 are limited to predominantly isolated examples that do not challenge the mainstream. Some are so constrained by the mainstream that they struggle to even exist; they are what is termed 'locked out' (del Rio & Unruh 2007). In order to overcome lock-out of innovative modes of infrastructure operation they must first survive, then evolve to overcome the constraining influence of regimes in the short-term and drive transformations in the regime structures over the longer term (Smith et al. 2010). Breaking lock-out is not an instantaneous or spontaneous change (Unruh 2002). Seyfang and Haxeltine describe three ways that innovations could evolve (Seyfang & Haxeltine 2012):

- Scaling out (or replication) – proliferation of the mode of operation at a similar geographic scale and scope so that it becomes more numerous.
- Scaling up – changing the mode of operation so that it encompasses a larger geographic area or a wider scope of services or infrastructures.
- Translation – adoption of the mode of operation in another sector or setting.

This thesis conceptualises governance constraints as processes of policy or regulation that prevent alternative modes of innovative operation from evolving in one of these ways.

4.2.3 Analytical framework (representing governance constraints)

In order to assist with the development of a conceptual framework of system interactions through case study analysis an analytical framework has been created to structure the investigation and to help make sense of empirical data. The framework is based on the MLP and co-evolutionary approaches to analyse innovations in infrastructure operation that were introduced in chapter 2. This thesis does not aim to integrate or hybridise the two approaches into a universal theory, rather to explore the crossovers. In this sense it takes a similar perspective to Hargreaves et al (2013) who explore the crossover between the MLP and practice theory.

The distinction between different levels of analysis in the MLP is useful in identifying constraints that policy and regulation (part of the regime) place on innovations in modes of operation (within niches) and in particular the influence of the current policy paradigm (the landscape) on the role and approach of policy makers and regulators. However, the structural focus of the MLP tends to underplay evolutionary aspects which are critical to the conceptualisation of policy and regulatory constraints (Smith et al. 2005; Foxon 2011). To overcome this limitation the analytical framework used in this thesis also draws on Foxon's co-evolutionary framework (Foxon 2011). This provides a greater focus on the interactions and relationships between actors. The inclusion of co-evolutionary aspects provides a better representation of infrastructure systems as complex adaptive systems, where changes emerge as a result of actor interactions. The two dimensions are represented in figure 4.2.

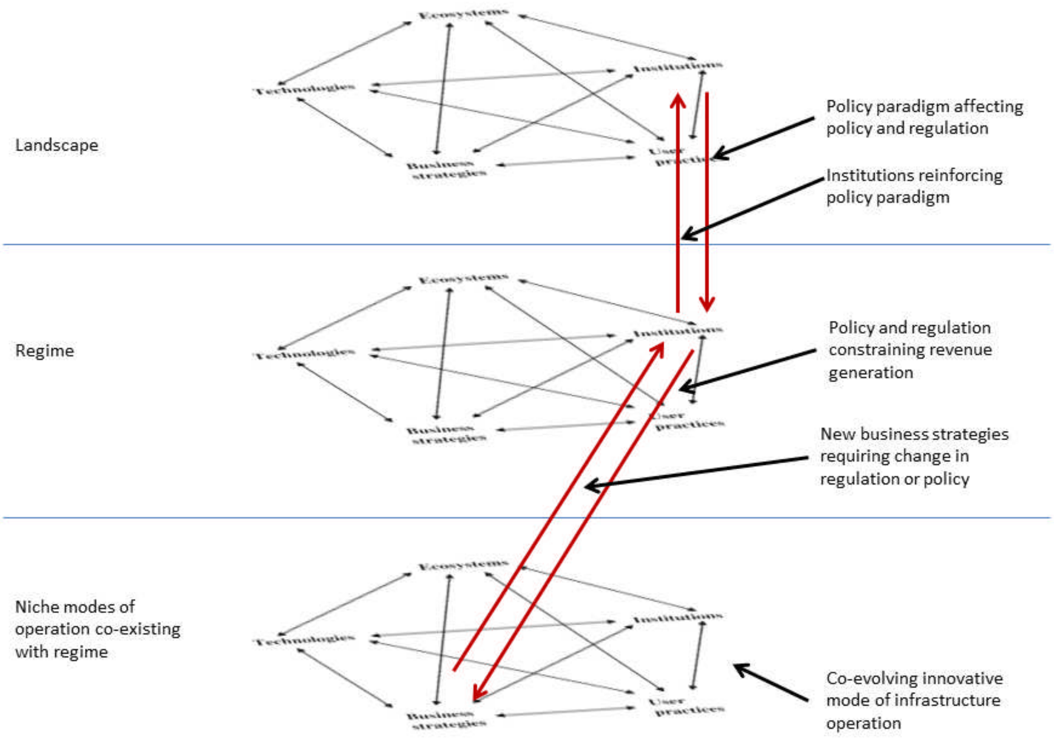


Figure 4.2: Analytical framework showing horizontal dimension of co-evolutionary framework and vertical dimension as inter-level interactions. Interactions have been illustrated for the institutional system in the landscape and regime but it is recognized that there will be interactions between other systems.

4.2.4 Case study of alternative modes of infrastructure operation

Rather than analyse one mode of alternative operation in detail, this thesis analyses a series individual cases as examples of alternative modes of operation. In this sense, the individual cases contribute to the meta-case study of alternative modes of operation. This approach is taken to reflect the fact that the future infrastructure system must retain diversity to avoid future lock-in and become more flexible to future change (Berkhout 2002; Stirling 2009). Achieving this diversity will require that a range of modes of operation, involving a variety of actors not traditionally associated with infrastructure provision, are able to co-exist. Studying one mode of operation in isolation would not enable analysis of the conditions which constrain diversity or how policy and regulation might better encourage this diversity.

Therefore, a series of different case studies are selected to represent different actors, with different motivations and different means of achieving those motivations (as represented in Figure 1.3). Where possible, each of the typologies of alternative modes of operation identified in figure 1.3 is represented by one of the individual cases. In that way, this thesis attempts to analyse the constraints that policy and regulation place on moving towards a system where any or all of these modes of operation could co-exist.

The principal characteristics of the nine individual cases contributing to the meta-case study are summarised in table 4.3. This demonstrates the purposeful sampling used to cover a range of initiators and revenue delivery mechanisms. Cases were also selected to represent projects at different stages of completion from those that had been operating for some years (such as Woking Borough Council) to those which were still in the planning phases (such as Welborne and Byker estate). This provided useful insights into experiences at different stages of developments but was not without challenges.

The number of interviews undertaken for each individual case varied depending on the detail available from the literature. Some cases had detailed evaluation reports associated with them and were also the subject of detailed analysis in parallel projects; therefore, less primary data was necessary to develop an understanding of the relationships, resources and constraints. A full list of interviewees and interview dates is provided in Appendix 1.1.

Table 4.3: Summary of case studies used in support of this research

Case study	Description	Market sector	Mode of operation	Initiator	No inter-viewees
Woking Borough Council	A Special Purpose Vehicle (including participation of the private sector) designed to save energy in council property and recycle financial savings	Commercial, public sector & domestic	Heat and electricity supply plus efficiency	Local authority	0[13]
Olympic Park Energy Centre	The largest Combined Heat and Power Plant in the UK providing heat to the Olympic park. Now extending heat supply to surrounding developments.	Domestic & commercial	Heat supply	Government Agency	2
Yorkshire Water Water Saving Trials	Pilot schemes trialling water saving device roll out to provide evidence on water saving potential in response to Ofwat's requirements to provide demand and supply balance and demonstrate action on demand management.	Domestic	Water saving	Utility Company	1[13]
Kimberley Clark Water Recycling	"Long loop" recycling of effluent from paper mill re-used as process water following intensive treatment. Novel financing mechanism (piggy-backing on a high payback scheme).	Industrial	Water saving	Private sector	1
Eco-Island	Retrofit and low-carbon energy schemes on the Isle of Wight. Planned to be undertaken by an ESCo run in partnership with the Community Interest Company.	Domestic & commercial	Energy service	Community Interest Co.	2[14]
Welborne	Proposed development of 3,000 house and employment land. In addition to high levels of sustainability (and updatability), it is proposed that properties and infrastructure will be managed by the community that occupies the site.	Domestic & commercial	Energy, water and transport service	Landowner	4
Chale Green Housing Trust	Retrofit of social housing using novel financing scheme to 'rent' roof space to a solar panel provider. Insulation and air source heat pumps were provided to increase thermal comfort and reduce energy bills.	Domestic	Heat and electricity supply plus efficiency	Social housing trust	1[13]
Newcastle City Council	Joint venture to deliver district heating networks in council and private property. Less profitable schemes bundled with more financially viable schemes.	Domestic & commercial	Heat and electricity supply	Local authority	1[13]
Byker Community Trust	Transfer of estate district heating system from local authority to community trust	Domestic & public sector	Heat and hot water	Community trust	2

---

[13] A lower number of interviews was undertaken because a detailed evaluation report was available: Woking (Thorp 2011), Yorkshire Water (Waterwise 2011), Chale Green (DECC 2012b), Newcastle (Jones 2012).

[14] One interview was not used for ethical reasons. See section 4.2.9 for full explanation.

4.2.5 Application of analytical framework

The analytical framework is applied to the case studies in the following way; firstly relationships, capabilities and process that contributed the development or stability of the case study are identified through a process of literature review and interviews. Particular attention is paid to identify these aspects across the five systems in the framework, and their co-evolution. It is anticipated that most of these relationships will be confined to the niche level of the analytical framework but positive relationships between regime actors and the niche are identified. These relationships, processes and capabilities are considered to be characteristics of the alternative mode of operation, which differ from the mainstream mode of operation. As a result, alternatives have different resources and capacity to those required to respond to selection pressures and enable transitions described by Smith et al (2005).

The second aspect of analysis identifies how processes in the regime level or relationships between the niche and regime limit the potential of alternative modes of operation to scale out, scale up or translate to other sectors. This addresses a limitation of the MLP, which considers that niches are protected from selection pressures. The framework illustrates how the development and evolution of alternative modes of operation are shaped by the regime. This interaction is represented by the arrows between the regime and the niche in figure 4.2. The analytical framework considers that institutions in the regime level can constrain alternative modes of infrastructure operation in two ways; by limiting their capacity to respond to selection pressures (by not supporting the necessary resources) and by creating negative selection pressures (directly constraining innovations). This provides greater insight into how policy might respond to remove constraints.

Importantly this thesis includes a third element of analysis, which recognises the influence that the policy paradigm, in the landscape, has on the institutions in the regime. In this way it draws on the work of Hall (1993), Mitchell (2010) and Kern (2010) that stresses the constraints that the policy paradigm place on the ability of policy to enable transitions. The policy paradigm is considered in both the analysis of constraints (as a driving force behind policy) and as a limitation to implementing policy responses.

4.2.6 Theory building from case study research

The analytical framework is applied to the case studies using the “theory building from case studies” approach advocated by Eisenhardt. This method is ‘*appropriate in the early stages of research on a topic*’ (Eisenhardt, 1989: 548) and is particularly relevant for research into constraints to alternative modes of infrastructure operation, which has received little research attention to date. This is an inductive method of analysis which uses qualitative data generated from case studies to elaborate theory in an iterative process, which draws on the design of case study research and grounded theory, amongst others (Yin 2009; Glaser & Strauss 1967). Grounded theory emphasises the role of data and observation in generating theory. The approach taken in this thesis is more representative of the Straussian form of Grounded Theory, which recognises the role of pre-existing theory in shaping interviews and identifying insights from empirical data, rather than relying on data alone (Ezzy 2002; Corbin & Strauss 2008).

The theory building from case studies method relies on the continuous comparison of data and theory, emphasising the emergence of theory from evidence and an incremental approach to case selection and data gathering. A theory is used by an analyst to specify which working parts of a framework are considered useful to explain diverse outcomes and how they relate to one another (Ostrom 2010). The goal of this analysis is to identify generalizable theories which explain why alternative modes of operation may or may not survive and prosper in the current mainstream system, and to identify the most effective points of influence for policy intervention or regulatory change.

The key stages of the approach are described, along with specific steps taken by the author, in table 4.4.

Table 4.4: The process of building theories from case studies

<table>
<tr><th>Step</th><th>Activity</th><th>Reason</th></tr>
<tr><td>1. Getting started</td><td>Definition of research question</td><td>Focuses efforts</td></tr>
<tr><td></td><td>Possibility of <i>a priori</i> constructs</td><td>Provides better grounding of construct measures</td></tr>
<tr><td></td><td>Neither theory nor hypothesis</td><td>Retains theoretical flexibility</td></tr>
<tr><td colspan="3">An initial definition of the research question is essential to set the focus for the research and support the identification of both potential case studies and the data to be gathered during case study analysis. The principal research questions for this stage of the research project are <i>“what processes and relationships have supported the development of alternative modes of infrastructure operation?”</i> and “<i>how have processes in the regime affected this process of development and the potential for scaling out, up and translation of alternative modes of infrastructure operation?”</i>. The aim of the case study analysis is to identify constraints to the success of these alternative configurations to provide insight about the most effective points of influence for policy intervention.</td></tr>
</table>

Step	Activity	Reason
2. Selecting cases	Specified population	Constrains extraneous variation and sharpens external validity
	Theoretical, not random, sampling	Focuses efforts on theoretically useful case studies –those that replicate or extend theory
A systematic literature review is undertaken to identify examples of alternative modes of operation in the UK, such as demand management, local authority provision or community provision. Examples are gathered across water and energy at a lower level of detail to inform selection. Case studies are chosen for theoretical, rather than statistical reasons. Ideally, an equal number of examples from water and energy would be selected for detailed analysis. Examples are selected that were initiated by different partners to assess the potential to maintain diversity in modes of operation.		
3. Crafting instruments and protocols	Multiple data collection methods	Strengthens grounding of theory by triangulation of evidence
	Qualitative and quantitative data	Synergistic view of evidence
	Multiple investigators	Fosters divergent perspectives and strengthens grounding
A combination of literature review and semi-structured interviews is used to strengthen the grounding of theory development and provide a synergistic view of evidence. The events that define the case study are collated into a timeline. The timeline provides a deeper understanding of the evolution of internal and external factors that lead to the success of the alternative mode of operation. The observed structure of the agents and assets involved in the alternative mode of operation is mapped and compared with the structure of a mainstream mode of operation. The mapping exercise provides a clear view into the relationships and interactions that contribute to the emergence of the case study within the mainstream. This yields insights into the influence of the regime on the development of alternative modes of operation. In order to generate additional, primary data, a number of semi-structured interviews is conducted for case studies. The analytical framework is used to structure high-level interview questions to ensure that all relevant potential constraints to scaling-out, scaling-up or translating the model to other infrastructure sectors.		
4. Entering into the field	Overlap data collection and analysis, including field notes	Speeds analysis and reveals adjustments to data collection
	Flexible and opportunistic data collection methods	Allows investigators to take advantage of emergent themes and unique case features
Data collection is undertaken progressively, with analysis of initial case studies undertaken before completion of data collection for subsequent cases. This allows adaptation of the interview protocol to 'probe particular themes that emerged' (Eisenhardt, 1989: 539). The use of semi-structured interviews allows the interviewer to delve deeper into emergent themes. This helps the researcher to understand each case individually and in as much detail as is feasible. This is not unsystematic, rather it allows the researcher to take advantage of the emergence of new themes and improve resultant theory.		

<table>
<tr><th>Step</th><th>Activity</th><th>Reason</th></tr>
<tr><td>**5. Analysing data**</td><td>Within-case analysis
Cross-case pattern search using divergent techniques</td><td>Gains familiarity with data and preliminary theory generation
Forces investigators to look beyond initial impressions and see evidence through multiple lenses</td></tr>
<tr><td colspan="3">A two-stage, structured approach to data analysis is used to draw out key findings from the wealth of data generated during the collection phase.
Within case analysis: Interview transcriptions are manually coded to segregate data relevant to the five systems described in Foxon's co-evolutionary framework. Detailed write-ups are produced for each case to describe the background to the case, its evolution and its potential future development. An example of an individual case write-up is provided in Appendix 1.2. This includes insights from literature review, timeline and mapping as well as from interview analysis. This improves familiarity with the data and supports preliminary theory generation; identifying the motivations, enablers and constraints for the development of the alternative modes of operation and the perceived constraints to scaling up, scaling out or translating of the case. This analysis provides preliminary hypotheses of potential constraints to alternative modes of infrastructure operation.
Cross- case pattern search: In the second stage of analysis, cross-case pattern searches is used to overcome information processing biases. The cases are compared to identify similarities and differences, both in the characteristics of alternative modes of operation and in the constraints from the regime. This juxtaposition of contrasting cases forces the investigator to look beyond initial impressions and see evidence through multiple lenses. It can break simplistic frames and lead to a more sophisticated understanding of case study results.</td></tr>
<tr><td>**6. Shaping hypotheses**</td><td>Iterative tabulation of evidence for each construct
Replication, not sampling, logic across cases
Search evidence for 'why' behind relationships</td><td>Improves construct definition, validity and measurability
Confirms, extends and sharpens theory
Builds internal validity</td></tr>
<tr><td colspan="3">Emerging theories from each case are tabulated and systematically compared with the evidence from other cases to *" assess how well or poorly it fits with case data"* (Eisenhardt, 1989: 541). Importantly replication logic is used and theories are compared with each case, not the aggregate cases to ensure that the resultant theory holds for each individual case and not just for the summary of the findings (Eisenhardt, 1989: 542). Qualitative data is particularly relevant at this stage to *uncover "the underlying theoretical reasons for why relationships exist"* and therefore establish internal validity (*ibid.).*</td></tr>
</table>

<table>
<tr><th>Step</th><th>Activity</th><th>Reason</th></tr>
<tr><td>7. Enfolding literature</td><td>Comparison with conflicting literature
Comparison with similar literature</td><td>Raises theoretical level, sharpens construct definitions and generalizability</td></tr>
<tr><td colspan="3">Emerging theories are compared with both conflicting and consensual secondary literature to enhance the ‘internal validity, generalizability, and theoretical level of theory-building from case study research’ (Eisenhardt, 1989: 545). This is crucial in case study research because findings are often based on a limited number of cases. Emerging findings of the case study are compared with theories and results in existing literature. Therefore, findings presented in Chapter 8: Policy and regulatory constraints include both empirical findings and supporting findings from the literature.</td></tr>
<tr><td>8. Reaching closure</td><td>Theoretical saturation when possible</td><td>End process when marginal improvements become small</td></tr>
<tr><td colspan="3">The process of adding new cases and iterating between theory and evidence is ended when theoretical saturation has been achieved, that is, <i>“the point at which incremental learning is minimal because the researchers are observing phenomena seen before”</i> (Eisenhardt 1989: 545). However, Eisenhardt acknowledges, <i>“theoretical saturation often combines with pragmatic considerations such as time and money</i>” (ibid.). Nine cases have been included in this thesis to represent a broad range of modes of operation, actors and organisations. This aimed to widen the generalizability and heighten the theoretical level of the findings.</td></tr>
</table>

4.2.7 Shaping hypotheses

Early hypotheses were tested at a series of deliberative workshops to enable feedback from a range of stakeholders, including those involved in the cases themselves. Deliberative workshops are structured, facilitated group discussions which encourage participants to consider an issue in depth. They are widely used in qualitative research exploring perceptions in emerging areas of science and technology and increasing public participation in these issues (Corner et al. 2013; Chilvers 2010; CST 2005). Workshops are specifically structured to create deliberations between participants with different kinds of experience and expertise (Evans & Kotchetkova 2009).

The first workshop was a stakeholder workshop attended by 28 representatives from industry, policy and academia[15]. Participants were invited to represent a broad range of expertise and institutions and to ensure a diverse range of viewpoints would be included in the workshop. Participants were engaged in two activities which required them to critique emerging findings from case study analysis. In the first activity participants were provided with a selection of constraints identified during case study analysis, along with a series of actions proposed to reduce or remove constraints. They were asked to identify additional constraints that they had

---

[15] This was organised by the whole project team and included activities from the wider project within which this research sits. Specific sessions were planned and managed by the author to address the findings of this thesis.

experienced or perceived and identify high level actions required to remove them. The second activity required participants to identify the specific activities needed to implement these actions and the parties that needed to be engaged in the process. Activities were facilitated by a researcher and took place over one day, interspersed with presentations about the research findings.

The second event was a more focussed workshop held by the Core Cities, a group of the eight largest cities outside London. The results of the analysis, relevant to local authority infrastructure delivery, were presented to representatives from the eight cities. Specific questions were posed to shape subsequent discussions related to the specific regulatory barriers they had faced and the barriers that were most important to them.

The deliberative workshops provided an important opportunity for participants to discuss their experience and perceptions of alternative modes of infrastructure operation and the constraints posed by policy and regulation. It became apparent from discussions that the policy and regulatory environment was so complex that it was extremely difficult for participants to pin down specifics. It was this complexity of constraints that presented the most significant challenge to potential alternative modes of operation. This identified an important lesson for this research; although important to understand the complexity of the constraints it was perhaps more important to articulate the specific implications of constraints and potential responses. This is a challenge when the system under analysis is complex and neither cause and effect, nor response to intervention, are linear.

4.2.8 Strengths and weaknesses of methodology

The principal advantage of case study research to this thesis is its ability to retain "*the holistic and meaningful characteristics of real-life events"* (Yin 2009: 4). It allows the researcher to investigate relationships and processes, rather than frequencies and incidents and provide explanatory detail to support theory development. Case studies also allow researchers to use a variety of evidence and data, both qualitative and quantitative. Theories generated through this method tend to be novel as a result of the juxtaposition of contradictory and paradoxical evidence. This can help to overcome researcher bias and can challenge initial impressions. Emergent theories are testable and empirically valid because they have been generated through an iterative process of theorising and testing with evidence and existing literature.

Common concerns about case study research include lack of rigour and difficulties with generalising results (Yin 2009; Blaikie 2008). Concerns of lack of rigour relate to lack of systematic processes and allowing equivocal evidence or biased views to influence research findings. This is addressed in this methodology by implementing a clear and structured case study protocol and by analysing a number of contrasting cases to reduce the influence of biased views (Yin 2009). Generalising results is addressed through cross case analysis of multiple, contrasting cases and comparison of emerging results with the literature (Eisenhardt 1989). This goal of this thesis is to achieve analytic generalisation (to expand and generalise theories) not statistical generalisation (enumeration of frequencies).

Building theories from case studies has a number of limitations, which must be recognised during analysis of results. The approach can produce complex theory that tries to capture the detail of all the empirical evidence used; it is often hard to assess which relationships are the most important. This can result in a theory which is *"very rich in detail but lacks the simplicity of overall perspective"* (Eisenhardt 1989: 547). Building theory from the specifics of empirical evidence may result in narrow theory which describes a very specific phenomenon. Such theories are testable, novel and valid but may be unable to achieve a high level of generality. This may require a series of theory building and theory testing studies.

4.2.9 Ethical considerations

A review of the ethical implications of this research was undertaken in order to identify and address any risks associated with data collection and interviewee participation. The author took care to obtain informed consent to participate in the research process before every interview, in accordance with the University of Leeds' ethical approval procedures. The research project was reviewed and approved by the Ethics Committee at the University of Leeds (ref no. LTEARS-005).

Interviewees were provided with information about the content and purpose of the research and the contribution of their interview to that research. Furthermore, interviewees were given the opportunity to identify statements that they would not like to be used in the research during the interview and were sent a copy of the interview transcript to correct any misinterpretation or identify further quotes to exclude from analysis.

Perhaps one of the most significant ethical considerations in this thesis was the use of data from failed case studies. Eco-Island Community Interest Company (CIC) went into voluntary liquidation stating lack of funding as the cause: "*The existing partnership funding structure was hit badly by the current economic climate and the new Energy Company initiative, in spite of numerous funding streams in the pipeline has not come online in time to provide the CIC with* *the necessary funding it required"* (Findon 2013b). Shortly after this the project was subject to a police investigations relating to mis-use of funds which the CIC has been awarded by the Isle of Wight Council to deliver Green Deal projects (Findon 2013a). The CIC founder was found dead two days later (Island Echo 2013). It was considered that using data from interviews with this participant would be unethical. In the absence of a clear outcome from the fraud investigation, it was not certain that the interview data was representative of the real situation. It was however, decided to use data provided by other participants in Eco-Island interviews, who provided valuable insights into the attitudes of private companies in novel modes of infrastructure operation, which was not affected by the liquidation of Eco-Island.

4.3 Summary

A mixed-methods approach to analysis has been used in this thesis to better represent the different nature of constraints under investigation. This also allows comparison of different approaches in relation to the type of knowledge generated and the relevance it has in identifying responses to constraints.

Critical material constraints occur from human-environment interactions and are considered to be best suited to an empirical-analytical approach using quantitative, objective criteria to evaluate and compare the criticality constraints faced by different technology pathways. The objective criteria is operationalised as an indicator set to represent the multiple dimensions of criticality, including supply disruption potential and the impact of this disruption on the goal of low-carbon energy generation. A composite indicator is developed using an iterative process and individual indicators are quantified separately to improve transparency and identify key drivers of critical material constraints. Analysis is dynamic, but only in relation to changes in the individual indicators over time. The method for aggregation of individual indicators represents the relationship between indicators to some extent but does not fully capture the dynamic nature of these relationships. This would require advanced modelling, which is beyond the scope of this thesis.

Constraints from policy and regulation are investigated using theory building from case study analysis. An analytical framework, combining insights from MLP and co-evolutionary thinking, is used to structure case study analysis and make sense of empirical data. A series of individual cases of alternative modes of infrastructure are analysed which represent different actors, with different motivations and different means of achieving those motivations. These individual cases contribute to the meta-case study of alternative modes of infrastructure operation. Emerging hypothesis of the characteristics of alternative modes of operation and the constraints from the regime are compared between individual cases to identify those common to all cases analysed. These characteristics and constraints are considered to be representative of the meta-case study. The use of case studies can result in a more holistic and meaningful characterisation of real-life event but can also result in theory that is rich in detail but lacking in simplicity.

5 Material criticality constraints

This chapter responds to research question three and investigates the nature and scale of constraints from critical material supply disruption. It describes the results of the research undertaken to develop an indicator set that quantifies critical material constraints. Constraints from critical material are represented as a risk to the roll out of low-carbon technologies and the indicator set is structured to reflect this representation. In the second part of the chapter, the indicator set is applied to a case study of decarbonisation of the electricity generation system in the UK. This case study serves to demonstrate the relative constraints faced by different pathways to decarbonisation.

5.1 Indicators of material criticality constraints

Criticality is conceptualised as the potential for disruption of the supply of critical materials to slow down or halt the roll out of low-carbon technologies in the UK necessary to achieve UK emissions reductions targets. The indicator set described in this chapter aims to represent this conceptualisation. This approach was selected because it is better able to account for the short- and medium-term disruption of supply, which is so important to low-carbon transition. Furthermore, the approach used addresses the multitude of factors that affect supply disruption, beyond availability of critical materials.

Analogously to risk, criticality is defined as the product of the probability of an event and the severity of harm resulting from that event. Two principal indices are created to represent these dimensions of risk:

- Supply Disruption Potential (*P*), which quantifies the likelihood that access to a particular material could be restricted.
- Exposure to disruption (*E*), which quantifies the effect of disruption on the goal in question.

When combined through multiplication, the two indices provide an assessment of the risk that material criticality poses to a low-carbon electricity system transition:

$$C(t) = P(t) * E(t) \qquad \text{(eq. 1)}$$

The relative importance of indices is a source of contention and the weights given to individual indices and metrics to represent this importance can have a significant effect on the overall composite indicator (OECD 2008). There are a number of techniques that could be used to develop weighting factors including statistical means, such as factor analysis, and participatory approaches, which engage stakeholders to assign weights. Whichever technique is used, weights are essentially value judgements based on expert judgement and policy priorities (OECD 2008). In the absence of evidence to the contrary, and in common with most composite indicators, the indices and metrics in this thesis are equally weighted when combined. One limitation of this approach is that it assumes the trade-off between indicators is comparable – that an increase in one metric can be off-set by a corresponding decrease in another. In reality trade-offs will be far more complex than this and more detailed analysis would be required to fully explore this complexity (Knoeri et al. 2013).

Importantly, both indices are produced as a forecasted time-series, which allows the estimation of criticality over time and the identification of trends of increasing (or decreasing) criticality. Each index is composed of a series of metrics, the trends in which can be tracked individually. This is essential to provide more detailed insights into the drivers of criticality for particular materials or technologies and the associated policy interventions that might reduce criticality.

The drivers of each index are numerous and the relationships between them are complex. This presents a significant challenge when attempting to represent criticality, particularly the change in criticality over time. The importance of individual drivers and the relationships between them are likely to change. Detailed analysis of this complexity would require dynamic models which can better represent the evolution of relationships between drivers but which are difficult and time consuming to construct (Knoeri et al. 2013). In this thesis, it was decided that a less complex but more transparent approach would be used where indices are aggregated using the same approach throughout the period of analysis. In this sense the dynamic nature of the analysis is related to temporality, the relationships between metrics remain static.

The combination of metrics contributing to indices is summarized in figure 5.1 and the metrics themselves are described below.

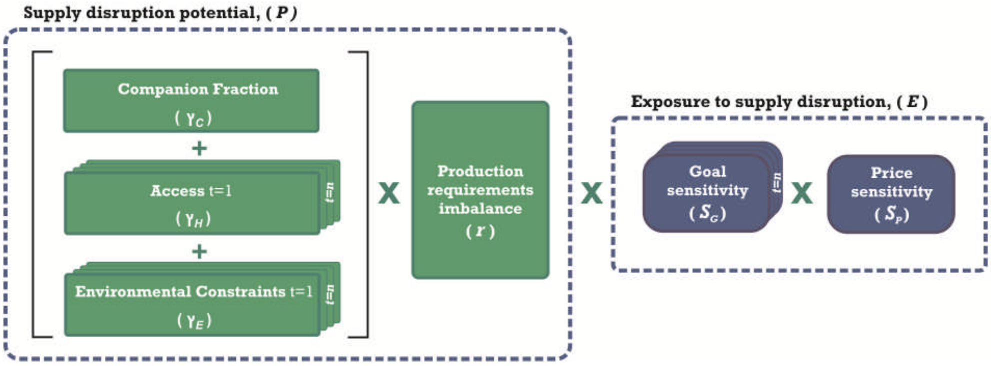


Figure 5.1: Components of index for assessing the material criticality constraints to infrastructure transitions

Using this method it is possible to analyse the relative criticality of a range of transition pathways, such as those outlined by DECC in the Carbon Plan (HM Government 2011a). This allows the user to compare the material constraints of different pathways to the same goal and identify pathways that might leave them most exposed to criticality. This is demonstrated in section 5.2.

5.1.1 Supply Disruption Potential, P

The Supply Disruption Potential index represents the likelihood that access to a particular material could be restricted as a result of an imbalance between production and requirements, which could be exacerbated by a range of factors that could constrain future increases in production. Therefore a metric 'r' is produced which represents the potential scale and frequency of imbalance over the period of analysis and a series of exacerbating factors γ_X. Comprehensive analysis of each exacerbating factor would be complex and require advanced modelling. Therefore, three factors have been selected that are considered to have significant and direct influence on production-requirements imbalance, have widely recognised metrics associated with them, and are readily quantifiably. These three factors are: co-production (γ_C) (many critical materials are not produced as primary products but as co-products of other materials); geographic distribution of production (γ_H) (geographic monopolies in production may tempt policymakers to impose supply restrictions for geopolitical purposes); and environmental constraints (γ_E) (the environmental sensitivity of land surrounding mines may give rise to restrictive legislation).

For a given material, it is assumed that the exacerbating factors tempering the production-requirement imbalance *r* (namely $\gamma_{C,\,H,\,E}$) are independent and equally weighted. The sum of the exacerbating factors is multiplied by the production-requirements imbalance to provide an overall assessment of the potential for supply disruption.

In order to compare criticalities of materials values are normalised with respect to some well-characterized element (e.g. iron), denoted by the subscript *0* using a method developed by Roelich et al (2014b). This allows the expression of *relative* criticality: the magnitude of the increase in criticality (e.g. "moving to the new technology will increase the risk of probability of disruption by a factor of p/p_0"). Normalizing with respect to a reference material, it is possible to write (eq. 2):

$$\frac{p(t)}{p_0} = P(t) = \left(\frac{\sum \gamma(t)}{\sum \gamma_0} \cdot \frac{r}{r_0}\right) \qquad \text{(eq. 2)}$$

The production of individual metrics contributing to P is described below.

5.1.1.1 Production-Requirements imbalance, r

Historically, production of materials has increased in response to market signals driven by greater demand. However, a number of emerging technologies, such as wind turbines, electric vehicle and solar photovoltaics rely on materials identified as potentially critical (European Commission 2010b; US Department of Energy 2011; Morley & Eatherley 2008). These technologies are forecast to be rolled out at an unprecedented scale and rate, increasing demand for materials such as neodymium and indium by up to 700% and 800% respectively (European Commission 2010b). Expanding the production of mines can be slow; the lead times involved can be in the order of years or even decades (Wellmer & Dalheimer 2012). This limits the ability of production to respond to the projected increase in requirements within the timescales required (Morley & Eatherley 2008); normal market signals will not be effective in stimulating new production.

This increase in pressure on production is measured using a ratio of requirements to production over the period under investigation (2012 to 2050). The European Commission (EC) study on critical raw materials uses a version of this ratio with a static level of production and only taking into account requirements from new technologies (European Commission 2010b). This ratio has been adapted by Roelich et al (2014b) to include a forecast of production and requirements and to include requirements from all uses, not just new technologies. A shortfall between potential production and forecasted requirements implies that there is potential for disruption of supply.

The metric of imbalance r, depends on the expected level of global production $M(t)$ and the expected global requirements $R(t)$ at a point in time. $M(t)$ is forecast by projecting historical trends in production. $R(t)$ includes future requirements for low-carbon technologies *as well as* future requirements for other uses, which has often been excluded from previous assessments (Candelise et al. 2012; Andersson 2001a; Speirs, Houari, et al. 2013).

Two metrics are derived to quantify the severity of a supply imbalance between *M(t) and R(t),* which are multiplied to provide at total severity estimate *r*.

1. *n* : Likelihood of supply disruption is the probability that *R(t) > M(t)* over the period considered i.e.;

$$n = \frac{a_{R>M}}{a} \quad \text{(eq. 3)}$$

Where, $a_{R>M}$ is the number of years in which *R > M* and *a* is the total number of years. *n* therefore varies between 0 (requirements exceeds production in none of the years) to 1 (requirements exceeds production in all of the years).

2. σ: Scale of potential supply disruption is the average of (*R – M*)/(*R + M*) over the time period of analysis, counting only years where there is a deficit (i.e. where *R > M*) and thus effectively assuming a worst case scenario where surpluses cannot be carried forward; σ therefore varies between 0 (production exceeds requirements in all years) to 1 (production is insignificant compared to requirements in all years). In notation form: for *a* years;

$$\text{where } M_t > R_t \text{ then } \sigma_t = 0, \quad \text{(eq. 4)}$$

$$\text{and; where } M_t < R_t\text{; then } \sigma_t = \frac{R_t - M_t}{R_t + M_t} \quad \text{(eq. 5)}$$

The severity of disruption is the product of the frequency and scale, therefore:

$$r = n\sigma = \frac{a_{R>M}}{a} * \frac{\sum \sigma_t}{a} \qquad \text{(eq. 6)}$$

The metric r gives a general indication of the potential scale and frequency of market imbalances over the period.

5.1.1.2 Companion fraction (γ_C)

A large proportion of materials currently considered critical are not mined in their own right, but rather as a co-product of a primary material, usually a ‘major’ metal with very high demand across a range of economic sectors, such as copper or zinc (Ayres & Peiró 2013). If a critical metal constitutes only a small proportion (in terms of tonnage and/or price) of the output of a mine, it is unlikely that production would increase solely as a result of a rise in demand for this material, since this would result in a surplus (and thus price suppression) of the primary metal, potentially making the mine less economic overall. This effect is particularly significant where the critical metal price is low and the reduced price for the primary metal cannot be compensated for by the increase in production of the critical metal co-product.

To reflect this constraint a metric γ_C is included which captures the product-co-product balance. In common with other assessments this includes a measure of the proportion of critical material by mass in the output of mines where it is produced. However, this approach is enhanced by incorporating the percentage contribution of the material to the price of one unit of mine output; the economic value of output. Thus:

$$\gamma_c = 1 - \left(\frac{C_m + C_p}{2}\right) \qquad \text{(eq. 7)}$$

Where C_m is the contribution of critical material to the mass of mine output and C_p is the contribution to the economic value of mine output. The contribution from mass and price is weighted equally in the absence of evidence that they are not equally important. The inverse is taken to ensure the scale represents the same as other exacerbating factors i.e. 0 is low and 1 is high. Thus, a value of γ_C approaching 0 would indicate that the material is mined in its own right; a value approaching 1 would indicate that it is almost entirely mined as a co-product of another material.

The proportion of material considered by mass in the output of mines at which it is produced (the mass fraction C_m) is calculated thus:

$$C_m = \frac{m_i}{m_x} \qquad \text{(eq. 8)}$$

Where m_i is the mass of material *i* produced at mine *x* and m_x is the total mass of material produced at mine *x* (data from (US Geological Survey 2013)). In the event of insufficient data on mine outputs, the typical mass fraction of the most commonly produced ore grade is used.

To account for the effect of the price of the material considered to increase the likelihood of accelerating production the percentage contribution of the material to the price of one unit of mine output (the price fraction C_p) is calculated thus:

$$C_p = \frac{p_i * m_i}{p} \qquad \text{(eq. 9)}$$

Where p_i is price per unit of material (using 2010 figures taken from USGS Mineral Commodity Summaries (US Geological Survey 2010)) and p is total monetary output of the mine.

It is possible that the companion fraction of critical materials will change over time; however, there is insufficient data to forecast how this might develop. Therefore it is assumed that the companion fraction stays the same over the period of analysis.

5.1.1.3 Access (γ_H)

Mineral deposits, by virtue of the processes by which they are formed, tend to be concentrated in a specific geographic location, which has implications for access to these materials in countries that do not have their own deposits. This geographic concentration of materials does not directly constrain the acceleration of production; however, the monopoly created by this concentration of production can restrict access to produced materials, further distorting the balance between requirements (outside the country of production) and available production. There is potential for producing countries to pursue industrial and/or geopolitical strategies to reserve resources for their exclusive use though trade restrictions, taxations and investment policies. An example of this is the tightening restrictions that China has placed on rare earth element exports (Danlu 2012; Yang 2012; Bradsher 2012). The geographic concentration of production at present is not necessarily indicative of concentration in the future. For example, China currently produces over 97 per cent of rare earth elements; however, it only holds 36 per cent of reported reserves (US Geological Survey 2010). This would imply that geopolitics could potentially become a less significant factor in the potential for supply disruption and that a dynamic measure of access is essential.

The Herfindahl-Hirschman Index (HHI) was used in previous criticality assessment to quantify the level of concentration of worldwide production (European Commission 2010b; Graedel et al. 2012). In this thesis the HHI is also used but a forecast to 2050 is created to support dynamic analysis. HHI is calculated by squaring the share of production from each country for a given year and summing the result for all producing countries (eq. 10):

$$\gamma_{H(t)} = HHI(t) = \sum_{i=1}^{n} H_{it}^2 \qquad \text{(eq. 10)}$$

Where H_{it} is the share of production of country i in year t and n is the number of producing countries.

The HHI falls onto a 0-1 scale where HHI approaching 1 would represent a concentration of production in a single country approaching a monopoly and HHI approaching 0 would suggest very distributed production and a competitive supply chain. A low level of competition (or a high γ_H) would increase the supply disruption potential as a result of its magnification of the risk associated with the production-requirements imbalance.

The distribution of production of each critical material is forecast to estimate how γ_H might change over time. The forecast is produced by interpolating between the distribution of current production and the distribution of reserves. This assumes that production distribution at the end of the period (2050) is the same as the estimated reserve distribution in 2012. This is a simplification of the real situation but is used to indicate how γ_H might evolve over time based on current reserves and market responses.

5.1.1.4 Environmental Constraints (γ_E)

The production of metals can have significant environmental impacts as a result of pollutant discharge to air, land and water and waste production (Moriguchi 2010) and consumption of energy and water, which will increase as ore grades deteriorate (Norgate 2010). In an attempt to contain these impacts, and as a result of international treaties, environmental regulation is becoming increasingly stringent. This is presenting a barrier to expansion of existing mining operations or the development of mines by increasing the cost of production (European Commission 2010b).

Quantifying the extent of environmental constraints on mining or the direct effect of regulation on new mining operations is rather difficult. As an alternative, previous assessments have used the Environmental Performance Index (EPI), as a measure of “*the risk that measures might be taken by countries with the intention of protecting the environment and by doing so endangering the supply of raw materials…”* (European Commission 2010b). The EPI represents a very broad and general assessment of countries environmental regulations based on the outcomes they achieve, rather than an assessment of the existence or enforcement of regulation itself. The EPI track outcome-oriented indicators, which have two overall objectives; to reduce environmental stresses on human health and promote ecosystem vitality. The performance against these objectives is tracked across ten policy categories which are assessed on 22 performance indicators. This structure is shown in figure 5.2, taken from the EPI 2012 report (Emerson et al. 2012).

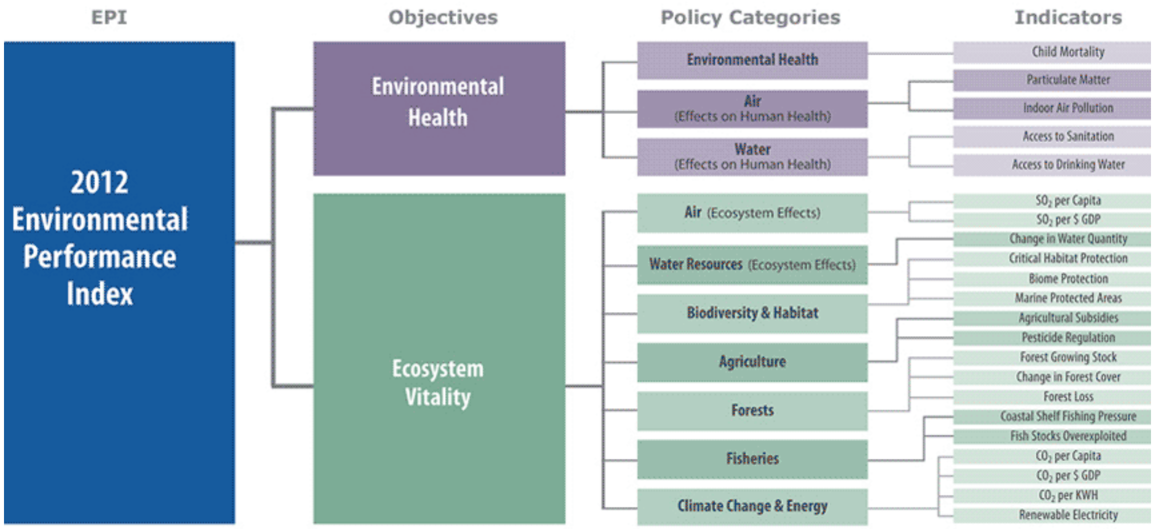


Figure 5.2: An illustration of how the EPI is compiled from 22 indicators grouped into 10 policy categories that fall into two environmental objectives (Source: Emerson et al. 2012)

Although several of the indicators used for the EPI are directly relevant to mining operations, e.g. Change in Water Quantity and Forest Loss, EPI is not included as a direct indicator for environmental regulations that impact mining operations. Rather, it provides a country level comparison of the potential for a country to already have, or institute new environmental regulations that may constrain mining operations (Emerson et al. 2012). The majority of critical metals are mined in more than one country, therefore, it is necessary to combine the EPI of individual countries to determine what the European Commission terms the Environmental Country Risk (ECR) (European Commission 2010b), here termed γ_E.

To calculate γ_E for a given year, the EPI is divided for each producing country by 100 to convert to a 0-1 scale, this is then weighted by the share of production from each country and the resulting figures are summed:

$$\gamma_{E(t)} = ECR(t) = \sum \left(\frac{EPI_i}{100}\right) * \left(\frac{M_{it}}{M_x}\right) \qquad \text{(eq. 11)}$$

Where EPI_i is the environmental performance index of country *i*, M_{it} is the production of a specific material in country *i* in year *t* and M_x is total production of that material. A value of γ_E approaching 0 indicates that there is unlikely to be any constraints to the development of new mining operations from environmental regulation.

It is likely that the environmental performance of countries (and hence the EPI) will improve over time, however there is insufficient historical data on which to base any forecasts about the rate of this improvement and this could be balanced (or even outweighed) by the fact that environmental legislation is likely to become more restrictive. Therefore, the EPI for each country is held static over the period of analysis. However, the proportions of production in different countries (and therefore the contribution of each EPI to γ_E) are likely to change. The forecast split of global production between countries is obtained from the HHI forecasts described above.

5.1.2 Exposure to Supply Disruption, E.

This index represents the effect of supply disruption on the transition to a low-carbon infrastructure system. In an overall electricity system in transition towards a low-carbon paradigm, any disruption to the roll-out of the technology on which this transition depends has obvious implications for energy security. Unlike the supply disruption potential, which is a material property, exposure is a property of the technical system and, therefore, must be assessed at the level of the goal that is being analysed i.e. decarbonisation of the infrastructure system. Exposure is operationalized as the product of the proportion of the goal affected by any disruption (the goal sensitivity S_G), and the likely effect of increasing price resulting from disruption (the price sensitivity S_P);

$$E(t) = S_{G(t)} * S_P \qquad \text{(eq. 12)}$$

The two metrics contributing to the exposure indicator are described in detail below.

5.1.2.1 Goal sensitivity S_G

The overall goal of transition to a low-carbon infrastructure system is operationalized as pathways or scenarios of technology roll out required to achieve decarbonisation[16]. The goal sensitivity, or the impact of a supply disruption on the overall goal, is measured in this index as the proportion of the decarbonisation pathway that relies on the technology or technologies affected by the potential material supply disruption. This assessment is based on the contribution of the technology of interest to the low-carbon goal (MW of electricity generation capacity, or tonnes of CO_2 reduction) as defined in the relevant pre-defined pathway (in this analysis this is taken to be DECC's low-carbon pathways (HM Government 2011a));

$$S_G = \frac{G_{tech}}{G_{goal}} \qquad \text{(eq. 13)}$$

---

[16] Within the country or region of interest – note that this may differ from the global roll out of low-carbon technologies used to determine future global requirements within the production-requirements imbalance metric is the scale of decarbonisation is sub-global.

Where: S_G is goal sensitivity; G_{tech} is the amount contributed to the low-carbon goal (miles, MW, CO_2 reduction etc.) by the technology (defined in relevant scenario); and G_{goal} is the total amount required to achieve goal (defined in the relevant pathway).

A high value of S_G (i.e. approaching unity) would imply that constraining the roll out of the technology of interest could completely derail the goal of low-carbon transition. A low value of S_G (i.e. approaching zero) would mean that the goal was relatively insensitive to the roll out of the technology of interest.

S_G is produced as a time series because the contribution of a particular technology in any year varies according to the pathway.

5.1.2.2 Price sensitivity S_P

As well as having the potential to physically constrain technology roll out; supply disruption could cause an increase in price, which could have further implications. To capture this effect, the price sensitivity metric quantifies the proportion of the total technology cost contributed by the cost of the material at risk of supply disruption:

$$S_P = \frac{V_{mat}}{V_{tech}} \qquad \text{(eq. 14)}$$

Where: S_P is price sensitivity; V_{mat} is the price of material in technology of interest (i.e. price per tonne multiplied by mass of material in technology); and V_{tech} is the price of technology of interest.

A high value of S_P (i.e. approaching unity) would imply that the technology cost was very sensitive to fluctuations in material price. A low value of S_P (i.e. approaching zero) would imply that the technology cost was relatively insensitive to price fluctuations and material supply disruption was less likely to constrain the required technology roll out as a result of increasing prices. It is recognized that the scale and design of technologies will change over time, affecting the price sensitivity. However, this change cannot be quantified to any degree of certainty at this time; therefore, this metric is assumed to be static for now.

5.2 Case study: Material criticality constraints to low-carbon electricity in the UK

The criticality assessment method is demonstrated using a case study of neodymium criticality constraints to UK low-carbon electricity, with a focus on wind turbines. Rare earth elements, predominantly neodymium, are used in permanent magnets required for gearless, direct drive wind turbines. Wind power has the potential to contribute significantly to the decarbonisation of UK electricity generation and is central to many of DECC's 2050 Pathways (HM Government 2011a). An ambitious roll-out programme of wind turbines is required to replace fossil-fuel powered generation. Neodymium is already identified by many recent reports as being at risk of supply disruption as a result of the concentration of its production in China (European Commission 2010b; Moss et al. 2011; US Department of Energy 2011; Du & Graedel 2011b).

This case study is used to determine how this potential supply disruption might affect the deployment of low-carbon electricity generation in the UK. It is recognized that this is only a first approximation as it is necessary to take into account the fact that almost all significant technologies are exposed to criticality via multiple materials and that individual critical elements are essential to the operation of multiple technologies.

The results are presented for the composite indicator first and then each contributing metric, along with the data used to calculate it, is discussed in turn.

5.2.1 Criticality of low-carbon electricity

The criticality of two of DECC's 2050 Pathways; Core Pathway and its Renewable Pathway (HM Government 2011a); has been calculated for the period from 2012 to 2050 using the indicator set described in section 5.1. The results for the composite criticality index (presented in figure 5.3) show that criticality in the Core Pathway increases four-fold over the period from 2012 to 2050, with a step-change occurring in 2030, as shown with reference to 2012 values. This trend is even more dramatic in the Renewables Pathway with an almost ten-fold increase.

The results of the composite index show a significant increase but provide little insight into how policy makers or commercial organizations might intervene to reduce this criticality. This requires analysis of the metrics contributing to the composite indicator, which is discussed in the following sections.

Figure 5.3: Criticality of two scenarios of transition to low-carbon electricity generation in the UK 2012-2050

5.2.2 Neodymium Supply Disruption Potential

Despite the significant increase in criticality, the Supply Disruption Potential (P) of neodymium (normalized to iron) decreases over the period of analysis from 11.65 to 8.47 a 27% reduction. This indicates that the overall trend in criticality of UK electricity system decarbonisation is driven strongly by the UK's increasing exposure to supply disruption and that intervention would be most effective if focused on reducing this exposure. This is not to say that intervention should not aim to accelerate the reduction in supply disruption potential, therefore the metrics which constitute this index are discussed in more detail below.

5.2.2.1 Neodymium production-requirements imbalance (r)

Future requirements forecast (R(t)) for neodymium

R(t) is comprised of estimates of future global requirements for neodymium from low-carbon technologies and from other sectors. The global growth in requirements from low-carbon technologies is projected to outstrip growth in demand from more traditional sectors (European Commission 2010b; Moss et al. 2011) so has been treated separately.

The future requirements from global low-carbon technology deployment is estimated using forecasts of low carbon technology roll-out and estimates of neodymium requirements per unit of technology. The global roll out of low-carbon technologies is derived from the International Energy Agency Technology Roadmap scenario to 2050 (IEA 2012), which provides capacity at 10 year intervals. A smooth roll out was assumed and the increase in capacity over each 10 year period was distributed evenly between intermediate years to provide the additional installed capacity (in MW) each year (data provided in Appendix 2). Note that this is a conservative assumption since it does not include additional capacity required to replace turbines which come to the end of their lifetimes. Not all wind turbines contain permanent magnets, which is the component that relies on neodymium, so the additional capacity of technologies must be reduced to reflect only the capacity that contains permanent magnets, and thus neodymium. It was assumed that the proportion of turbines containing permanent magnets changed each year as turbine technology became more mature. Kara et al (2010) assume that the proportion of turbines containing magnets increase up to 2040 (see table 5.1). After this time the increasing trend was projected to give the proportion up to 2050 (because Kara et al only forecast to 2040).

Table 5.1: Proportion of wind turbines containing permanent magnets (and thus neodymium)

Year	PMG share	Year	PMG share	Year	PMG share	Year	PMG share
2012	12%	2021	21%	2031	27%	2041	41%
2013	13%	2022	21%	2032	28%	2042	42%
2014	14%	2023	22%	2033	30%	2043	43%
2015	15%	2024	22%	2034	31%	2044	44%
2016	16%	2025	23%	2035	33%	2045	45%
2017	17%	2026	23%	2036	34%	2046	46%
2018	18%	2027	24%	2037	36%	2047	47%
2019	19%	2028	24%	2038	37%	2048	48%
2020	20%	2029	25%	2039	39%	2049	49%
		2030	25%	2040	40%	2050	50%

This data on the additional MW of technology requiring neodymium installed each year is combined with coefficients of neodymium requirements per MW of technology (table 5.2) to estimate the global neodymium requirements from global deployment of low-carbon technologies.

Table 5.2: **Material intensity coefficients**

Technology	Nd Content	Unit	Source
EV/PHEV	1.5	kg/vehicle	(Kara et al. 2010)
ICE Vehicle	0.06	kg/vehicle	(Kara et al. 2010)
PMG wind turbine	166	kg/MW	Average of (Kara et al. 2010; US Department of Energy 2011)

These data were combined to calculate total neodymium requirements for each low-carbon technology;

$$R_t = \sum T_{it} * I_i \quad \text{(eq. 15)}$$

Where R_t is neodymium requirements at year t, T_{it} is installed capacity of technology i at year t and I_i is the material intensity of each unit of technology i.

Requirements from other uses in the base year were calculated by subtracting low-carbon uses (calculated for that year using the method described above) from total production in that year[17]. Thereafter, the ‘other uses’ total was increased by 3% annually, which is a simplifying assumption that demand for most materials is likely to mirror global economic growth and that annual global GDP growth will remain around 3% (PWC Economics 2013). The results of these calculations are provided in Appendix 2.

Future production forecast

The future production of neodymium (M(t)) was estimated by projecting the historical trend in growth of production from the base year to 2050. This trend was determined using data for rare earth oxide production[18] for the period from 1990-2011 (US Geological Survey 2013) and is shown in figure 5.4. The historical trend was linear and could best be explained using equation 16:

$$y = 4413x + 42703 \text{ (eq. 16)}$$

This growth function was used to forecast future production of REOs. The production of neodymium was assumed to be 18% of REO production (Kara et al. 2010). Data resulting from this analysis is presented in Appendix 2.

---

[17] In the base year it is assumed that the amount produced is equivalent to the total amount required and that there is no surplus.

[18] Production data is not provided for neodymium, which is one product of rare earth oxides, so the trend was inferred from mining of the parent material.

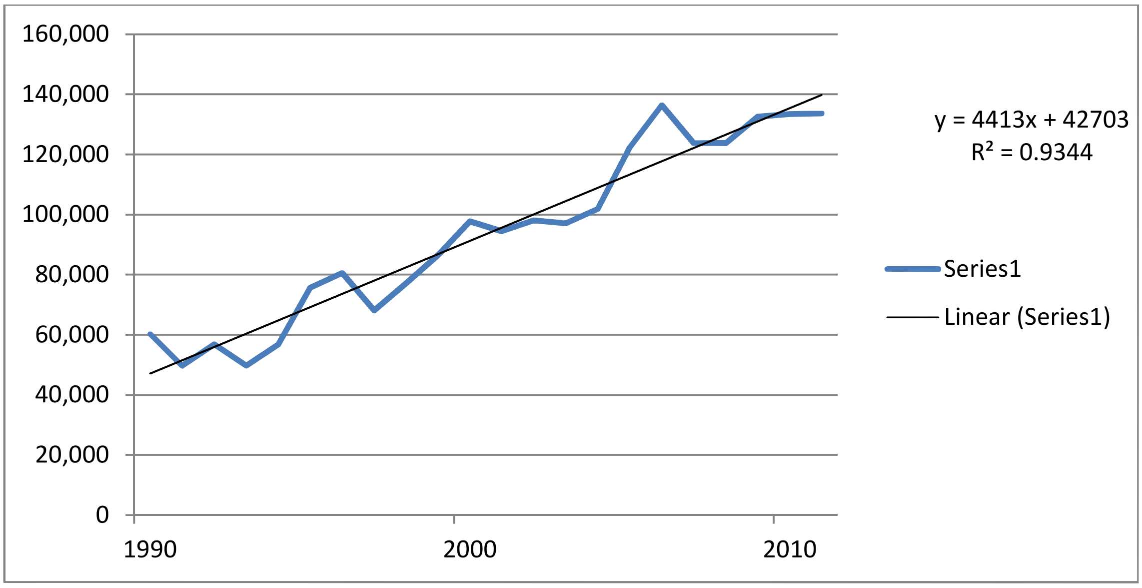


Figure 5.4: Trend in rare earth oxide (REO) production 1990-2011

Calculating frequency and scale of deficit

The requirements and production data are used to calculate the frequency and scale of deficit, using equations 3, 4, 5 and 6, to generate r for neodymium. The results of these calculations are presented in table 5.3.

Table 5.3: Analysis of frequency and scale of deficit

Frequency of deficit				Scale of deficit		
Year	Requirements (R(t))	Production (M(t))	M:R	R(t)-M(t)	R(t)+M(t)	σ
2012	25985	24,839	0.96	1147	50824	0.02
2013	27864	25,633	0.92	2231	53497	0.04
2014	29851	26,427	0.89	3424	56279	0.06
2015	31947	27,222	0.85	4725	59169	0.08
2016	34152	28,016	0.82	6136	62168	0.10
2017	36467	28,810	0.79	7656	65277	0.12
2018	38891	29,605	0.76	9287	68496	0.14
2019	41427	30,399	0.73	11028	71826	0.15
2020	44074	31,193	0.71	12880	75267	0.17
2021	48014	31,988	0.67	16027	80002	0.20
2022	52215	32,782	0.63	19432	84997	0.23
2023	56246	33,576	0.60	22669	89822	0.25
2024	60108	34,371	0.57	25738	94479	0.27
2025	63803	35,165	0.55	28638	98969	0.29
2026	67331	35,960	0.53	31372	103291	0.30
2027	70692	36,754	0.52	33939	107446	0.32
2028	73888	37,548	0.51	36340	111436	0.33
2029	76918	38,343	0.50	38576	115261	0.33
2030	79784	39,137	0.49	40647	118921	0.34
2031	86568	39,931	0.46	46637	126499	0.37
2032	92805	40,726	0.44	52080	133531	0.39

	Frequency of deficit			Scale of deficit		
Year	Requirements (R(t))	Production (M(t))	M:R	R(t)-M(t)	R(t)+M(t)	ơ
2033	99307	41,520	0.42	57787	140827	0.41
2034	106074	42,314	0.40	63760	148388	0.43
2035	113107	43,109	0.38	69998	156215	0.45
2036	120406	43,903	0.36	76503	164309	0.47
2037	127973	44,697	0.35	83275	172670	0.48
2038	135808	45,492	0.33	90316	181299	0.50
2039	143912	46,286	0.32	97626	190198	0.51
2040	152287	47,080	0.31	105207	199367	0.53
2041	156959	47,875	0.31	109084	204833	0.53
2042	162427	48,669	0.30	113758	211096	0.54
2043	168004	49,463	0.29	118541	217468	0.55
2044	173690	50,258	0.29	123433	223948	0.55
2045	179486	51,052	0.28	128434	230538	0.56
2046	185394	51,846	0.28	133548	237240	0.56
2047	191414	52,641	0.28	138773	244055	0.57
2048	197548	53,435	0.27	144113	250983	0.57
2049	203797	54,229	0.27	149567	258026	0.58
2050	210162	55,024	0.26	155138	265186	0.59
	Frequency of deficit			Scale of deficit		
	occurrence		39	Average **ơ**		0.36
	Percentage (n)		1.00			

The results of this analysis show that there is an acceleration in projected requirements in 2030 as a result of global demand for low-carbon technologies. Even exponential growth in production cannot keep up with expansion in the number of technologies using neodymium and the scale of roll out of these technologies (see figure 5.5). Consequently both the frequency (i.e. 1 indicating that requirements are higher than production for all of the period) and scale of potential disruption (i.e. 0.36 indicating that requirements are higher than production) are high. This results in a production-requirements imbalance (r) over the period 2012-2050 of 0.36 compared to 0.09 for iron, indicating that there is a high potential for supply disruption.

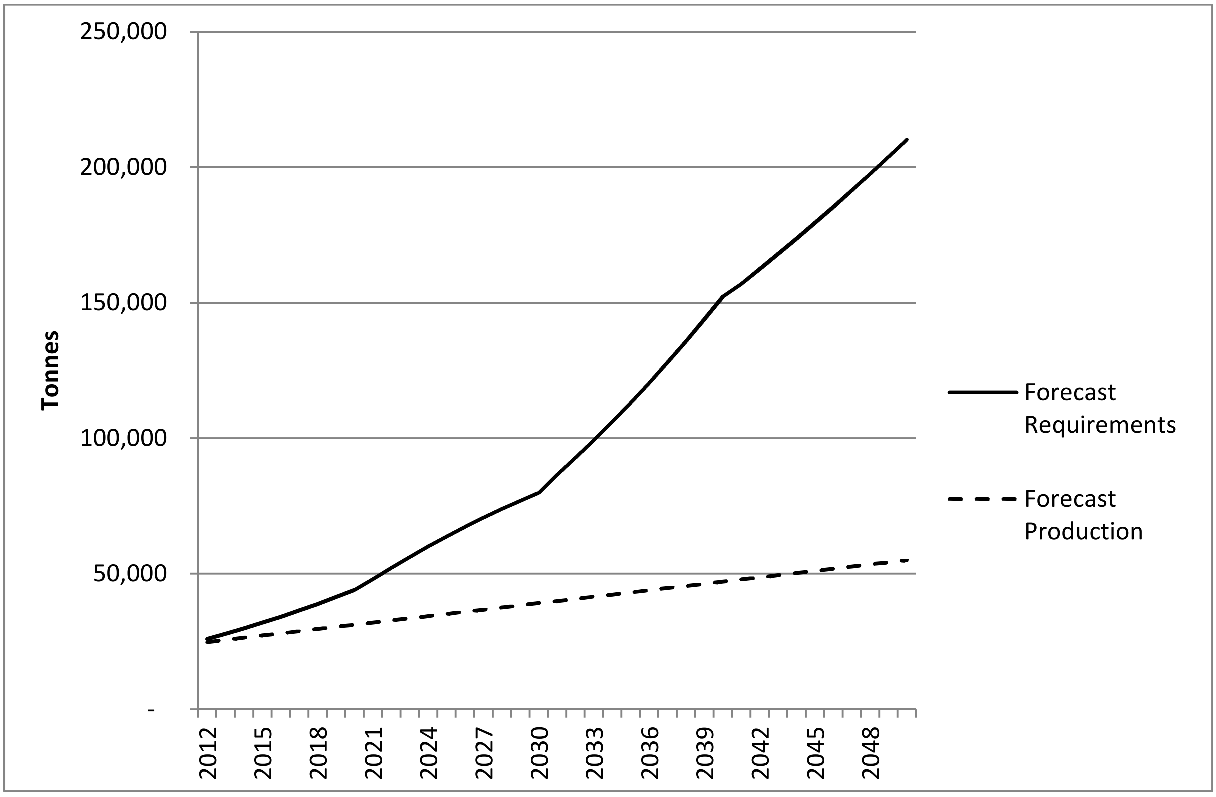


Figure 5.5: Forecasts of neodymium production (M) and requirements (R) 2012-2050

5.2.2.2 Exacerbating factors (γ)

This potential for neodymium supply disruption is exacerbated by a series of factors, which are discussed in turn below.

Environmental constraints (γ_E)

The risk that environmental legislation could constrain the development of new reserves of neodymium, is relatively low in 2012 (0.42) as a result of the dominance of China (which has a low EPI) in its production. However, as the distribution of production is increased, the contribution of countries with a higher level of environmental legislation, such as Australia and the USA, increases and γ_E increases slightly by 2050 to reflect this increase in constraint from regulation (to 0.47) (figure 5.6).

Access (γ_H)

The current production of neodymium is almost a monopoly, with the majority produced in China; therefore γ_H is almost unity in 2012 (0.92). However, neodymium reserves are less geographically concentrated than current production would suggest (US Geological Survey 2010) so the dynamic analysis of γ_H for is important and the change in production distribution must be calculated.

To reflect the potential for reducing geographic concentration of neodymium production, the distribution of neodymium production in 2050 is taken to be the same as the reserve distribution provided by US Geological Survey in Mineral Commodity Summaries (US Geological Survey 2010). Figures for the relative production in each country are estimated by dividing the total production in 2010 by the percent of reserves held by the country of interest, thus:

$$M_{i2050} = M_{2010} * D_{i2050} \qquad \text{(eq. 17)}$$

Where M_{i2050} is the production in country *i* in 2050, M_{2010} is total production in 2010 and D_{i2050} is the proportion of reserves held by country *i*. It is recognised that production will, in reality, grow but it is the distribution that is of interest – the total is not relevant in this context. Production for each country in intervening years is interpolated between M_{i2010} and M_{i2050}.

Reserves of rare earth elements are distributed between seven named countries and an 'other' country category. This final category accounts for almost 20% of reserves and in most estimates of HHI it acts as a single country. In reality it will be many different countries with smaller reserves, which could increase competition in the supply chain. This is reflected by dividing the 'other' category into the 15 countries identified in the British Geological Survey's (BGS) Rare Earth Elements summary (BGS 2011).

Neodymium is assumed to be 18% of both current production and reserves (Kara et al. 2010). Data relating to forecasting production distribution is provided in Appendix 2.

When production distribution is forecast towards reserve distribution γ_H reduces from 0.92 to 0.28 by 2050 (figure 5.6). This reflects the likely future evolution of a far more competitive supply chain, which could mitigate the high disruption potential, such as the reopening of historic mines outside of China (Bradsher 2010).

Companion fraction (γ_C)

Neodymium is mined as a co-product of other rare earth metals and represents only 18% of rare earth mine output (C_m=0.18). It also has a relatively low contribution to the economic value of mine output (16%; C_P=0.16)) so could be expected to have limited influence over total mine production when compared to the remaining rare earth elements. This results in a high score for γ_C (0.85), which indicates that co-mining has a high potential to exacerbate the production-requirements imbalance (figure 5.6).

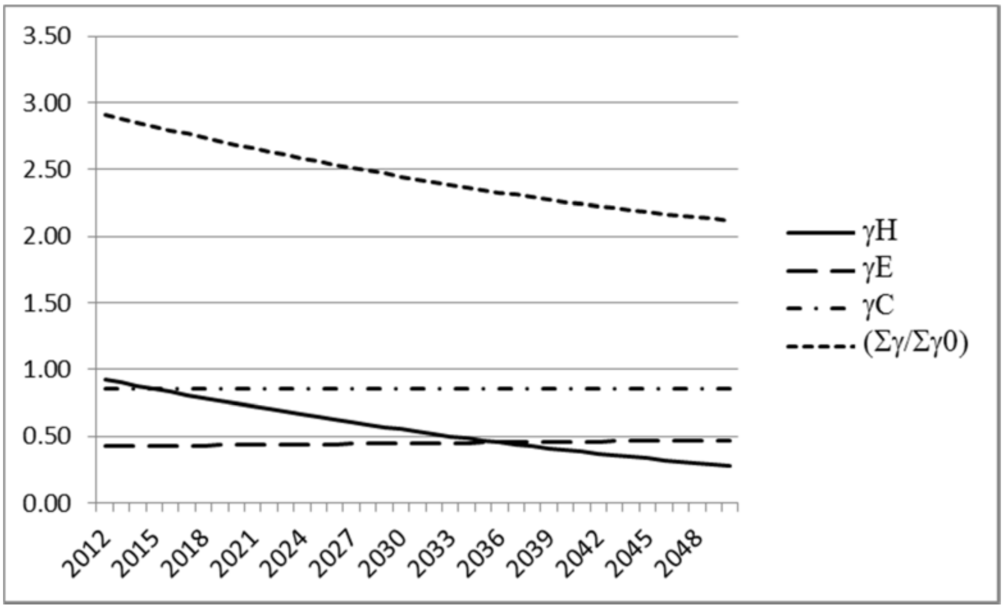


Figure 5.6: Time series of factors exacerbating production-requirements balance – individual factors and sum of factors (sum normalised to iron) 2012-2050

5.2.3 Exposure of wind power to neodymium supply disruption

The evolution of exposure of low-carbon electricity generation in the UK to neodymium supply disruption from 2012 to 2050, is shown in figure 5.7. This uses DECC's scenarios, which reflect wind turbine roll-out in the UK only, rather than the global scenarios used to calculate global demand for neodymium. The results are indexed to a 2012 baseline because the numerical value of E is very small and the scale of increase is hard to see the relative trend.

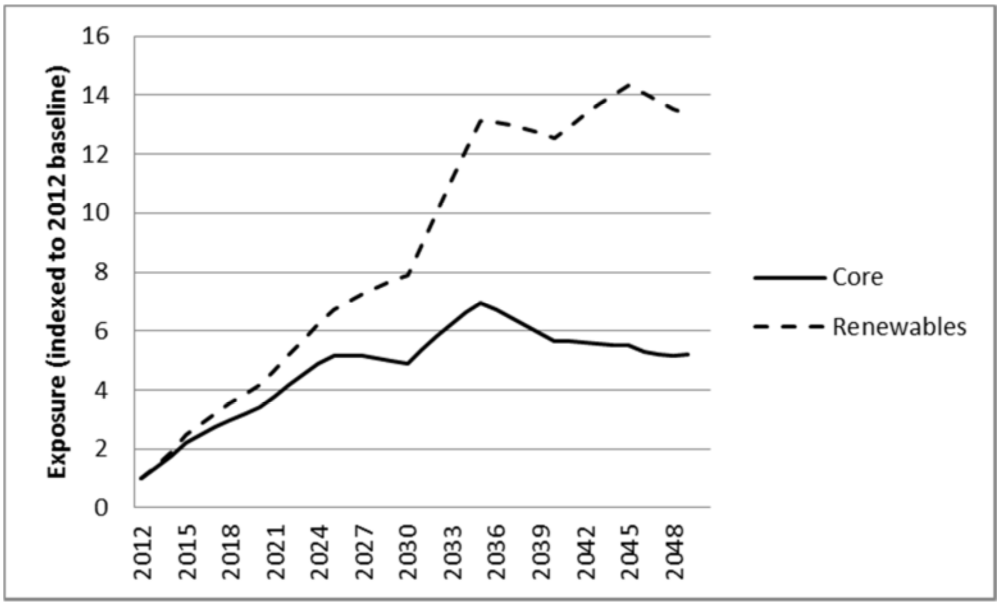


Figure 5.7: Exposure of two scenarios of transition to low-carbon electricity generation to neodymium supply disruption 2012-2050

The trend is the opposite to that of the supply disruption potential, with exposure increasing significantly over the period under investigation, as wind turbines that rely on neodymium become more prevalent in the energy mix. The peaks and troughs in the exposure trend are artefacts of the uneven roll out of wind turbines in the DECC pathways. This exposure is a product of the goal sensitivity and the price sensitivity:

Goal sensitivity (S_G)

The 'goal' of low-carbon electricity generation, which is exposed to supply disruption, is taken to be the capacity of low-carbon electricity generation required each year. Goal sensitivity was determined for the period from 2012 to 2050 using scenarios of technology roll out required to meet the UK's goal of decarbonising electricity generation. The scenarios were predetermined by DECC in their Low Carbon Transition Plan (HM Government 2011a). The roll-out of all wind turbines is provided in GW at five-year increments for both the 'Core' and 'Renewables' scenarios and intermediate years were interpolated to provide annual GW. It was assumed that the proportion of wind turbines containing permanent magnets (and therefore, neodymium) increased at the same rate as was used to calculate global requirements from low-carbon technologies (table 5.2). The total GW relying on neodymium were calculated using interpolated data from DECC scenarios and the proportion of wind turbines each year containing PMGs; these data are provided in table 5.4. The total GW of low-carbon electricity was provided in each DECC scenario, again at five year increments and was interpolated to provide annual GW of total low-carbon electricity.

Table 5.4: Proportion of technology roll out exposed to neodymium supply disruption

	Core			Renewables		
	GW PMG	GW total	S_G	GW PMG	GW total	S_G
2012	1.12	71.69	0.02	1.43	73.34	0.02
2013	1.49	70.43	0.02	2.02	72.93	0.03
2014	1.92	69.44	0.03	2.72	72.79	0.04
2015	2.40	68.85	0.03	3.54	72.98	0.05
2016	2.67	68.76	0.04	4.08	73.56	0.06
2017	2.95	69.12	0.04	4.65	74.59	0.06
2018	3.24	69.87	0.05	5.26	76.14	0.07
2019	3.53	70.93	0.05	5.90	78.25	0.08
2020	3.84	72.23	0.05	6.60	80.98	0.08
2021	4.36	73.70	0.06	7.71	84.35	0.09
2022	4.91	75.28	0.07	8.94	88.15	0.10
2023	5.46	76.92	0.07	10.27	92.15	0.11
2024	5.99	78.55	0.08	11.68	96.09	0.12
2025	6.47	80.12	0.08	13.15	99.74	0.13
2026	6.62	81.58	0.08	14.09	102.90	0.14

	Core			Renewables		
	GW PMG	GW total	S_G	GW PMG	GW total	S_G
2027	6.68	82.91	0.08	14.98	105.60	0.14
2028	6.69	84.08	0.08	15.81	107.91	0.15
2029	6.65	85.08	0.08	16.55	109.92	0.15
2030	6.57	85.89	0.08	17.20	111.71	0.15
2031	7.27	86.51	0.08	19.88	113.36	0.18
2032	7.93	87.05	0.09	22.59	114.97	0.20
2033	8.56	87.63	0.10	25.33	116.66	0.22
2034	9.16	88.35	0.10	28.10	118.53	0.24
2035	9.72	89.34	0.11	30.92	120.69	0.26
2036	9.54	90.68	0.11	31.45	123.22	0.26
2037	9.35	92.29	0.10	31.96	126.02	0.25
2038	9.14	94.07	0.10	32.41	128.98	0.25
2039	8.90	95.89	0.09	32.78	131.97	0.25
2040	8.64	97.65	0.09	33.04	134.86	0.24
2041	8.76	99.25	0.09	34.84	137.56	0.25
2042	8.83	100.71	0.09	36.54	140.10	0.26
2043	8.88	102.05	0.09	38.15	142.52	0.27
2044	8.92	103.31	0.09	39.69	144.88	0.27
2045	9.00	104.52	0.09	41.20	147.23	0.28
2046	8.79	105.72	0.08	41.06	149.63	0.27
2047	8.70	106.99	0.08	40.98	152.20	0.27
2048	8.76	108.43	0.08	41.01	155.03	0.26
2049	9.00	110.12	0.08	41.20	158.25	0.26

The goal sensitivity S_G increases dramatically over the period under investigation from 0.02 to 0.08 for the Core Scenario and from 0.02 to 0.26 for the Renewables Scenario (i.e. up to 26% of low-carbon electricity generation is provided by wind turbines with permanent magnets).

Price sensitivity (S_p)

The price sensitivity of the goal S_P is static and was estimated using a 3MW direct-drive wind turbine, which is representative of the type of turbine that would be deployed in the initial period of the study (Polinder et al. 2005). Data used to calculate price sensitivity is provided in table 5.5.

Table 5.5: wind turbine price data (in 2005 values) used to calculate price sensitivity

	Value	Source
Value of critical material in technology (V_{mat})	£2,827	68 kg/turbine (Vestas n.d.); £41.48/kg (US Geological Survey 2010) (converted from USD using 0.66£/USD))
Total value of technology (V_{tech})	£1,569,167	(Polinder et al. 2005) (converted from Euros using 0.83 £/Euro)
Price sensitivity (S_p)	1.8×10^{-3}	

5.3 Correlation analysis

A correlation analysis was undertaken to examine the independence of indicators and associated assumptions made when combining metrics. Matlab was used to create a correlation coefficient matrix which quantified the strength of a linear relationship between each indicator. A value approaching zero indicates that there is no linear relationship between indicators. A value approaching 1 indicates there is a strong linear relationship.

5.3.1 Indicator Correlation Analysis

Results of correlation analysis for the high level indices that constitute C are presented in table 5.7 below. Results show that there is a higher degree of correlation between criticality and exposure (E) than between criticality and Supply Disruption Potential (P), supporting the conclusion that exposure dominates the trend in criticality.

Table 5.6: Correlation between indicators (R^2 values)

	Year	C	P	E
Year	1	0.42	0.99	0.62
C	0.42	1	0.53	0.96
P	0.99	0.53	1	0.72
E	0.62	0.96	0.72	1

5.3.2 Metric Correlation Analysis

Results of correlation analysis for metrics, which contribute to indices are presented in table 5.8 below.

Table 5.7: Correlation between metrics (R^2 values)

	r	γ_H	γ_E	γ_C	S_P	S_G
r	1	0	0	1	1	0
γ_H	0	1	0.94	0	0	0.71
γ_E	0	0.94	1	0	0	0.58
γ_C	1	0	0	1	1	0
S_P	1	0	0	1	1	0
S_G	0	0.71	0.58	0	0	1

A number of the metrics are shown as having no correlation or total correlation because those indicators are static (r, γ_C and S_P). There is a very high degree of correlation between γ_H and γ_E because γ_E uses HHI (γ_H), to weight the EPI to produce ECR. When γ_E is forecast, the EPI remains static so the trend is entirely driven by HHI, which explain the high degree of correlation. This means that the metrics are not entirely independent, which is implied by the additive combination of indicators. However, metrics used are readily quantifiable, without the need for expert judgment, and clearly articulate the conceptualization of criticality proposed in this preliminary study, providing a basis for more sophisticated conceptions.

5.4 Discussion of criticality results and indicator limitations

The results of the case study demonstrate the importance of considering the temporally dynamic analysis of the risk of material criticality. In the case of low-carbon electricity from wind turbines in the UK, the likely decrease in *P* for the key critical material is outweighed by the increase in the exposure *E* of the goal to that material as the electricity system becomes increasingly reliant on wind turbines; thus the overall criticality C increases over the analysis period. The dynamic approach described in this thesis allows analysis of the nature of the change in criticality over time. The results showed a steep increase in criticality after 2030, when roll out of direct drive turbines is projected to increase dramatically and existing turbines come to the end of their life. This means that wind turbines need to be deployed to not only increase capacity but also to replace existing capacity that is coming out of service. It will be more difficult to devise policy responses to such steep changes than to static high levels of criticality.

The systemic nature and goal orientation of this approach allows the analysis of the relative risk to different pathways to achieve the goal of low-carbon transition (demonstrated using DECC's Core and Renewables pathways). This provides more specific and relevant information to support decision making under uncertainty and may prevent reliance on pathways and technologies that could become highly critical in the future (creating 'lock-in'). The UK aims to replace a significant proportion of current generating capacity with wind turbines, in pursuit of low-carbon goals. Therefore; this analysis could also improve decision making in support of energy security to avoid supply shortfall as a result of restricted roll-out of technology on the scale planned.

A further advantage of the method is its inclusion of demand for critical material from sectors and countries outside the scope of the analysis; to represent the competition with these sectors and countries for material access. The majority of studies of material criticality and low-carbon technologies consider only demand for material from these technologies (Angerer et al. 2009; Speirs et al. 2011; Andersson 2001a; Moss et al. 2011), (with the exception of Alonso et al (2012)), which excludes a significant source of demand. Demand is included from these sources in forecasts of global requirements, which affects the production-requirements imbalance and; therefore, the Supply Disruption Potential.

The case study used to demonstrate the methodology was chosen for its relative simplicity but it is possible to extend this application to analysis of multiple materials and technologies (Dawson et al. 2014) or to other societal goals. A further extension could be to undertake additional analysis to explore the effect of substitution (of materials or technologies) or recycling on criticality through exogenous scenarios. This would support analysis of the effectiveness of different interventions to help shape policy approaches to reduce the criticality of particular goals.

It is recognized that the use of composite indicators has limitations, particularly in relation to indicator weighting, aggregation and combination (European Commission 2010b). A correlation analysis has been undertaken to examine the independence of indicators and the associated assumptions made when combining metrics. This identified a high degree of correlation between some of the metrics used to construct the Supply Disruption Potential index. This means that the metrics are not entirely independent, which is implied by the additive combination of indicators. Equal weighting of these correlated metrics could result in an element of double counting. Further work is required to explore the implications of this double counting.

The equal weighting of metrics and indicators implies that the trade-off between indicators is comparable – that an increase in one metric can be off-set by a corresponding decrease in another. Criticality is a contentious and politically sensitive issue, therefore, using participatory approaches, which engage stakeholders might be of particular benefit (Munda 2005).

Some metrics which might contribute to the understanding of criticality have been excluded because robust data is not widely available or they are not widely supported as reliable indicators by criticality investigators. For example, some studies have included measures of the stability of government and the potential for instability to constrain production. The Worldwide Governance Indicator (WGI) of ‘political stability and absence of violence’ has been used in some studies as a proxy for this (Graedel et al. 2012; European Commission 2010b). However, the WGIs are highly criticized (Thomas 2010; Langbein & Knack 2010) and there is no reported correlation between political instability and production; therefore this measure is not included in this analysis.

The assessment of supply disruption potential only considers disruption at the point of production. It is also recognized that the supply chain of critical materials has additional down-stream stages where disruption may occur. This has been explored qualitatively (IEA-RETD

2012) but a quantitative analysis will require significant further data gathering, investigation and analysis.

A number of the indicators have been forecast in order to demonstrate the benefits of a more dynamic analysis of criticality. It has been necessary to make a series of assumptions to support this forecasting that are of course contestable. The method used to forecast production and requirements rely on the continuation of historic trends to forecast future change. Future production and requirements will in fact be driven and constrained by a multitude of interrelated physical and economic factors, many of which cannot be predicted (e.g. disruptive new technologies, changes in patterns of consumption) (Candelise et al. 2012; Speirs, Houari, et al. 2013).

The dynamic nature of criticality analysis is limited to temporal dynamics; the change in importance of or relationships between metrics is not analysed. This was a pragmatic decision and the current method is considered to be appropriate for its intended purpose of providing an indication of the potential for an imbalance between production and requirements in the future. A deeper understanding of the relationships between metrics and their consequences would require more detailed analysis of agent interactions and decisions (Knoeri et al. 2013).

In common with Graedel et al (Nassar et al. 2012; Graedel et al. 2012) there is no discussion of a threshold of criticality, because these indicators are intended to be used to compare the relative criticality of different pathways to low-carbon infrastructure, rather than defining the point at which criticality becomes unacceptable. This is not to say that it is not possible to define a threshold of this nature rather that it is not the intention of this thesis; criticality thresholds will need to be informed by a combination of political and economic factors as well as a technical analysis of criticality.

5.5 Summary

This chapter describes an indicator set developed to represent the risk of disruption in the supply of critical materials. Analogous to risk, criticality is defined as the product of the potential for disruption in supply and the effect of that disruption on the goal of low-carbon infrastructure transition. This approach was selected because it is better able to represent the short and medium-term disruption than long-term forecasts of critical material availability. The index of potential for supply disruption takes into account the effect of multiple factors on the imbalance between production and requirements, which is the cause of supply disruption. In the same way, the exposure index is comprised of two factors which represent the effect of disruption on low carbon transition. The individual metrics that contribute to supply disruption potential and exposure can be tracked individually to provide more detailed insights into the drivers of criticality for particular materials and technologies.

The relationship between the metrics is represented, to some extent, by the method used to aggregate them. However, the relationships are complex and likely to change. Detailed analysis of this complexity would require dynamic models which can better represent the evolution of relationships between drivers but which are difficult and time consuming to construct. In this thesis, it was decided that a less complex but more transparent approach would be used. Metrics contributing to supply disruption potential are forecast to create a temporally dynamic measure of supply disruption potential. Exposure of low carbon transition to this disruption is also quantified annually. In this sense the dynamic nature of the analysis is related to temporality, the relationships between metrics remain static. Nevertheless, the temporally dynamic analysis allows the estimation of criticality over time and identifies trends of increasing or decreasing criticality.

When demonstrated using the case of neodymium and wind turbines in the UK, the indicator set identifies that material criticality will increase up to ten-fold between 2012 and 2050. The supply disruption potential of neodymium decreases over the period as a result of increasing distribution of production. However, the exposure of low-carbon electricity generation increases significantly as a result of increasing reliance on direct drive turbines.

6 Responses to material criticality constraints

The indicator set described in Chapter 5 aims to not only quantify the constraints from material criticality but also to support analysis of how these constraints could be reduced. The indicator set is purposefully transparent to support analysis of the causes of material criticality constraints and to allow responses to be targeted at the most important causes. The potential responses of UK policy makers to material criticality constraints to low-carbon energy in the UK are described below, grouped by the index to which they would contribute. The case of UK responses to material criticality constraints from neodymium in wind turbines containing permanent magnets has been used as an example to illustrate key points. The effectiveness of responses will vary depending on the material; therefore, the appropriate combination of responses will need to be assessed on a case-by-case basis.

In response to private sector concern about the availability of some raw materials the UK Government has undertaken a review of risks from supply disruption (AEA Technology 2010). As a result, the Department for the Environment, Food and Rural Affairs (Defra) has developed the Resource Security Action Plan (RSAP), which provides a "*framework for business action to address resource risks and sets out high level actions to build on developing partnerships between government and business to address resource concerns*" (Defra 2012). Actions include; addressing barriers to greater recovery; facilitating provision of relevant information; promoting and supporting innovation and research (relating to recovery of resource from waste materials); and engaging the EU and international partners to promote the appropriate international framework.

The government's RSAP is concerned with risks to UK businesses, not to other government strategies, such as the low-carbon transition. It is perhaps unsurprising therefore that some of the constraints described in this thesis are not addressed by the RSAP. However, it is important to take a broader perspective of resource risks and associated responses to address the constraints to low-carbon technology roll out identified in Chapter 5.

6.1 UK response to supply disruption potential

6.1.1 Production-requirements imbalance

The principal points of intervention to reduce the production-requirements imbalance are to increase global neodymium production and reduce the global requirements for neodymium. The UK has greater potential to address the latter through reducing its own requirements for neodymium since it has no neodymium economically viable reserves of its own (BGS 2011). It has four approaches to doing this: reducing total consumption of neodymium through consuming less to deliver the same output, using less resource per unit of technology, substituting a critical material for a less critical material and recovering secondary neodymium to displace requirements for primary material. These responses would also reduce the UK's exposure to supply disruption potential so would have a cumulative effect on material criticality constraints.

6.1.1.1 Consuming less to deliver the same output; resource sufficiency

One of the most effective approaches to reduce material consumption is to provide the same service with fewer units of technology, often called resource sufficiency (Allwood & Cullen 2012). This could be achieved by increasing the output efficiency of each product so fewer are needed to deliver the same level of output. For example, increasing the energy generation efficiency of wind turbines will mean that fewer turbines, and by association fewer neodymium containing permanent magnets, are required to deliver the same energy capacity.

Many technologies, particularly consumer electronics, are scrapped and replaced before the end of their useful lives (Cooper 2005). This not only creates waste but also increases requirements for primary material in replacement products. Increasing the lifetime of products, through behavioural change or improved maintenance could dramatically reduce material consumption and associated environmental impacts, such as GHG emissions (WRAP 2011; Barrett & Scott 2012). Resource sufficiency strategies can be more effective in reducing resource consumption than resource efficiency, which aims to reduce the material used per unit of product or technology (Barrett & Scott 2012; Allwood & Cullen 2012). This is relevant to neodymium, which is used in magnets in hard disk drives.

Resource sufficiency strategies may require alternative business models that allow companies to retain ownership of products, incentivising refurbishment and maintenance. These business models are often grouped under the term “product service systems”, and can extend product lifetime and lead to higher efficiency levels in the supply stream (Mont & Tukker 2006). An often-cited example of this is the Rolls Royce model of selling hours of airtime, rather than jet engines. This incentivises Roll Royce to create robust engines and invest in maintenance regimes to maximise the service life.

An extension of this concept is communal consumption schemes, like car sharing, which further decouple service from number of products required. Instead of owning a number of cars, which are not used for the majority of the time, a group shares ownership of one car.

Therefore the service stays the same, but is delivered with far fewer cars and lower material requirements (Mont 2004). This is particularly relevant to neodymium, which is used in electric vehicles; if the number of vehicles required to deliver the service of mobility could be reduced, this would reduce the consumption of neodymium (and also our exposure to any remaining disruption).

6.1.1.2 Using less resource per unit of technology; resource efficiency

Improving product design or the efficiency of technology production processes could reduce the critical material content of technologies; however, the potential to do this in existing production processes through waste management or lean production is minimal (Morley and Eatherley 2008). The magnetic strength of the most recent generation of neodymium magnets is believed to be close to fundamental and technical limits of this material (Kara et al. 2010). The scarcity and price of these materials has already led companies to develop production processes that are designed to maximise efficiency and minimise waste production. However, Fromer et al (2011) identify material efficiency strategies, which could reduce neodymium content of electric vehicles (Fromer et al. 2011). These include reducing the grain size of compounds in permanent magnets and simplifying shapes for manufacturing. It is unlikely that further reductions could be achieved without associated change in technology or substitution for another material, which is covered in the next section.

6.1.1.3 Substituting for a less critical material

There are a number of examples where a critical material has been replaced with one that is significantly less critical. For example in the 1980s the US military was able to replace cobalt with nickel-aluminium-molybdenum alloys in jet engines owing to increasing cost and security concerns associated with cobalt. However, critical materials have properties and characteristic which make them uniquely suitable for specific applications. For example, samarium-cobalt magnets have similar performance at high temperatures to neodymium but have only half the magnetic strength, making them less suitable for use in electric vehicle motors.

It is also important that an insecure material is not substituted with a material that is, or could become more insecure than the material it is replacing. This is a particular problem among groups of scarce materials with similar properties, such as the platinum group metals or rare earth metals. This is illustrated for materials with properties making the suitable for optoelectronic applications, such as photovoltaics, in figure 6.1. Many of the substitutes for metals in use, having suitable physical and chemical properties, are already in use for this purpose and are already critical themselves. Furthermore, the material with similar properties are often produced as by-products from the same ores, therefore are subject to the same risks of supply disruption (Graedel 2011).

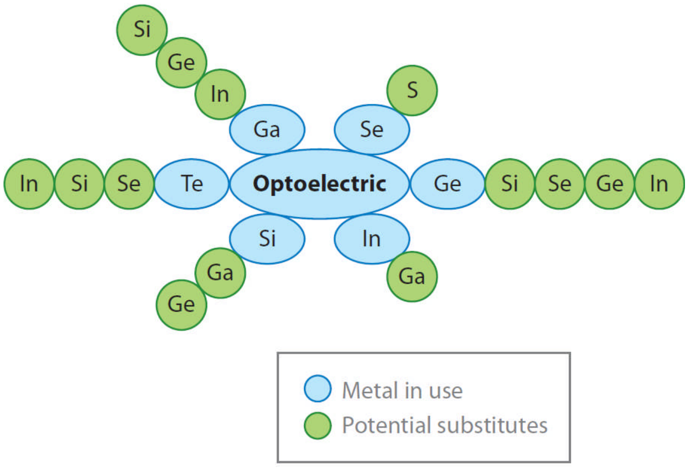


Figure 6.1: Potential substitution of metals for optoelectronics (Source: Graedel 2011)

The only appropriate replacements for neodymium in permanent magnets are dysprosium or praseodymium, both of which are more expensive and potentially more critical than neodymium (Kara et al. 2010). However, this diversification of materials used does spread the risk and lowers the overall risk if substitution is not complete. Nanotechnology is improving the opportunities for material substitution. For example nanocomposites combining NdFeB and FeCo nanoparticles could also reduce the rare earth content of permanent magnets without affecting performance (Kara et al. 2010).

In some cases, the specific properties of the critical material are not substitutable within the current technological configuration (for example conventional automobile engines require a minimum proportion of platinum in catalytic converters). In these cases, it is necessary to amend or substitute the technology itself to avoid risk of material criticality constraints. For example, recent developments have identified a series of technology substitutions that might reduce neodymium content of motors in electric vehicles and wind turbines. Speirs et al (2013) undertook a detailed review of technology substitution for neodymium in low-carbon technologies (Speirs, Houari, et al. 2013). Neodymium is needed in permanent magnets to retain magnetic power at high temperature; therefore, improving cooling systems in electric vehicle motors could reduce neodymium requirements (Fromer et al. 2011). The Japanese car manufacturer Toyota is developing a new induction motor to reduce its reliance on China for rare earth metals (Electric Vehicle News 2011). The new motor uses electromagnets, rather than permanent magnets avoiding the need for neodymium. Another example is hybrid motors, which are under development that combine permanent magnet motors with switched resistance motors, which require less neodymium (Schüler et al. 2011). Similar hybrid systems have been proposed for wind turbines, which include a permanent magnet direct drive train and a gearbox, which can reduce rare earth metal content by up to 80% (de Vries 2011).

6.1.1.4 Recovering secondary material

An important way to reduce primary material requirements is to recover material that is wasted during the manufacturing process or to recover material from goods that have come to the end of their life, avoiding dispersal of the materials into the environment. The recovered material can displace primary material, which reduces global requirements. Recovery is predominantly discussed in terms of recycling, which implies some form of processing to extract individual materials from compounds or composites, but materials and components can often be re-used without reprocessing or technology can be reconditioned.

Theoretical recycling efficiency of goods containing critical metals can be high; for example the efficiency of recycling neodymium from permanent magnets can be as high as 90% (Binnemans et al. 2013) and the matter of critical material recycling is rapidly moving up the political agenda (European Commission 2011b). However, potential end of life recycling rates of critical materials far exceed the actual performance. A study by the UN's International Resources Panel found that end of life recycling rates of many critical metals necessary for low-carbon technologies (including neodymium, indium, tellurium and lithium) were less than one per cent (see figure 6.2).

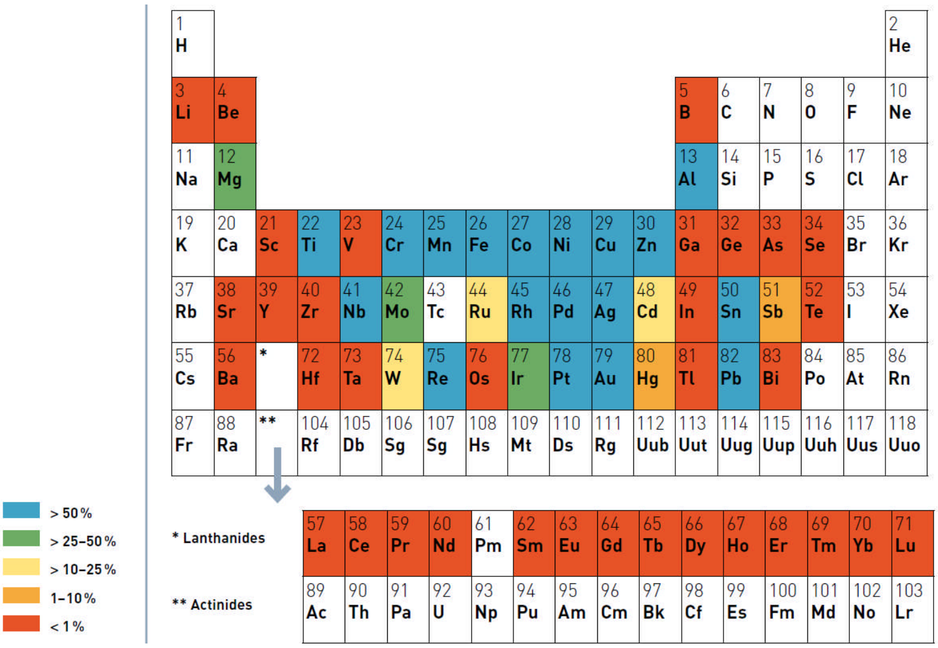


Figure 6.2: Global average end of life (post-consumer) recycling rate for sixty metals. Recycling is defined as a process whereby the physical and chemical properties that made the material desirable in the first place are retained for subsequent use (Source: UNEP 2011)

Recovery of materials from end of life goods has a great deal of potential; Du and Graedel estimate that there was 137,000 tonnes of neodymium in stocks of goods in 2007, almost ten times the annual production of primary neodymium (Du & Graedel 2011b). The stocks of neodymium in permanent magnets alone are almost four times the annual extraction rate (Du & Graedel 2011a). The efficient recovery of these stocks could provide a valuable supplement to primary production. However, critical materials are predominantly used in extremely small quantities or in composites and alloys, which makes recovery challenging as a result of inefficient collection, difficulties in separation and the high cost of recovery processes (Reck & Graedel 2012; UNEP 2011). As a result, the majority of used permanent magnets are disposed of as scrap (Du & Graedel 2011b; UNEP 2011).

The diffuse use of neodymium in small magnets in short-lived consumer electronics means that collection rates are very low (Rademaker et al. 2013). More concentrated sources, such as wind turbines and electric vehicles, are starting to enter the market but have long lifetimes so will take a long time to reach recycling plants. This reduces both the efficiency of the overall recycling process (Binnemans et al. 2013) and the viability of recycling facilities, which need significant feedstocks to achieve economies of scale. Producing forecasts of stocks coming out of use (such as (Busch et al 2014)) will enable better planning of collection infrastructure and associated recycling facilities.

The majority of electronic waste undergoes shredding and sorting prior to recycling, which breaks the brittle magnets into granules, which, because of their residual magnetism, stick to other ferrous waste and become difficult to separate. Ideally, magnets would be separated before the shredding phase, but this is difficult because of the complex, and varying, physical binding structure for magnets in end of life products. Greater attention should be paid to the design of products to enable more effective separation, improving their recycling potential and reducing the cost and environmental impact of recycling. Design measures could include labelling alloys and other composite materials to make sure people are aware of the valuable materials embedded in the alloy (increasing the likelihood or reuse or recovery) and designing products that are easier to disassemble.

Despite a significant amount of research into rare earth recycling technologies there are very few commercially developed. The majority of active recycling facilities are based around three principal techniques: pyrometallurgy; hydrometallurgy; and electrometallurgy (Binnemans et al. 2013). Pyrometallurgy uses high temperature to melt composites and chemically separate metals. This is a very energy intensive process and can cause some oxidisation of rare earths, which makes full recovery difficult. Hydrometallurgy uses strong acidic or alkali solutions to selectively dissolve and then precipitate metals of interest from a pre-processed powder form. The fundamental processes used for recycling are the same as those used for raw ore. Electrometallurgy uses a current to separate rare earths from a liquid-leach solution containing the metal by either electroplating or electrowinning. Forecasts of future stock coming out of use will help to drive the development of recycling facilities but this may need to be supported by investment (Du and Graedel 2013b).

The recycling process can be very costly and have a high environmental impact. For example, pyrometallurgical processes require high temperatures (over 1400C), which requires a great deal of energy, and therefore high carbon emissions (Binnemans et al 2013). Therefore, the lifecycle impacts of recycling strategies should be considered when evaluating them in a low-carbon economy.

The high environmental and economic costs of recycling can be avoided if the magnets from end of life products are re-used or reconditioned without processing. Direct re-use is only really possible for large magnets from wind turbines, or possibly electric vehicles (Binnemans et al. 2013). Smaller magnets, from hard disk drives for example, can be reconditioned through powder processing and re-sintering of magnet granules; however, this is only possible if there is a narrow distribution of magnet composition among manufacturers. In other cases, it would be necessary to separate neodymium from other elements in the alloy in a recycling process.

6.1.2 Access (HHI)

A lack of competition in neodymium production could be overcome through encouraging the diversification of production, which is currently predominantly in China. This could be achieved through exploiting UK deposits or supporting mining in other countries. There has been no comprehensive evaluation of rare earth resources in the UK; however rare earths are known to occur in a number of locations (BGS 2011). Most are small occurrences that are not economic to extract so it is unlikely that they could be exploited.

Perhaps a more productive approach would be to encourage mining in countries with more significant deposits than the UK. This could be through foreign investment in mining projects and in associated infrastructure in countries with more significant deposits. A series of mines outside China are at an advanced stage of development which would contribute to this diversification (BGS 2011). Anti-competitive behaviour could be further reduced through trade policy, by developing partnerships with producing countries. This is being promoted by the European Union (EU), as part of its Raw Materials Initiative and it would be beneficial for the

UK to engage fully in these activities (European Commission 2008b; European Commission 2011a). An EU trade strategy for raw materials has been developed which promotes trade negotiations, work to tackle barriers to raw material trade and outreach activities (European Commission 2012b). Free Trade Agreements and Partnership Agreements have been established with a number of countries, which include rules to achieve sustainable supply of raw materials. Furthermore, the EU is active in challenging measures which violate World Trade Organisation or bilateral commitments; such as export restrictions.

6.1.3 Environmental Constraints (ECR)

It has been suggested that restrictions on expansion of mining could be alleviated by simplifying mining regulations and reducing permit processing times (European Commission 2008b). In support of this policy the EU has produced guidance on the management of the competing objectives of ensuring a high level of environmental protection in Natura 2000[19] areas and the development of competitive extractive activities (European Commission 2010a). In particular it provides guidance on undertaking Appropriate Assessments for mining activities and promotes mitigating activities, such as rehabilitation and biodiversity off-sets so that mining is not automatically excluded in protected areas. There is little potential for rare earth element mining within Europe but this guidance could be shared with countries that do have significant reserves.

6.2 UK response to exposure

6.2.1 Goal sensitivity

The principal cause of material criticality constraints to low-carbon technology roll-out is the increasing exposure to disruption in critical material supply. This exposure results from an increasing reliance on specific technologies, which in turn are reliant on critical materials. One of the most effective responses to this exposure could be to reduce this reliance on individual technologies by diversifying the technology contributing to low-carbon transition. This has the added advantage of also contributing to a reduction in global demand for critical materials and reducing supply disruption potential.

Diversity can also mitigate future lock-in, hedges ignorance (with regard to both future demand and supply) and also offers a means to promote innovation (Stirling 2007). As an example of this; electric motors which require permanent magnets are favoured in the UK and the majority of future transport scenarios are entirely based on this technology, leaving us

---

[19] A network of sites designed to safeguard Europe’s rarest and most endangered species and habitat types defined in the Habitats Directive (European Commission 1992)

exposed to supply disruption in neodymium. Technology diversity would require a move away from a sole focus on vehicles containing electric motors to encompass hydrogen or other fuel cell vehicles. However, diversity, and the ensuing resilience under shock and robustness under stress, is more than just the presence of many technologies: "*diversity is generally a state under which an observed system is seen to display: (1)* ***even balance*** *across (2) a* ***variety*** *of (3)* ***mutually disparate*** *categories"*(Stirling 2011). In the example given above; balance would mean that shares between electric vehicles, hydrogen and other fuel cell vehicles would have to be similar; variety and mutual disparity would mean that the vehicle technologies would not rely on the same critical material.

This diversity in low-carbon technology has implications for supporting infrastructure in the UK and the wider UK economy. A diversity of technologies requires a flexible and varied network of supporting infrastructure. Taking the vehicle example given above; a system of both electric and hydrogen charging stations would be required, but of a much smaller scale than if one technology was dominant. This means that if there is some change in the future which precluded the use of electric vehicles, or a more appropriate alternative was discovered, there is less investment in a stranded asset and the UK is less likely to be locked into this technology. In this way the transition to a low-carbon future is more adaptable. Furthermore, diversifying the output of the UK vehicle manufacturing sector makes the sector itself is more adaptable to future changes in technology and infrastructure.

6.2.2 Price sensitivity

In addition to reducing the amount of material used per unit (as discussed above) or substituting critical materials for other, cheaper materials, the UK could take action to reduce the volatility of critical material prices. There was approximately a 15-fold increase in finance and investment flows into commodity derivative markets between 2003 and 2008, which is considered to have exacerbated the price volatility of critical metals (European Commission 2011a). The relationship between price volatility and financial markets is complex and uncertain but the EU is proposing a suite of regulations to improve raw material market transparency and avoid market distortion. The EU market access partnerships described above may also serve to stabilise price volatility as a result of long term relationships between producers and consumers. The UK could support these proposals and use its trade relations to reinforce action at an EU level.

6.3 Summary

There is a great deal of potential to reduce global demand for critical material through lifetime extension, design and production efficiency and recovery of secondary materials. However, there remains a myriad of challenges to implementing technically feasible opportunities. Research is needed to support this implementation, including into alternative business models to reduce product consumption. There is a major role for policy makers to create the conditions to encourage and enable this reduction of material consumption in the UK, including regulatory frameworks requiring design for longevity and recovery. This role includes more targeted support for alternative business models and product design to extend the lifetime of products and make it easier to recover critical materials from end-of-life products.

One of the most important responses to material criticality constraints is to reduce exposure to supply disruption by increasing the diversity of technologies. This requires not just a number of different technologies but also a balance between technologies that rely on different critical materials. Furthermore, the supporting networks which connect technologies must be adaptable to enable this diversity and respond to future changes in technology roll out.

The responses described in this chapter required cross-departmental policy and co-ordination; reducing resource consumption rests predominantly within Defra, however, technology diversity is more likely to come under the remit of DECC and the Department for Business, Innovation and Skills (BIS). Responsibility for the RSAP is shared between Defra and BIS so, to some extent, there is a precedent for this cross-department action. However, this must be expanded further to include DECC to ensure co-ordination between objectives and strategies.

7 Mainstream infrastructure operation and alternative modes of operation

This chapter contrasts the characteristics of the mainstream mode of infrastructure operation with that of the cases of alternative operation analysed as part of the meta-case study. This is necessary because policy and regulation has co-evolved with the characteristics of the mainstream (in terms of assets, actors, processes and relationships) so it is important that these characteristics and the relationship with policy and regulation are clearly understood prior to analysis. As a result of this co-evolution, selection pressures in the regime tend to favour mainstream modes of operation, which can create significant barriers to the development of alternative modes of operation (Mitchell & Woodman 2010; Smith et al. 2005). Therefore, in this chapter the characteristics of the mainstream mode of energy and water infrastructure operation, with which the current policy and regulation has co-evolved, are briefly described in section 7.1.

The individual cases analysed in support of this thesis were intentionally diverse in terms of both the initiator of the alternative mode of operation and the nature of revenue generation. Nevertheless, analysis uncovered a series of common characteristics of the meta-case study of alternative modes of infrastructure operation. When contrasted with those of mainstream infrastructure operation these alternative characteristics reveal important differences which are crucial in explaining unfavourable selection pressures or a lack of resources available to respond to these selection pressures (Smith et al. 2005). This will help to expose the root cause of constraints described in chapter 8 and will also inform recommendations for responses to these constraints in chapter 9. The common characteristics of the meta-case study of alternative modes of infrastructure operation examined in this thesis are described in section 7.2.

7.1 Mainstream infrastructure operation

Both the water and energy sectors are dominated by a small number of large private utility companies. For example, despite regulation to disband monopolies and create a ‘free’ infrastructure market, the market is dominated by a small number of large, international utility companies, for example 98% of household gas and electricity is supplied by only six energy companies (Ofgem 2013a). The operation of both energy and water infrastructure is characterised by a throughput-based model, whereby greater profit is made by increasing the number of units of utility product sold (for example kWh electricity), or by increasing the marginal cost efficiency of producing a unit of energy. This mode of operation, in combination with the prioritisation of short-term income, discourages demand reduction (which reduces throughput) and investment in low-carbon technologies (which increase marginal costs in the short-term) (Roelich et al. 2014). Both sectors have been able to deliver significant efficiency savings by 'sweating assets'[20] built and funded during previous state-owned eras (Helm 2009; House of Lords Science and Technology Committee 2007). The specific characteristics of mainstream energy and water operation are described below.

7.1.1 Energy

In the early 20th century, energy was provided at a municipal level by a range of public and private actors and the UK's energy system was dominated by coal (Fouquet & Pearson 1999). Gas and electricity systems were small and localised and evolved to serve specific users and locations (Hughes 1983). As technology (particularly electric motors) progressed electricity networks were developed to connect new users to sources of electricity generation. Initial networks were private and independent with little harmonisation of technical standards (Patterson 1999). The 1920s saw the start of a phase of standardisation and centralisation to improve economies of scale. The national grid emerged during the inter-war period and universal technical standards were imposed by national government. Electricity supply was nationalised in 1947 and gas a year later, accompanied by amalgamation of over 1000 companies into 12 regional gas boards (National Grid 2005). This centralisation was reinforced by the development of progressively larger power plants, including the world's first civil nuclear power station (Marshall 2010; Sherry 1984).

The energy sector underwent a process of privatisation in from 1986 to 1995. During the 1990s, generation and supply were separated and the retail markets were liberalised to enable competition for both electricity and gas (Marshall 2010). At the end of 2012 there were 34 major power producers in the UK, a number of which are under the ownership of international companies (Office of Fair Trading 2010). The majority of electricity (88%) is produced at large, thermal power plants, which are typically in the order of 1-2GW[21] (DECC 2013a). Thermal plants generate electricity by heating water to convert it into steam, which then powers steam turbines. Sources of heat include combustion of coal, gas, oil and biomass or nuclear fission and thermal efficiency is typically 35-50%[22].

---

[20] Maximising the value extracted from existing assets whilst minimising investment.

[21] This is highly dependent on the fuel and can range from 0.05GW for gas oil to 4GW for coal.

[22] Thermal efficiency is the efficiency with which heat energy contained in fuel is converted into electrical energy. It is calculated for fossil fuel burning stations by expressing electricity generated as a percentage of the total energy content of the fuel consumed (based on average gross calorific values).

The electricity generated by power producers is sold through the wholesale market, BETTA[23], using a centralised balancing system as well as long term contracts and bilateral trading (see figure 7.1). There are similar wholesale market arrangements for gas. Since liberalisation there have been commercial moves towards vertical reintegration between generation, distribution and electricity supply, so the same parent company might own an energy generator, distribution network operator and an energy retailer (Ofgem 2011). This means that vertically integrated companies conduct internal deals in order to reduce the risk (and penalties) of imbalance. As a result, very little electricity is sold through the balancing system (Bolton 2011). There have been moves to limit the proportion of electricity that can be sold internally but it still outweighs freely-traded electricity (Ofgem 2011).

National Grid owns and operates the high voltage transmission system in England and Wales. The transmission network transports power from generation to sub-stations which are closer to sources of demand and a charge is levied to generators (see figure 7.1). In 2005 National Grid took on responsibility for operating the system in Scotland and became the overall transmission system operator (TSO) responsible to ensure demand for electricity can be met at all times (National Grid 2005). This includes operation of a balancing mechanism in the wholesale market to ensure that demand and generation are kept in balance, and transmission constraints are respected.

Electricity is transported from substations through the distribution network to end-users. The 14 regional distribution networks are owned by nine distribution network operators (DNOs). Unlike the high-voltage transmissions system, electricity flows one-way in the distribution systems and they are not designed to accommodate input of distributed generation. Suppliers pay a charge to DNOs for the use of the distribution system.

For nuclear stations it is calculated using the quantity of heat released as a result of fission of the nuclear fuel inside the reactor.

[23] The British Electricity Trading Arrangements

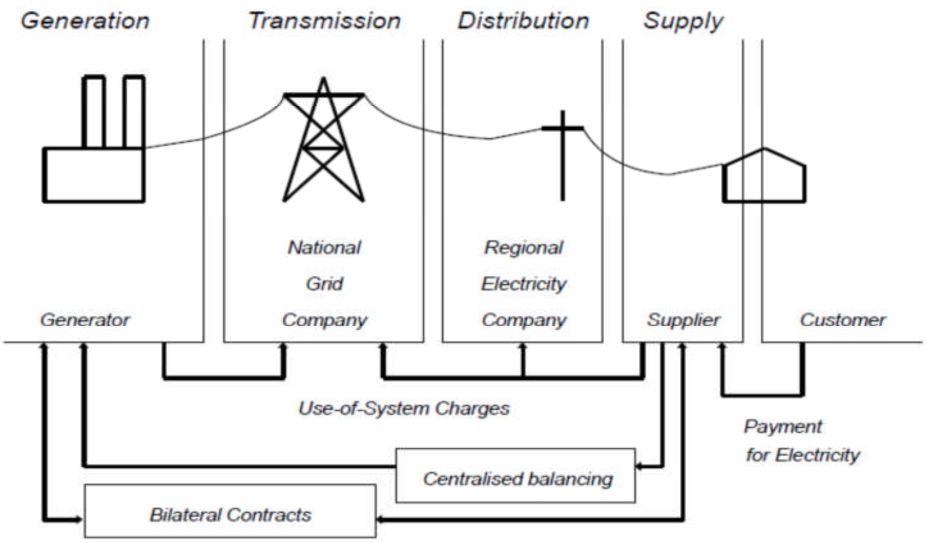


Figure 7.1: The electricity market in the UK (Green 2011 reproduced in Bolton 2011)

There are more than 70 licenced electricity supply companies but the retail market is dominated by six companies; Centrica, EDF energy, E.On, Npower, Scottish Power and Scottish and Southern, often called the 'Big Six'. The majority of these companies are vertically integrated, international companies. Vertical integration has a number of implications for competitiveness in the retail sector; bilateral contracts reduce transparency of energy trading, making company profits difficult to trace and making it difficult to have true competition (Ofgem 2011). As a result, retail energy prices have continued to rise despite little increase in the wholesale cost of energy (Lee 2013).

Business arrangements in the energy sector treat energy as a commodity; the more units of gas or electricity that are sold, the greater a company's revenue (Patterson 2013). As such, both energy producers' and retailers' interests are in maximising sales and reducing the short-term costs of producing and distributing each unit of energy. Neither of these drivers is compatible with the goals of reducing energy demand or increasing uptake of low-carbon technologies.

7.1.2 Water

Much of the existing water infrastructure in the United Kingdom originated in the mid-nineteenth century and, between this time and the end of the Second World War, was under the control of local authorities (Hall et al. 2012). The Water Act 1945 amalgamated local authorities with local boards to set up River Boards, which were then replaced in 1963 by River Authorities (HM Government 1945). In 1973, perhaps the most important change in England and Wales came about through the Water Act 1973, which established ten Regional Water

Authorities (RWA), further removing water from municipal control (HM Government 1973). Up to this point, capital investment in the industry had increased year on year, tripling in real terms since 1953. However in 1979 the Government instructed the state-funded RWAs to reduce investment. By the 1980s, investment had fallen to between a quarter and half of what it had been ten years previously and the water infrastructure had deteriorated to such an extent that an estimated £24 million was needed to upgrade it to meet obligations under the EC's water-related directives (House of Lords Science and Technology Committee 2006). The Government was reluctant to fund this from the public sector so it hoped that privatisation would allow necessary funds to be raised by the private sector (House of Lords Science and Technology Committee 2006).

The English water sector was the first in Europe to be fully privatised and the ten RWA and associated water infrastructure assets were placed into companies which were floated on the London Stock Exchange under the Water Act 1989 (House of Lords Science and Technology Committee 2006). Scottish Water and Northern Ireland Water have been retained as state-owned companies, rather than private companies. The government originally retained a share of the privatised water companies but this expired in the mid-1990s and since then, the ten original English and Welsh water companies have changed ownership a number of times. Ownership was originally dominated by foreign investors but regulation of prices introduced in the late 1990s led most foreign companies to withdraw from the water sector (Hall et al. 2012). Current ownership is dominated by private limited companies and investment or pension funds (Office of Fair Trading 2010). This can have a significant influence over decision making processes: " *[Water Company] was still four or five years ago a listed company…I would say that since it's been bought out by investment banks it has been delisted, that has started to change it a little bit. Innovation is no longer seen as a long-term priority. The business is all about short-term, quick-fix solutions"* (Water Company Manager 2013).

Since privatisation a number of smaller-scale water and companies have been set up to supply water and wastewater to specific localities, such as Albion Water and Veolia Water Projects, and some which supply water only, such as Bristol Water and Portsmouth Water Plc (Water UK 2014). However, the majority of water is supplied by the ten vertically integrated, regional water and sewerage companies created by privatisation. In contrast to the energy sector, liberalisation of the water was resisted, with only a small portion of the market open to retail competition (Defra 2011b).

Water supply infrastructure comprises a regional or local network distinguished by hydrological boundaries and, to some extent, historical ownership patterns. There are few strategic interconnections between regions, with each region relying on its own sources of water. Waste water infrastructure is characterised by sewers transferring wastewater to a centralised treatment system. The network and treatment systems vary in size but most people are served by a small number of large systems. This is deemed to be more efficient, particularly as the intensity of treatment increases as a result of increasing discharge standards (Hall et al. 2012). Business arrangements in the water sector mirror those of the energy sector with water sold as a commodity. This is reinforced by economic regulation which controls the price per unit of water (Bakker 2005).

7.2 Alternative modes of infrastructure operation

The individual cases used in support of analysis of this meta-case study were selected to represent a range of modes of operation. This was done deliberately to enable examination of the diversity of motivations and configurations of alternative modes of operation that would need to be supported by future infrastructure governance systems. In this way, recommendations can be made on the basis of encouraging diversity, to avoid future lock-in to another dominant but unsustainable mode of operation.

7.2.1 Individual case summaries

The individual cases, summarised in table 7.1, were selected to cover a range of initiators and revenue generation mechanisms. An initial desk study was undertaken to identify potential cases and a shortlist of five was generated. The remaining cases were identified in the course of undertaking these initial cases. For example, during the interview with Newcastle City Council, it became apparent that the Byker Community Trust was planning a very relevant and potentially useful project. This opportunistic approach to case selection identified a number of cases which had yet to receive much attention but which were very relevant to this analysis.

Table 7.1: Summary of individual cases used in support of this thesis

Case	Description
Woking Borough Council	A Special Purpose Vehicle (including participation of the private sector) designed to save energy in council property and recycle financial savings to fund low-carbon energy generation and further energy saving measures.
Olympic Park Energy Centre	The largest Combined Heat and Power Plant in the UK providing heat to the Olympic park. Now extending heat supply to surrounding developments.
Water Company Water Saving Trials	Pilot schemes trialling water saving device roll out to provide evidence on water saving potential in response to Ofwat's requirements to provide demand and supply balance and demonstrate action on demand management.

Case	Description
Kimberley Clark Water Recycling	"Long loop" recycling of effluent from paper mill re-used as process water following intensive treatment. Novel financing mechanism (piggy-backing on a high payback scheme).
Eco-Island	Retrofit and low-carbon energy schemes on the Isle of Wight by a not for profit CIC. Planned to be undertaken by an ESCo run in partnership with the CIC.
Welborne	Proposed development of 3,000 house and employment land. In addition to high levels of sustainability (and updatability), it is proposed that properties and infrastructure will be managed by the community that occupies the site.
Chale Green Housing Trust	Retrofit of social housing using novel financing scheme to 'rent' roof space to a solar panel provider. Insulation and air source heat pumps were provided to increase thermal comfort and reduce energy bills.
Newcastle City Council	Proposed joint venture to deliver district heating networks in council and private property. Less profitable schemes bundled with more financially viable schemes.
Byker Community Trust	Transfer of ownership and operation of estate district heating system from local authority to community trust

Where secondary data was insufficient, interviews with key participants were undertaken to identify the relationships, processes and capabilities which had enabled their development. The interviews went on to explore the barriers that individual cases had faced or that they felt that other actors would face if they tried to replicate the mode of operation in other circumstances or on a bigger scale. This provided direct insights into regulatory and policy constraints but also provided data relating to the different characteristics and organisation of actors and assets, which informed more in-depth analysis of constraints. These aspects of the case study analysis are discussed in the following sections.

7.2.2 Differences from mainstream infrastructure operation

Data from literature and participant interviews were used to develop a representation of the actors and assets involved in each case study. This was initially structured using the five systems from Foxon's co-evolutionary framework and the stages of the infrastructure supply chain, from generation to end user to ensure that the full range of actors and assets was addressed (Foxon 2011). A similar process was used to create a representation of the mainstream mode of operation for the infrastructure which was the focus of the case study. An example of these representations for Woking Borough Council is provided in figure 7.2.

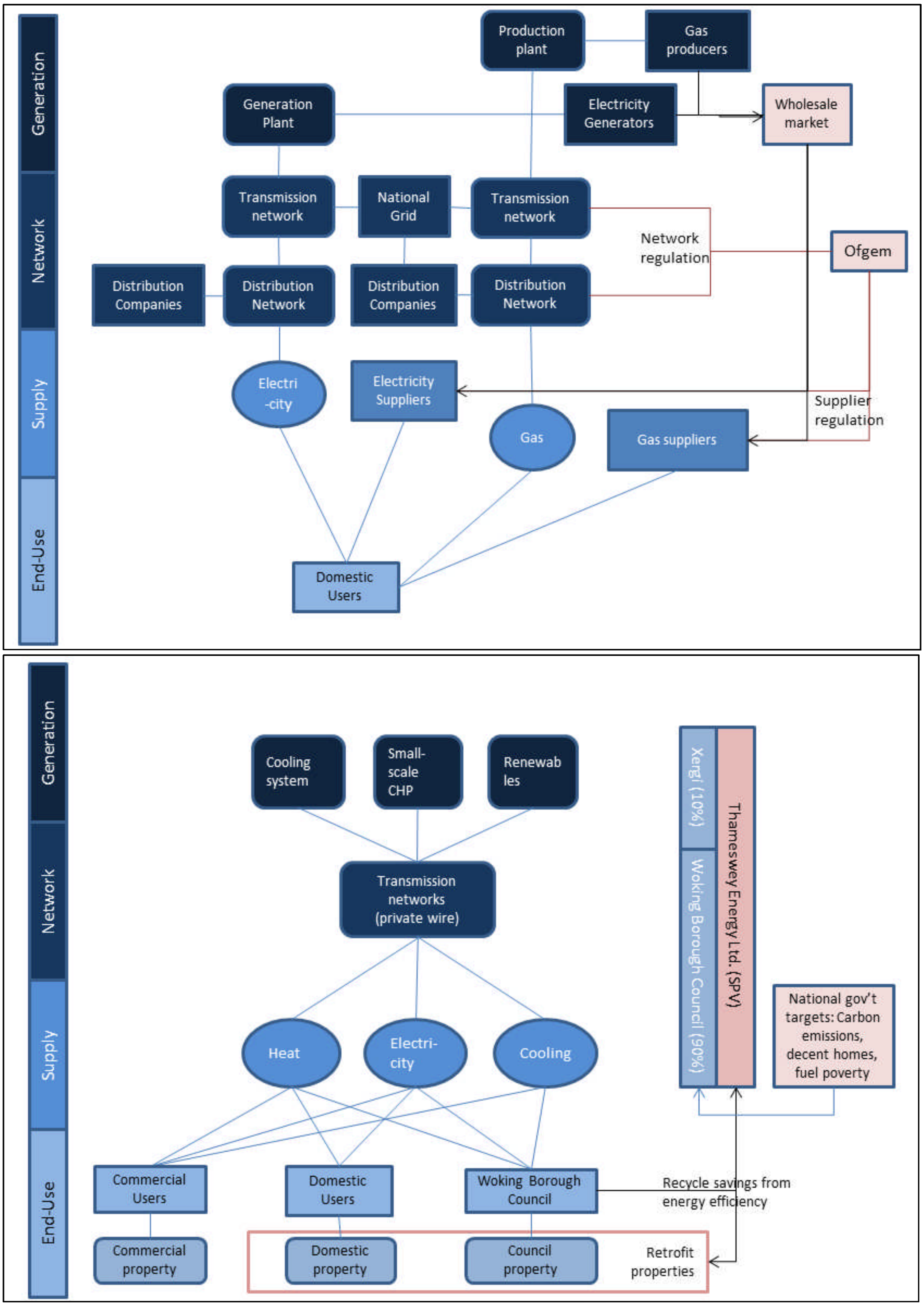


Figure 7.2: Contrasting representations of energy operation; a – mainstream; b - Woking Borough Council

The five systems from Foxon's framework have been removed from the representation for clarity. Comparison of the two representations helps to articulate the principal differences between the individual modes of operation and helps to understand where current policy and regulatory approaches may be unsupportive.

Further to this, a detailed description of the case study was produced, grouping insights by each of the five systems identified in the analytical framework. An example of an individual case write-up is provided in Appendix 1.2. Emerging insights, in relation to the characteristics which differentiate the case from the mainstream, and the constraints to further roll-out of the case were also recorded in this write up (see Appendix 1.2 for an example).

This analysis was undertaken for each individual case (within case analysis) then the emerging insights from all cases were compared to identify differences in characteristics common to the majority of cases (between case analysis). These common characteristics were taken to be representative of the meta-case of alternative modes of infrastructure operation. The differences identified include the motivation of initiators, the scale of operation, the extent of integration of operation, the prevalence of partnerships, the presence of a long-term vision and iterative development and the use of alternative financial models. Each difference is described in more detail below using examples from individual cases for illustration.

7.2.2.1 Motivation and value derived from infrastructure operation

Many project initiators had motivations, other than revenue generation, for becoming involved in infrastructure operation. These motivations were varied but tended to be socially oriented, such as fuel poverty (in the case of Newcastle City Council, Chale Green and Byker Community Trust), self-sufficiency (in the case of Eco-Island and Welborne), resilience (in the case of Byker Community Trust) and community cohesion (in the case of Welborne). These motivators strongly influenced the configuration and organisation of alternative modes of operation. For example; the creation of a network of heat networks in Newcastle aimed to use revenue generated by commercially attractive networks to off-set reduced returns from schemes which would deliver a significant reduction in fuel poverty (Jones 2013; Jones 2012). Assets were often configured in a less cost effective way if they provided greater value during operation. For example, Byker Community Trust had an over-capacity of generation but had done this to diversify the fuel mix and reduce the exposure of residents to price rises resulting from volatile fuel costs (Beadle 2014). A common feature of these alternative motivations, and the social and environmental value derived from alternative modes of infrastructure operation is the challenge associated with capturing the value and using it in support of decisions (Gibbon & Dey 2011).

Many of the cases had multiple motivations, a finding corroborated by Seyfang et al (2013) who uncovered a wide range of goals from community energy projects and found an average of eight objectives per project. This plurality in motivations extends to local authorities, who engage in infrastructure for a range of reasons, including cost and environmental impact reduction and improvement in the quality of life of citizens (Bale et al 2012).

7.2.2.2 Integrated outcomes.

The breadth of goals displayed by the cases provided more opportunities for actors from different parts of the system to engage in the same project. A more integrated, outcome-focussed approach to problem framing and vision development was observed to support the alignment of stakeholder strategies, maximise the range and scale of potential benefits delivered and reduce negative unintended consequences. This finding is supported by work undertaken by researchers at the STEPS centre at Sussex University (Smith 2007; Leach et al. 2010). This engagement is exemplified by the Climate Change Strategy at Woking Borough Council which "*by encompassing a wide range of different objectives, spanning the breadth of services and activities undertaken by the Council, the Strategy ensures that the objectives for responding to climate change are embraced holistically across the Council"* (Thorp, 2011: 77).

The integration of supply and end-use was central to the effective implementation of many of the alternative modes of operation. This is evident in the representation of the Woking case study in figure 7.2; property retrofit was core to the business model of the project. Perhaps one of the most successful examples of this was at the Chale Green housing project where a project headquarters was set up within the community to involve local residents as the project developed. Following installation it acted as an advice centre to support appropriate operation of new equipment and encourage efficient use of low-carbon energy. On completion of the low-carbon project the residents were reticent to leave the centre and it has since become a hub for other community activities such as local food production and training. The close engagement of the end users is cited as a key factor of the extraordinary rate of uptake of technology and has resulted in demonstrable reductions in crime and vandalism (Stanford Clark 2013a; DECC 2012b). These findings are supported by Walker and Devine-Wright, who find that "*direct and substantial involvement of local people in a project can contribute to greater project acceptance and support*" (Walker & Devine-Wright 2008: 499). This "substantial involvement" goes beyond simply consulting or engaging end-users in pre-determined projects; it entails the involvement of end-users in planning, design, delivering, managing or evaluating a project (Yu et al. 2012). This involvement is increasingly termed co-production and is of great relevance to infrastructure. As well as facilitating adoption of new technologies, co-production could have a positive impact on understanding and support for sustainability (Walker & Devine-Wright 2008).

7.2.2.3 Scale of operation

The scale of the alternative modes of operation was far smaller than mainstream modes of operation in terms of both the technology employed and the market served. The former incumbent suppliers (the six vertically-integrated supply companies) supply over 98% of UK domestic customers, which dwarfs alternative providers (Ofgem 2013a)(Ofgem 2011). The largest active cases explored in this thesis were Byker Community Trust and Woking Borough Council, who supplied in the order of 2000 customers each. These cases are at the higher end of the scale of alternative providers yet operate facilities that generate in the order of MW, not GW of energy. Woking Borough Council is one of the more advanced schemes of its type but generates 2.6MW electricity, 3.2MW heat and 1.7MW cooling from a range of facilities across the city (Thorp 2007).

The scale doesn’t only affect the technology and number of customers but also the resources available to the initiators. Many alternative modes of infrastructure operation must find forms of financing which can deal with their smaller scale, which can mean accepting higher rates of interest or more restrictive terms. Furthermore, their smaller scale can mean that organisations lack the breadth of technical expertise available to larger incumbents, for example the technical, financial, contractual and regulatory capabilities necessary to navigate project development and operation. In the case of Byker Community Trust, the trust, which managed the estate and would take on the district heating scheme had only six employees. This is supported in the literature for both community groups (Seyfang et al. 2013) and local authorities (Hawkey et al. 2013; Bale et al. 2012) who identify that these organisations struggle to resource infrastructure projects as a result of limitations on finance, expertise and time.

7.2.2.4 Partnerships

Many of the case studies developed partnerships between different types of organisations to reduce financial risk for both the private and public sector, increase trust from end-users and increase resources and capabilities available to the project. Local authorities tend to be risk averse as a result of budgeting processes and historic factors (Hawkey et al. 2013). However, involvement of a private sector partner to create a joint venture can mitigate some of the risk of investing in and operating new facilities: “*and I was always of the opinion that, especially for a local authority, you don’t want to take all the risk yourself. You want to have some of the* *benefits and you want to have some of the ownership and the rights that come through it, but you don't want to take all the risk"* (Jones 2013). Public private partnerships helped to overcome this barrier in both Woking Borough Council and Newcastle City Council (Thorp 2011; Jones 2013).

Local authorities are in a very privileged position of trust and their involvement can increase acceptance from citizens and increase engagement of end-users (Jones 2013). Infrastructure operation requires a considerable range of skills and resources and the inclusion of partners can increase the scope and potential of the project. The housing officer at Chale Green ultimately drove the project but the intervention of a smart technology expert at IBM led to the project taking a much greater focus on energy management and demand management, rather than just sustainable energy supply (Stanford Clark 2013a).

All cases examined needed some form of agreement or rules and ways of working to maintain alignment of interests and ensure that risks and benefits were fairly distributed. These rules and ways of working, and the organisations or actors needed to implement and monitor them, tended to be unique to each case study and often took a great deal of time and expertise to establish. There have been some attempts at standardisation of partnerships, particularly public-private partnerships, but these have struggled to reconcile private sector participation and sustainability (Koppenjan & Bert 2009).

7.2.2.5 Long term vision

Many of the cases had clearly articulated and long-term visions, which had a very positive effect on project development and stability. A vision of desirable project outcomes provided a means with which to engage stakeholders, allowing them to envisage their role in the project more easily and identify what they might get out of being involved. *"[I] try and sell to them the concept and the vision in the hope that I can get them to take a longer term view about putting money in here to create the type of infrastructure we need, to create the longer term facilities that the community needs; at the same time balancing that against their short to medium term returns that they want from the project"* (Bench et al 2013).

A long-term vision also meant that decisions in the short term were made on the basis of their contribution to the long-term plan, rather than just short-term profit. An excellent example of this was the refurbishment of the district heating system at the Byker Estate. A full technical study was done to identify all work needed to upgrade the generation equipment, heating network and in-house controls and integrate additional connections to the main heating system. The work was planned in such a way than income from one part of the project (such as Renewable Heat Incentive from a biomass boiler) was used to fund others, such as additional connections (Beadle 2014). It also meant that individual measures could be assessed based on their contribution to long-term, system-wide efficiency. This meant that measures which did not offer sufficient returns in the short-term could be justified in terms of their benefits in the longer term. Furthermore, this increased the scale of the overall project, which was more effective in attracting partners, such as technical designers. The Trust was unable to attract any of the Big Six to invest in the primary network, under the Community Energy Saving Programme (CESP) until the project was aggregated with a replacement boiler, after which it was of sufficient scale to be of interest but still only one energy company bid to engage in the project (Beadle 2014).

The alignment of the long-term vision with external drivers or worldviews is important to gaining support. This is very challenging to maintain in the face of the constantly shifting personnel and leadership in local authority and civil service departments. This can affect projects running both within and outside a local authority. The energy masterplanning project at Newcastle City Council was conceived during Liberal Democrat Leadership of the council and neighbourhood regeneration was one of the primary focusses of the leadership. The energy masterplan contributed to both the sustainability and decent homes goals of the regeneration department and was seen to be central to its success. However, just before the masterplan reached the implementation phase the council leadership reverted to Labour and the regeneration department was disbanded. At the time of interviewing the energy master planner, it was not clear where the energy masterplan sat within the reorganised council and whether the project would proceed to procurement (Jones 2013).

7.2.2.6 Incremental development

The majority of the case studies investigated developed the alternative mode of operation in an incremental manner. This helped to reduce the financial risk of the projects and encouraged learning. For example, Woking Borough Council started with small-scale energy efficiency projects within its own properties which were used to create capital to develop Thameswey Ltd.; an ESCo. The profits of this ESCo were used to support further energy efficiency investment and the scheme was expanded to include a wider range of sustainability issues, including water resources and climate change adaptation. In this way *“the financial consequences for the Council’s budget in progressing the sustainability agenda are therefore minimised”* (Thorp, 2011: 82). In a similar way, the Energy Centre at the Olympic Park was "*designed to be as modular as we can...so what we have actually done is…enabled for the legacy", "there are two engines installed with space for a third and again there is two or three boilers shown…with space for two more. So basically we can grow the plant in line with legacy"* (Carr 2013). This allowed the business case to be developed based on the initial technology set-up but enabled it to expand as new opportunities for connection arose.

This incremental development allows interactive learning, reducing the initial social capital necessary to set up complex contractual and technical projects. This is also more likely to support development of locally embedded expertise (Hawkey et al. 2013). However, incremental development requires that governance arrangements evolve as the project does. For example the Byker Community Trust developed a series of legal documents to set out how its relationship with Newcastle City Council as the process of handover of the district heating networks evolved (Beadle 2014). Initially, the documents set out the terms under which the Council will supply heat to the development. A development agreement sets out the work that both the Council and the Trust will do to upgrade the system prior to transfer. A development agreement set out the terms under which the transfer would happen. Finally, following transfer, a non-residential supply agreement will be established to determine the terms under which the Trust will supply Council properties on the estate.

7.2.2.7 Alternative financial models

Despite having broader goals than delivering profit, alternative modes of infrastructure operation still need to match investment with revenue generation in order to survive and proliferate. Many of the cases achieved this through the use of innovative financial strategies, which often bypassed mainstream accounting practices. Examples include the fund recycling scheme developed by Woking Borough Council to retain savings from energy efficiency to pay for new supply capacity, which reduced the need for external financing and overcame the significant barrier to local authority lending (Thorp 2011). Private sector organisations, such as Kimberly Clark and Water Companies, have 'bundled' together investments which have long payback periods with those that pay back in substantially shorter period (Water Company Manager 2013; Wyatt 2013). This means that the overall 'investment bundle' meets the requirements of financial planning processes or regulators: "*what I did, and this is pretty unique to the industry actually, was devise first of all a 25 year strategy for investment and innovation in water and in [the Water Company]. And then also to devise that innovation in two bits. The first was aiming to deliver short-term quick-win benefits that provided business return to keep the regulator happy. The regulator doesn't want to invest in something like* *innovation if the return will actually barely struggle to be there. And then the other part of it was to create a longer term innovation programme or strategy that started to do the more fundamental type of work and the more fundamental research that will ultimately feed through into business benefits"* (Water Company Manager 2013).

A similar approach is taken at Newcastle City Council, who hope to bundle district heating schemes which make significant contributions to fuel poverty (but will not generate significant revenue) together with those that are commercially very attractive to ensure that they become more financially viable (Jones 2013).

In some circumstances, these alternative financial models required actors to create a new entity or partnership to circumvent mainstream processes. For example; Woking Borough Council created a new joint venture business *"to allow commercial projects to be run both inside and outside the Borough, to give enough autonomy from the election cycle to respond effectively to opportunities, to offer the capacity of autonomous budget management over the longer term, to harness external investment and to be profitable while remaining responsive to the Council's political objectives"* (Thorp 2011: 82).

A number of alternative modes of operation generated revenue from sources other than selling units of energy and water. This included service-oriented operation, such as Eco-Island which planned to charge a fixed price for energy supply and demand management measures, and Byker Community Trust, which had graduated charges, based on the floor area heated. This enabled a focus on system-wide efficiency, rather than marginal cost efficiency.

7.3 Summary

The alternative modes of operation examined in this thesis displayed markedly different characteristics from the mainstream mode of operation, which could increase the potential for unfavourable selection pressures from policy and regulation. Perhaps the most significant difference is the more diverse, and less profit-oriented, motivations for engaging in infrastructure operation. These alternative motivations strongly influenced the configuration and organisation of alternative modes of operation to maximise non-monetary value, such as resilience and self-sufficiency. However, capturing and reporting on these non-monetary values can be extremely challenging.

The presence of multiple motivations can result is a more integrated, outcome focussed approach that engages a wider range of actors from across the infrastructure system. This engagement is essential to many alternative modes of operation who work at a much smaller scale than mainstream operators. This smaller scale not only affects the type and efficiency of technology but also the capabilities available to manage this technology, customers and interact with regulatory systems. It can also have significant implications for a project’s access to finance. Perhaps in response to the limitations of scale many alternative modes of operation form partnerships to increase capabilities and balance issues such as attitude to risk with trust from consumers.

Many of the alternative modes examined had a long-term vision, which was delivered incrementally. This meant that measures which did not offer sufficient returns in the short-term could be justified in terms of their benefits in the longer term. This was supported by alternative financial models which were essential to justify investment in projects which contributed to the long-term goal of the project, or delivered non-monetary value. It also helped to address the issue of scale and access to finance; if an individual project is shown to be part of a wider package of work it is more likely to attract interest.

8 Policy and regulatory constraints

This chapter responds to research question three and describes the nature of constraints that alternative modes of infrastructure operation face from current policy and regulation. It begins with an analysis of the characteristics of the current system of policy and regulation of infrastructure operation. This serves two purposes; to provide context for later analysis of the specific constraints faced by cases examined in this thesis; and to provide an indication of the policy paradigms which underpin specific policy and regulatory processes.

Subsequent sections explore the constraints presented by policy and regulation to; scaling out alternative modes of operation (increasing the number of schemes of a similar scale); scaling up modes of operation (increasing the number of customers served or scope of service offered); and translation of the mode of operation to another sector. The constraints are represented as governance processes within the regime that have the potential to restrict the evolution of these modes of operation. This may be either in the form of unfavourable selection pressures, or the absence of appropriate support to enable alternative modes of operation to respond to these pressures. Selection pressures are processes which exert pressure for change, for example competition, regulation and cultural attitudes (Smith et al. 2005).

One of the key contributors to the success of the cases examined in this thesis was the avoidance of economic regulation, as result of the small scale of the project. Therefore, these sections include specific detail on the constraints presented when alternative modes of operation become subject to economic regulation (as is likely if they scale out or scale up). It should be noted that these findings are drawn from analysis of alternative modes of infrastructure operation and will not represent all constraints to low-carbon transition from policy and regulation.

The chapter concludes with a discussion of the role of the policy paradigm (the landscape level in the analytical framework) in creating constraints. This analysis will also inform the discussion of the potential for the paradigm to restrict policy responses to constraints in chapter 9.

8.1 Governance of energy and water infrastructure in the UK

This section analyses the characteristics of the current system of policy and regulation of infrastructure operation. In the UK, infrastructure governance has been dominated by national government action; which is briefly described below to characterise the actors and paradigm driving infrastructure governance.

8.1.1 National government and infrastructure governance

Prior to the second world war, UK water and energy infrastructure was decentralised and predominantly under the management of local councils (Marshall 2010). The war was followed by a period of nationalisation and centralisation in both the gas and electricity systems (National Grid 2005; Sherry 1984) and the later regionalisation of the water sector in the early 1970s. There was a high degree of government investment in network infrastructure and water supply to support this centralisation; primarily motivated by energy and water security issues (Marshall 2010).

In the 1980s central government transferred its state-owned water and energy assets into private hands (Hall et al. 2012). This privatisation was driven by the strong ideology of then the Conservative government, that state control of infrastructure was inefficient and undesirable (Roelich et al. 2014a). Privatisation was accompanied in some parts of the energy system by liberalisation and competition was introduced into the generation and supply of energy. It was considered that a ‘market’ context would be best suited to establish priorities for infrastructure investment and operation; however it was recognised that some regulation would be required to disband monopolies and create a ‘free’ infrastructure market[24] (Hall et al. 2012). Privatisation was accompanied by the establishment of independent economic and environmental regulators. This market-led ideology has been variously described as the Regulatory State Paradigm (Mitchell 2010) and the Pro-Market Policy Paradigm (Kern et al. 2014). The primary purpose of post-privatisation infrastructure government intervention was to introduce competition into the infrastructure system, to deliver greater economic efficiency and to protect consumer rights (Mitchell 2010). More recently, the focus of government intervention has shifted to include sustainability and resilience (Kern et al. 2014). Furthermore, until recently government intervention was led by economic (and to a lesser extend environmental) regulators; however, recently ministerial departments have taken a stronger role in directing the energy and water operators to ensure sustainability and resilience objectives are met.

The primary drivers and responsibilities of governmental institutions relating to energy and water infrastructure are briefly described below to provide an indication of the aims and character of this intervention. The links between ministerial departments, regulators and infrastructure sectors is represented in figure 8.1 and described in more detail below.

---

[24] Note the water market is not fully liberalised; retail competition has only been introduced for commercial organisations consuming more than five mega litres of water annually.

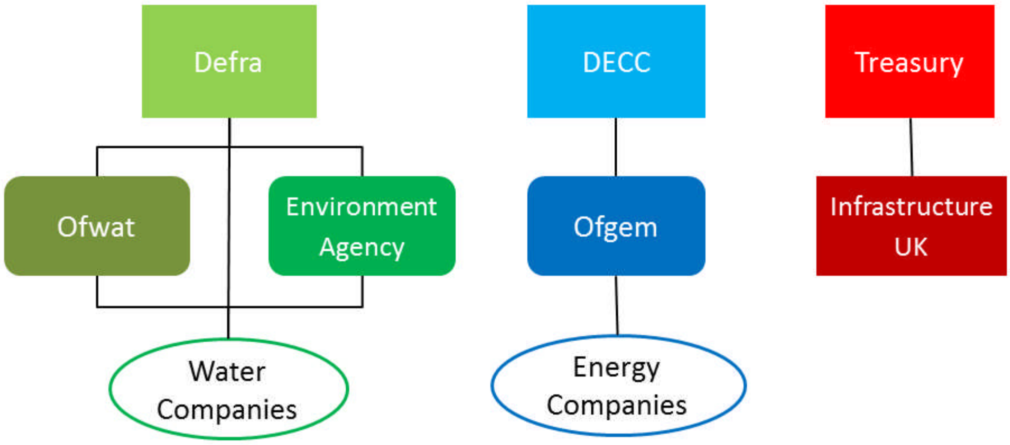


Figure 8.1: Ministerial departments and regulators directly involved in infrastructure operation

8.1.1.1 Infrastructure UK

Infrastructure UK is a unit within the Treasury which is responsible for improving the planning and financing of infrastructure projects. It sets out its plans in the annual National Infrastructure Plan (Infrastructure UK 2010; Infrastructure UK 2011). One of its principal priorities is to improve the co-ordination and planning of infrastructure projects and to develop a more integrated and long term plan for maintenance and improvement of all infrastructure sectors (Infrastructure UK 2011). It does this by identifying pipelines of planned investments to provide greater transparency and certainty. Furthermore, projects are prioritised to identify and promote those that are most crucial for economic growth. Infrastructure UK is also tasked with tackling the planning and regulatory delays that plague infrastructure projects.

Infrastructure UK has been set challenging targets to secure investment and mobilise finance for infrastructure projects and is particularly concerned with co-ordinating public and private investment. One of its principal mechanisms to secure investment is through attracting new sources of private investment to the UK and to reduce the risk of investment through provision of guarantees to support risks. However, it is also concerned with identifying and enabling new sources of revenue, like tolls on roads. There is an increasing interest in interdependencies and in particular the opportunities and Infrastructure UK is one of the few central government institutions which is able to take a cross-sectoral view of UK’s infrastructure as a system.

8.1.1.2 Department of Energy and Climate Change

DECC was created relatively recently, in 2008, primarily in response to the Climate Change Act (HM Government 2008a) and was formed to deliver on responsibilities enshrined in the Act, including ensuring the UK has secure, clean, affordable energy and promoting international action to mitigate climate change. It is responsible for setting the remit of the economic regulator for energy, Ofgem, and reviewing its performance. DECC clearly states that it aims to do policy in a way that maximises benefits to economy in terms of jobs, growth and investment (HM Government 2014b).

DECC develops policy to encourage investment in energy infrastructure and maintain reliable energy networks to improve security. It is responsible for the UK's target to source at least 15% of its energy from renewable sources by 2020 and has developed a renewable energy roadmap in parallel to a system of incentives and obligations for renewable technology deployment (DECC 2013c). DECC's role in ensuring affordability principally involves promoting energy efficiency in homes and non-domestic properties.

DECC leads the UK's contribution to climate mitigation through international action and cutting UK GHG emissions. The 2008 Climate Change Act established the world's first legally binding climate change targets of reducing the UK's GHG emissions by at least 80% by 2050 (HM Government 2008a). A series of budgets have been set to limit the amount of GHG the UK is allowed to emit over a specified time and the actions required to meet these budgets are set out in the Carbon Plan (HM Government 2011a). DECC also leads the UKs efforts to negotiate a comprehensive global climate change agreement. It administers support through the International Climate Fund and Reducing Emissions from Deforestation and Forest Degradation (REDD) to help developing countries to mitigate and adapt to climate change and reduce GHG emission from deforestation and forest degradation (HM Government 2014b).

8.1.1.3 Department for the Environment, Food and Rural Affairs

Defra has a wide remit to grow the rural economy, improve the environment and safeguard plant and animal health. Its responsibilities relevant to infrastructure are principally related to the water sector. This includes water resource management and water efficiency (HM Government 2014a). Defra sets the remit of the economic regulator for water, Ofwat, and reviews its performance.

Defra places particular emphasis on managing the water abstraction system to protect the natural environment and secure water supplies and has recently reviewed its approach to water resource management (HM Government 2011b). The review recognised that the current approach will need to become more interventionist and co-ordinated if these objectives are to be achieved in the face of demographic and climate change. Despite this, Defra still states that it will only intervene when there is a true market failure (Defra 2013).

Water efficiency is an important way to reduce demand and the associated pressure on the natural environment. Defra promotes efficiency in both the domestic and non-domestic sectors (HM Government 2014a). However, water efficiency doesn't have same link to affordability as energy because many domestic users are unmetered. Affordability is ensured through price controls regulated by Ofwat.

8.1.1.4 Economic regulators

The UK Government considers that "*competitive markets are the best way in the long run to deliver [infrastructure] services to consumers and provide incentives to invest and improve efficiency and service quality"* (BIS 2011:1). The Government recognises that network effects or economies of scale create natural monopolies, which limit effective competition. Therefore is has created independent economic regulators to promote competition, or to provide a proxy for competition where this is not possible[25], and protect consumers' interests (BIS 2011). This is typically thorough capping the prices that monopoly companies can charge in order to promote efficiency and fairness.

It has set out the principles of economic regulation (BIS 2011), which include;

- Accountability – the framework of duties and policies of regulators are set by a democratically elected government and there is a clear division of responsibilities to ensure legitimacy and transparency.
- Focus – regulation should protect end users by ensuring the operation of well-functioning and contestable markets or to replicate the outcomes of competitive markets.
- Predictability – regulation should provide a stable environment within which industry can make long-term decisions with confidence.
- Coherence – regulation should form a logical part of government's broader policy context and enable cross-sector delivery of policy goals where appropriate.
- Adaptability – regulation should have the capacity to evolve to respond to changing circumstances.
- Efficiency – intervention should be proportionate and cost effective.

Ofwat (the Water Services Regulation Authority, but formerly the Office of Water Services) was brought into power by the Water Act 1989 (HM Government 1989). Ofwat regulates

---

[25] For instance in the national energy network or in water supply companies.

monopoly water companies to ensure that customers are provided with a high quality service while bills are kept as low as possible. It does this through a five-yearly periodic review of both investment and pricing. Ofwat's secondary duties include promoting economy and efficiency and contributing to the achievement of sustainable development.

Gas and electricity markets were originally regulated separately by the Office of Gas Supply (Ofgas - brought into force by the Gas Act 1986) and the Office of Electricity Regulation (OFFFER – brought into force by the Electricity Act 1989). Ofgem (the Office of Gas and Electricity Markets) was created when the two regulators were merged (enforced by the Utilities Act 2000 (HM Government 2000)) Ofgem's principle objective is to protect the interest of present and future customers, which it enacts predominantly through revenue controls and supervision of wholesale and retail energy markets. It sets price controls on monopoly transmission and distribution networks and promotes competition (by enabling and encouraging customer switching) in non-monopoly suppliers. It has a duty to promote sustainability but when doing this its principal objective is to look after the interests of the customer.

8.1.1.5 Environmental and quality regulators

The Environment Agency is a non-departmental public body (NDPB) responsible to Defra and tasked with protecting and improving the environment and promoting sustainable development (HM Government 2014c). It delivers environmental priorities of central government through a series of permitting and licencing regimes. The environmental discharges from combustion plants are regulated through the Environmental Permitting regime with the aim of preventing pollution of air, water and land (Environment Agency 2013b). Discharges from water or wastewater treatment plants are regulated separately but with the same aims. The Environment Agency requires Water Companies to produce and maintain water resource management plans to demonstrate how a balance between supply and demand will be achieved (Environment Agency 2013a). Furthermore they are formally consulted on Ofwat's periodic review of pricing and investment.

The Drinking Water Inspectorate is also an NDPB appointed by Defra to ensure that drinking water is safe and acceptable to customers (DWI n.d.). It enforces drinking water standards set out in EU law and is also a consultee on Ofwat's periodic review of pricing and investment.

8.1.2 The influence of the European Union

The EU has been instrumental in bringing environmental protection into infrastructure policy and regulation in the UK, particularly around emissions reductions and water quality. The creation of the EU climate and energy package, and the associated 20-20-20 targets (European Commission 2008a), in 2007 arguably motivated the Climate Change Act and the formation of DECC. A series of directives have ensued, which formalise specific aspects of this package including renewable energy (European Commission 2009) and energy efficiency (European Commission 2012a). Both of these directives recognise that current policy and regulation has the potential to constrain alternative modes of infrastructure operation. The Energy Efficiency Directive, in particular argues that member states should *"identify and remove regulatory and non-regulatory barriers to the use of energy performance contracting and other third-party financing arrangements for energy savings"* (European Commission, 2012b: 8). Despite the UK leading the field in setting emissions reductions targets it has yet to achieve the levels of support for alternative modes of infrastructure operation provided in the EU and in other member states (Marino et al. 2010).

The EU influence over water infrastructure is perhaps even stronger and has a longer history. The Urban Wastewater Directive (European Commission 1991), Drinking Water Directive (European Commission 1998) and Bathing Water Directive (European Commission 2006) have steadily increased the standards of treatment of water and wastewater. This has driven the very high level of investment in particularly wastewater treatment (Hall et al. 2012). These very tight controls have strongly influenced the UK's regulatory system and the water supply licencing regime. The Water Framework Directive was brought into force in 2000 and has increased policy attention towards integrated protection, improvement and sustainable use of water through binding Environmental Quality Standards (European Commission 2000).

It is important to recognise the influence of European policy and that member states are required to transpose EU directives into national policy and regulation. Furthermore, although these directives have been instrumental in improving environmental standard there are occasions when policy objectives conflict. For example, the additional treatment processes required to meet increasingly stringent water quality standards are highly energy intensive and lead to trade-offs between targets to reduce energy demand and GHG emissions.

8.1.3 Overarching characteristics of energy and water infrastructure governance in the UK

Governance of regulated infrastructure systems, such as energy and water, displays a number of common characteristics relevant to this thesis, which are briefly described below and represented in figure 8.2.

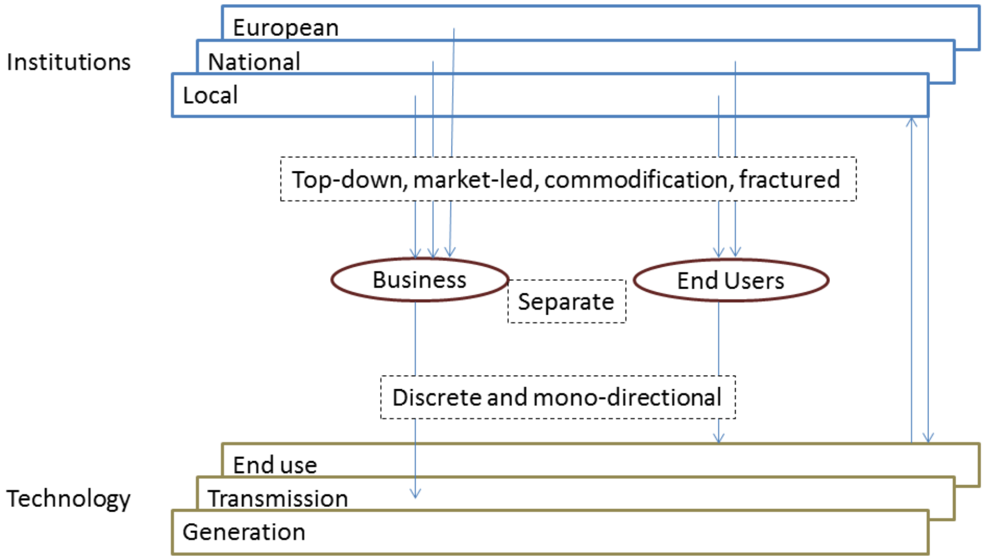


Figure 8.2: Representation of the current infrastructure governance system. Actors and assets are represented as boxes. The relationships between actors and assets are represented as directional arrows. Particular features of interest are highlighted in dashed boxes.

Despite being a liberalised system national infrastructure governance, the regulation of investment and pricing in particular, represents a ’top-down’ approach to governance: *“[Ofwat] controls 100% what a water company does…Ofwat is key or king rather and we respond to what the regulator wants us to do”* (Water Company Manager 2013). This approach has evolved to complement the rather monolithic and highly centralised mode of infrastructure operation and the interests of larger, powerful companies (Mitchell & Woodman 2010). This has implications for smaller and alternative modes of operation; top-down is not necessarily bad, if it enables diversity. However, the current system tends to maintain the current configuration of actors within the system, thereby preventing access by new entrants (Mitchell 2010).

Economic regulation is built on the premise that infrastructure should be part of a free and competitive market (BIS 2011). Energy and water are treated as commodities and the market is incentivised to deliver a unit of commodity (water or electricity) as cheaply as possible. This perspective is perhaps not surprising in the energy sector, which has evolved in the main from commodity markets, such as oil, gas and coal (Patterson 2008). However, this treatment of infrastructure products as commodities encourages business models which generate revenue through increased sales of commodity units and takes for granted the technology and physical assets necessary to produce, distribute and use water and energy. Furthermore, it focusses attention of units of water and energy, rather than the service (such as thermal comfort or cleanliness) that end users actually want (Patterson 2008; Roelich et al. 2014a).

In the energy sector the commodity market for electricity and gas actively hinders business models centred on renewable technologies and energy efficiency, which require all investment in the early stages of operation (Patterson 2008). The market creates short-term volatility which exposes these business models to significant risk and uncertainty, which deters incumbents and new entrants alike.

The full commodification of water has proved elusive to both government and economic regulator as a result of challenges in developing cost-reflective pricing mechanisms and a lack of direct competition[26] (Bakker 2005). Nevertheless, controls on price and investment programmes have been introduced to simulate the effects of competition and minimise the cost of each unit of water in a similar manner to commodity markets in energy. This price cap regulation was deemed to be "*preferable to rate of return regulation of profits [used in the United States] because it is simpler, less expensive and interventionary [sic] and less vulnerable to 'cost plus' disincentive effects*" (Littlechild 1986 quoted in Bakker 2005).

Competition among water companies occurs relative to the efficiency yardsticks calculated by the regulator. The valuation of water as a commodity was introduced to enable marginal water pricing[27] which was perceived to encourage optimisation of water use (Bakker 2005). However, cost reflexive pricing of water is inherently difficult because the externalities of water supply are very difficult to measure (Willis et al. 2002). Furthermore, water resource investment is incremental and requires high capital expenditure relative to revenue. This has severely limited the effectiveness of market-based regulation of water and has required an increasingly interventionist role for the state (Bakker 2005). Despite this, the market-led philosophy persists in economic regulation in the water sector.

---

[26] Which is prevented by the absence of transregional water networks

[27] The cost resulting from a unit increase in production at the margin. Where marginal costs and benefits are equal, according to neo-classical theory, efficiency is maximised. For example in an area where water supply is scarce any new demand will require new resources which are very expensive to exploit, therefore the marginal cost for new demand is very high and customers are encouraged to conserve water.

The pro-market stance permeates infrastructure governance beyond economic regulation; the majority of ministerial departments underpin their policy with assumptions that intervention should only occur in the event of a clear market failure and that market mechanisms are the most effective way to deliver policy outcomes (for example, Defra 2013). However, many academics argue that (technology blind) market mechanisms alone cannot achieve the government's sustainability and resilience aims (Mitchell & Woodman 2010; Gross et al. 2012).

Privatisation and liberalisation of utility sectors has led to the introduction of new actors responsible for the delivery, regulation and financing of utilities to correct market failures and protect customers. This has created a complex and fragmented governance structure (CST 2009a). The profusion of actors can result in duplication of regulation or conflict between regulators (Hall et al. 2012). For example the long-term Water Resources Management Plans required by the Environment Agency are not reviewed by Ofwat who instead had a separate requirement for supply and demand balance planning (Defra 2011a). Further complexity is added by the range of scales at which these issues are governed. Regulators can act at local, national or European level, often showing little consistency between scales.

Infrastructure systems are highly interconnected both through their physical layout and through their interaction with end users (Rinaldi et al. 2001; Roelich et al. 2014a). The governance of energy and water infrastructure systems has not evolved uniformly and rarely takes such interconnectedness into account (Hall et al. 2012). As a result synergies and interdependencies are largely ignored and policy and regulation tends to operate in sector-specific silos.

Utility companies, technology providers and end users are governed separately and their agency in the operation and transition of the infrastructure system is confined to limited parts of the system. These actors are conceived to interact with the physical infrastructure system (or technologies) at discrete and different levels. End-users interaction with technology is limited to end-use technologies, and their behaviour is seen to be independent of this technology. Businesses may interact with either generation or transmission technologies, which are regulated separately.

There is a great deal of agreement in the academic literature that the current system of infrastructure policy and regulation is poorly aligned with the goals of decarbonisation and energy security (Hall et al. 2012; Mitchell 2010; CST 2009b).

8.2 Scaling out case studies; becoming more numerous

This section describes the constraints from policy and regulation to increasing the number of modes of operations working at a similar scale to those examined in this thesis. The constraints are represented as stakeholders or processes within the regime that have the potential to restrict the evolution of these modes of operation either in the form of unfavourable selection pressures, or in the absence of appropriate support to enable alternative modes of operation to respond to these pressures.

8.2.1 Lack of Novelty

Novelty was central to a number of the cases and was a principal driver for the main actor or for key stakeholders. Engaging in novel projects can confer a number of benefits, including being seen as a first mover, or experimentation in a new market. This created a strong motivation for investing and co-ordinating resource in overcoming unfavourable selection pressures. This meant that actors place a higher value on non-monetary benefits, such as reputation or experimentation. In the absence of novelty, when these alternative modes of operation are repeated in a similar way in other locations, these motivations are unlikely to be present. If the projects are replicated the rather intuitive approach to benefit analysis is likely to be replaced by a more cautious and monetary approach to cost benefit analysis in commercial organisations; *“I would say from IBM’s perspective, there is no interest [in replicating the project] until we can prove how to monetise some of the things that happened in the Eco-Island experience”* (Stanford Clark 2013b). Replication of projects is likely to be more straightforward as a result of learning and resultant reductions in costs, but it is possible that the lack of novelty could result in poor co-ordination of resources and limited engagement of key stakeholders.

The majority of case studies analysed were financially viable in the long-run but required partners to manage a higher level of financial risk in the interim. For example, at Welborne, the initial fixed costs were very high, as a result of the high sustainability standards of the development, but which provided “*scope for better return in the long-term*” (Bench et al 2013). However the novelty of the site and its status as a pilot for a new approach to development enabled investors to overlook these risks and the reduced income in the short-term. In some cases, this meant that commercial organisations allowed longer pay-back periods or that organisations unaccustomed to managing risk, such as community organisations, were required to make extreme sacrifices while new modes of operation became established.

Actors were willing to accept this additional risk because they were experimenting. Moving from experimentation to replication requires that more actors are willing to take on this risk without the benefits accrued as a result of novelty, which is a significant constraint. For example; many local authorities are very risk averse and would not take on the risks that Woking did and Newcastle plan to (Hawkey et al. 2013). Furthermore, even actors that take on this risk during experimentation wouldn't do it again once the initial idea has been trialled; *"Yes. [Once you start replicating a business model] the expectation I think will be a faster go-to-market"* (Stanford Clark 2013b). A lack of policy and regulatory support to reduce risk, particularly financial risk in the early stages of projects, means that lack of novelty is a significant constraint to scaling out alternative modes of operation.

8.2.2 Lack of recognition of non-monetary benefits

Many of the cases analysed generated a wide range of benefits which are not included in a standard cost benefit analysis. Some of these are related to the novelty of the mode of operation, as discussed above. Other benefits are more related to the specific mode of operation, and particularly how it was developed and delivered. For example, the Chale Green project was delivered in a way that engaged residents in the process and the outcomes. This may have increased the resource required to deliver the project but resulted in benefits, such as crime and vandalism reduction and improved engagement with energy efficiency measures (Stanford Clark 2013a), which would not have been achieved by a less resource intensive and more cost efficient approach. Furthermore, many case studies were designed to deliver value in the long-term and were willing to receive a lower rate of short-term returns to enable this longer-term plan.

The value of these non-monetary and long-term benefits is not captured and included in assessments of efficiency of infrastructure (which pervade policy and regulation). Alternative modes of operation must compete with mainstream modes which have been optimised to maximise short-term economic value. This could mean that projects are either abandoned or that they are delivered in a different way to maximise revenue, rather than wider benefits. The inclusion of a wider range of objectives (for example, of community engagement at Chale, fuel poverty reduction at Woking and Newcastle and resilience at Byker) made the schemes viable because they achieved the multiple objectives of the project initiators. However, it is very difficult to justify these approaches in a fragmented policy environment which deals with community, the environment and climate change in different central and local government departments.

This fragmentation is reinforced by targets linked to government funding, which specify singular outcomes, such as carbon emissions reductions (Hawkey et al. 2013). For example, the assessment criteria for support from the Heat Network Delivery Unit (HNDU) does not include measures, such as fuel poverty, which can be central to the aims of local authority development of heat networks (Jones 2013). The lack of recognition of non-monetary benefits and focus on funding for individual aims acts as a significant constraint for alternative modes of operation. Those that might generate lower levels of financial revenue but far greater levels of non-monetary benefits will be out-competed in a system set up to promote the marginal cost efficiency of producing a unit of energy or water.

8.2.3 Institutional lock-in

The majority of cases were instigated by actors that are not conventionally associated with infrastructure provision. This is a promising development to increase the diversity and innovation of infrastructure operation. However; there may be a limit to the number of organisations such as local authorities or communities that are willing to do this as a result of institutional lock-in.

At an individual or community level, actors are locked into a passive system of receiving infrastructure services from external organisations (Knoeri et al. 2014) and there is a limited history of self-organisation. Breaking this institutional lock-in can require an influential individual, often termed champions, catalysts or change agents (Fell et al. 2009; Taylor et al. 2012). However; it's not clear whether there are enough champions to catalyse change on the necessary scale, in the face of the overwhelming constraints described in this thesis (Darby 2014). It is possible that once one person has done it the lessons learned can be transferred and influential individuals need to be less visionary. However, this knowledge transfer requires a strong network of intermediaries and the lack of government support for this kind of network is a real constraint (Hargreaves, Hielscher, et al. 2013).

Local authorities have not had a role in energy and water governance, beyond spatial planning, since these infrastructure systems were merged and nationalised in the 1940s (HM Government 1947) and 1960s respectively (HM Government 1963). Furthermore, a cultural ethos of aversion to risk and revenue generation limits the willingness of local authorities to engage in infrastructure operation: "*it's like anything else within the local authority; they're very risk-averse. So once you can even get the local people engaging, they are probably going to say 'oh we haven't got the skills to move this along or certain… You know, we'd have to bring in external advice"* (Jones 2013). This cultural ethos has evolved in response to central government restrictions on the role of the local authority (Hawkey 2013, Bale et al. 2012). Some of these restrictions have been lifted recently, in particular selling surplus power from renewable energy to the National Grid was lifted in 2010; however, the institutional lock-in created by historic constraints limits many local authorities to traditional ways of operating and a risk averse ethos persists (Hawkey 2013). Changes in financing and accounting practices could be slow in the face of this lock-in, limiting the number of local authorities willing to get involved (Bale et al. 2012).

It is thought that learning from other projects could help this but the lack of consistent financing (which is one of the biggest challenges in local authorities) means that examples from other authorities can't be followed, reducing learning effects (Hawkey 2013). The development stages of projects take a great deal of resources, in terms of time, and in the face of reducing budgets for core activities it is often difficult to find or justify this resource. It's not just the quantity of the resource but a lack of internal technical knowledge, which leads to a lack of confidence in decision making processes (Bale et al. 2012).

The perceived lack of resource to respond to selection pressures, as a result of institutional lock-in can represent a significant constraint to unconventional actors engaging in infrastructure operation.

8.2.4 Transaction costs

The majority of cases examined in this thesis had formed some sort of partnership to overcome financial, risk or technical barriers. These partnerships were underpinned by detailed contracts describing the distribution of risks and benefits and were very specific to the nature of the partnership and the mode of operation. Organisations like the Byker Community Trust had access to specialist resource which helped them to put together a series of contractual documents to manage the relationship between them and Newcastle City Council during the handover of the district heating system (Beadle 2014). These contracts take a great deal of time and expertise to draft; resulting in high transaction costs and acting as a constraint to wider roll out of the alternative modes of operation (Sorrell 2007).

Attempts have been made to develop standard contract terms; such as the RE:FIT scheme in London. This scheme aims to support public sector bodies to enter into contracts with the private sector to retrofit public sector buildings (Mayor of London 2013). A standardised procurement process and contract have been developed to overcome these transaction costs. The roll-out of this scheme to other parts of the public sector is being explored and is encouraged by the EU Energy Efficiency Directive (European Commission 2012a). However, this ignores the domestic sector, where transaction costs are arguably higher. Furthermore, standardised contracts are restrictive in their content and format and could exclude some of the benefits of bespoke governance arrangements (such as recognition of specific non-monetary values) and the increase in resilience afforded by enabling a diversity in modes of operation. The scheme had been rejected as a potential delivery mechanism by Newcastle City Council, because it did not allow them to address commercial and domestic properties (Jones 2013).

Even when contracts have been produced the monitoring and verification of the terms of the contract can be prohibitively complex and time consuming for smaller schemes where there is a direct relationship between the provider and the end user[28] (Sorrel 2007). This is a particular challenge for organisations which sell infrastructure services which requires more information and communication technology involvement than just measuring input flows (Roelich et al. 2014a; Steinberger et al. 2009).

The significant transaction costs associated with developing contractual and monitoring arrangements can act as an unfavourable selection pressure to alternative modes of infrastructure operation and constrain the scaling out of these alternatives.

8.3 Scaling up case studies; becoming larger or widening scope

This section describes the constraints from policy and regulation to increasing the size or scope of alternative modes of operation. The constraints are represented as stakeholders or processes within the regime that have the potential to restrict the evolution of these modes of operation either in the form of unfavourable selection pressures, or in the form of the absence of appropriate support to enable alternative modes of operation to respond to these pressures.

8.3.1 Trust

Trust has played a significant role in the success of many of the cases examined, particularly those operating at a smaller scale and involving the community. Relationships in the alternative modes of operation examined in this thesis appear to be strongly founded on trust. The proposals for community governance of the Welborne development are founded on trust and respect for communal goods *"I think the best thing is just to trust people and if you do that*

---

[28] i.e those schemes that do not simply generate and sell energy to the grid but sell utilities or infrastructure services directly to customers, whether they generate the utility products themselves or not.

people generally speaking like to be trusted. And they are more respectful" (Bench et al 2013). This increases engagement and commitment of end users and particularly the promotion of collective benefits over individual. The project officer at Chale Green spent a great deal of time in the early stages of the project building trust in the motivations for and benefits of the retrofit project. The selection on community members to act as energy advisors contributed to increasing the level of trust in this advice and consequently the success on the retrofit measures (Stanford Clark 2013a).

This supports the findings of Walker et al (2010), who concluded that trust between people and trust in local institutions enables cooperation and ensures that projects are locally appropriate and promote collective benefits (Walker et al. 2010). One of the greatest challenges of increasing the scale or scope of projects is the retention of this trust.

Walker et al (2010) provide an excellent example of initial support for a small-scale wind power project which was promoted to generate local income and boost the declining farming industry. The local community's trust in these motivations was eroded when the project expanded and the nature and distribution of benefits was questioned. Trust between the local community and the (local) developer was eroded to such an extent that an active opposition group was formed to stop the expansion of the development.

Interactions at the regime level are not all founded on trust (Mitchell 2010). This could be because trust is far less prevalent in interactions between individuals and organisations, where the values are less well aligned and self-interest is more prevalent, which is the case with the current energy and water regime (Mumford & Gray 2010). When projects try to scale up, and interactions between alternative mode of operation and regime get more significant, the benefits of trust in the alternative mode could be eroded by unfavourable selection pressures.

Trust is particularly low in conventional utility companies, who are perceived to act entirely in their own self-interest and are presumed to oppose environmental initiatives (Mumford & Gray 2010). Acting alone and without significant reconfiguration of models to demonstrate better alignment of values and fairer distribution of benefits these organisations are unlikely to increase their role in alternative business models and scale up these models. One way to overcome this lack of trust would be to create partnerships with organisations such as community groups or local authorities who are more trusted. However, it is important to be aware that collusions between groups who are expected to challenge each other will be seen as a betrayal of trust (Mumford & Gray 2010).

In addition to promoting collective benefits, trust is an important attribute in the management of risk and uncertainty, which are a key part of alternative modes of operation (Bellaby 2010). Participants in the cases examined accepted a higher level of risk and uncertainty on the understanding that the benefits of doing this will be higher. Participants are required to trust that the project will succeed and that when the project scales up that the risk and uncertainty is reduced. This might not be the case if projects expand and face new constraints from regulation and planning; this trust might not persist.

Trust is a vital resource in addressing selection pressures and a lack of support to co-ordinate and enhance this resource could result in a constraint to the scaling up of alternative modes of infrastructure operation.

8.3.2 Capabilities

Expanding the scale or scope of projects may increase interaction with processes and selection pressures in the regime, such as regulatory processes. This increasing interaction may require an expansion of certain knowledge, skills, competence and attributes, which are often termed capabilities (Sen 1997). For example, taking on additional customers increases the variability in demand and also creates the necessity for standardisation of customer interaction. This may require additional technical knowledge and skills to better balance supply and demand as well as competence in contract development and management and customer service skills. Additional customers might also require an increase in the scale and complexity of technological systems, which requires more specific technical knowledge and skills. An increase in the scale of technology can tip a project over regulatory thresholds, after which operators would have to engage with regulatory processes such as supplier licencing. These processes are complex and require capabilities in financial and regulatory systems, which may be beyond the abilities of community or local authority-led schemes.

Expanding technologies and customer base can also bring greater risks, which many operators do not have the capabilities to manage. Furthermore, larger operations will need new approaches to internal management; large scale projects are unlikely to be managed by committees which are prevalent in community-based schemes, so formal governance systems of procedures and accountability are required (Walker et al. 2010).

An increasing exposure to regime processes and selection pressures can reveal a lack of capabilities necessary to respond to these processes and pressures, which can present significant constraints to scaling up alternative modes of infrastructure operation.

8.3.3 Regulation

The majority of cases existed because they did not fall within the remit of regulators, particularly economic regulators. While their contribution to infrastructure service delivery remains small regulators are happy for alternative modes of operation to fly under the radar; however, increasing in scale or scope could bring them within the regulated sector. Moreover, their increasing contribution to infrastructure service provision is likely to trigger a change in regulation to incorporate their activities.

8.3.3.1 Constraints from energy system regulation

As alternative modes of operation increase in scale they are more likely to need to interact with regulated systems at the regime level. In the energy system, this means transacting with national markets for energy and using or creating energy networks. Interacting with national markets has very high transaction costs as a result of the structure and functioning of the market and high risks associated with balancing mechanisms[29] (Mitchell 2010). This is related to both selling to and buying from the wholesale market. Some of the cases examined in this thesis, such as Byker Trust, do interact with the wholesale market to purchase gas but use long-term contracts to hedge risks (Beadle 2014). However, this risk reduction is likely to increase costs, putting the scheme at a disadvantage compared to those better able to manage this risk. Amendments to wholesale market have been proposed to reduce the challenges faced by small supply companies in securing access to buy from the wholesale market and competing with vertically integrated energy companies (who can secure longer-term agreements for supply) (Ofgem 2013b). However, there is no associated proposal to reduce barriers to small generators who want to sell to the wholesale market and arguably face even higher barriers associated with forecasting and balancing mechanisms.

It is possible for generators to avoid interacting directly with wholesale markets by using a consolidator, who pools the output of a number of generators; however, this means that operators do not receive the full retail price for electricity. Revenue can be reduced by up to a quarter (Bolton & Foxon 2013), which can significantly affect the financial viability of the scheme (Haringey Council 2012).

Another way to avoid transacting with the wholesale market is by selling to designated customers using a public distribution network. However, using the public distribution network incurs charges, which historically have not been cost reflective, particularly in regions where

---

[29] Generators who fail to input the same amount of power they had agreed to are penalised for the amount of power they failed to generate (Green 2010)

the network is at or near capacity. The additional network charges can make the difference between a scheme being financially viable or unsustainable (Bolton 2011).

As a result, many schemes have invested in private wire networks to connect generation directly to local customers, bypassing the electricity network and wholesale market, as demonstrated by Woking Borough Council (Thorp 2011). Private wire networks allow an operator to retain the full retail price for output but bring a degree of investment risk and require the operator to balance supply and demand locally. Furthermore, the scale of operation of private wire networks is limited; above 5MW of power – of which up to 2.5 MW may be to domestic customers – operators must hold a supply licence and manage the associated administrative and transaction costs, which may require complex IT systems, specialist staff and entail significant legal costs (Fontenergy 2009).

An important implication of this licencing requirement is that operators would no longer be able to enter into longer-term contracts with customers, which are prevented under Ofgem's regulation. If private wire operations scale up and are no longer exempt from licencing it would mean customers would be able to switch at 28 days' notice. Longer-term contacts are essential to justify investments in costly infrastructure, such as heat networks: *"If it would just have been a 15 year contract, because of the cost of recovery, it would be very expensive for anybody who wants to connect to the system. Compared to a much longer concession* " (Carr 2013). This means that revenue generation could become unacceptably uncertain and unstable for smaller modes of operation and they would be unable to compete with large suppliers, who can accommodate changes in their much larger customer base.

Ofgem has introduced a 'Licence Lite' for smaller scale electricity suppliers who exceed the exemption threshold but lack the capabilities to manage a full licence. Under these arrangements (implemented in 2009) Licence Lite holders engage with a third party licenced supplier (TPLS) who provides services relating to market access, in accordance with a Supplier Services Agreement (SSA) (Ofgem 2013c). However, TPLS are under no obligation to enter into an SSA and the costing of services may be unpredictable and uneconomic for the potential Licence Lite holder; Ofgem has deemed that these factors should be left to the market. To date only the Greater London Authority has applied for a 'Licence Lite' (Arup 2011; Greater London Authority 2013).

8.3.3.2 Constraints from water system regulation

The situation is more challenging still for alternative modes of operation in the water sector, where new entrants face far greater constraints. Upstream services (anything associated with water supply and wastewater carriage and treatment, except retail) can be delivered by parties other than the local incumbent water company; by various means of transfer of raw/treated water from neighbouring incumbents (or transfer/treatment of wastewater by neighbouring incumbents), supply by third parties (to non-domestic customers only) or by self-supply/self-treatment (Defra 2013). However, there is a number of significant barriers to new entrants, including pricing, abstraction licencing, operational expenditure and regulatory limitations which mean that supply is dominated by local vertically integrated companies (Cave 2009). There is no retail competition other than for companies who use more than five mega litres per year – the overwhelming majority of customers in the UK use their local statutory undertaker. This severely limits the potential for alternative modes of operation.

Prices and standards are determined by Ofwat using comparative competition regulation; savings from cost efficiency over and above the industry average can be retained by the over-performing company, but only if savings are produced over a period of between five and seven and a half years. There is concern that this constrains more innovative approaches to water services, which have more uncertainties associated with outputs and costs of development and implementation (CST 2009b). If the combined costs of introducing and operating the new technology pay back over a longer period than this the companies are very unlikely to introduce it.

Both water prices and infrastructure investment are set in a five-year cycle, which forces short-term, incremental change and fosters risk aversion in water companies. Moreover, the regulator's focus on cost efficiency per unit of water sold means that investment in demand management is disincentivised: "*…at the end of the day, especially with the march towards more universal metering, it is not exactly in their interest to promote water efficiency…They are selling water at the end of the day; the more they sell the more profits they make. It is a little bit of a catch 22 for them"* (Water Company Manager 2013). Reduction in water consumption does not count as cost efficiency and the water company actually loses revenue (since revenue cannot be generated from domestic customers for services, other than a direct charge for water).

Recent changes to legislation have allowed limited competition in the retail market for water (Defra 2011b). Water companies are able to offer, and charge for, a broader range of services to industrial, commercial and public sector clients, and to offer these services to customers outside their licenced catchment area. This has resulted in some innovations in modes of operation, including the provision of water management services: "*The business model will be around what else we can sell them. It will reduce their water consumption but it gives us a shoe in to the business and we are talking to the commercial people and from there might be a whole host of different business products that we can sell them. So there is an opportunity there"* (Water Company Manager 2013). The commercial services arm of Scottish Water has gone so far as to offer 'gainshare' services to an investment constrained local authority, where they invested in water saving technologies and retained some of the £1.3 million savings in water bills (Business Stream 2013). These additional revenue streams compensate for losses associated with reduced water consumption. However, these alternative modes of operation are limited to non-domestic users.

There is a further limitation to the development and testing of new business models in the water sector; applied research (which would include market research and new business model development) must be written off in the year of expenditure (CST 2009b). This discourages research into new modes of operation because it cannot be reasonably demonstrated to be cost beneficial. This vicious circle of constraints to investment, research and revenue generation disincentivises efforts to develop alternative modes of operation within the water companies.

8.3.4 Financing

Finance at an early stage of complex projects is essential to support planning and development. However, very few sources of funding will support this critical phase; the majority is available for technology, and for implementation of detailed plans. Many incentives are only available when projects are operational, such as Feed in Tariffs, which doesn't help manage the financial risks of initial stages. Lack of funding at a crucial stage of development prior to operation was cited as the cause for the failure of the Eco-Island project (Findon 2013b). The incremental development of projects further limits funding available, since the small scale of the project and the risk associated with new ideas make it unattractive to investors (Hildyard 2011). This was demonstrated at the Byker district heating scheme, where a biomass boiler scheme was presented as a potential project to the six energy companies obliged to fund community energy schemes under the CESP scheme. Despite being legally obliged to undertake such schemes only one energy company responded because the scale of the scheme was insufficient to be of interest (Beadle 2014).

The current system of funding and incentives is so complicated it is often perceived that it’s not worth the additional effort to seek this funding. Inconsistent funding is also an issue, as noted by Carley et al, making “*it difficult for industry actors, community stakeholders and other involved parties to plan [energy-based economic development] projects”* (Carley et al. 2011: 293).

Projects must generate revenue to support operation and enable future development. Projects which involved regulated sectors (such as water and energy utilities as well as local authorities, whose financial processes are regulated) face significant challenges in generating revenues thorough new services. Water companies in particular must account for revenue generated from non-core business separately to that from other activities, which is a significant barrier to initiating other activities. Demand management activities therefore reduce profits, which disincentivises action in this area.

The availability, scale and complexity of financing can act as an unfavourable selection pressure to alternative modes of operation as a result of their small scale, the importance of funding at early stages of projects and the alternative approaches to revenue generation.

8.4 Translating case studies to other sectors

The majority of alternative modes of operation analysed in this thesis were in the energy sector. This section explores the constraints to translating this practice to the water sector or to more integrated modes of operation, covering both water and energy.

8.4.1 Alternative modes of operation in the water sector

The challenges of innovation in the water industry have been discussed above and it seems very unlikely that water companies will be able to extend the innovation displayed in modes of operation with business customers to the domestic market as a result of the tight regulation aiming to protect customers. The co-evolution of regulation and business strategy in the water sector is very strong: *“Whilst innovation and [research and development] R&D is occurring throughout the UK water industry, and along the supply chain, it is largely driven by the regulatory framework and appears to be uneven – it is concentrated in the supply chain”*(CST 2009: 13)

“*The risk appetite … is much lower today than it was. Some of the more speculative endeavours funded historically and in use today in the water and other sectors would not be undertaken in today’s regulatory structure.”*(CST 2009: 14)

Therefore, a major change in regulation would be required before alternative modes of operation could ever challenge the mainstream mode of operation in the water sector.

8.4.2 Integrating infrastructure operation

Perhaps the more interesting question is: what might be constraining a greater integration in the operation of infrastructure sectors? Infrastructure systems are highly interconnected both through their physical layout and through their interaction with end users (Rinaldi et al. 2001; Roelich et al. 2014a). The physical interdependence is well illustrated by the interconnectedness of the water and energy systems: water and wastewater treatment plants are significant energy users, and becoming more energy intensive as water quality standards become increasingly stringent (CST 2009b). Water companies place a burden on the energy system but also provide a great deal of potential to generate energy, supporting the energy system. For example, through anaerobic digestion of sewage sludge and the use of hydro turbines. A similar example at the end-user side is hot water, which accounts for 5.5 per cent of household energy use (Defra 2008). A reduction in hot water use would not only contribute to reductions in water consumption but also to a reduction in energy consumption (Waterwise 2011).

Infrastructure governance systems have not evolved uniformly across utility streams and rarely take such interconnectedness into account. The governance arrangements have evolved in response to the changes within the individual utility systems and thus exhibit dramatic differences between sectors. For example, the tight regulation of the water sector, with the economic regulator playing an important role in setting prices and making investment decisions, is in stark contrast to the waste sector, where, investment and pricing of waste management services is left entirely to the market (Hall et al. 2012).

Governance continues to operate in sector-specific silos – synergies and interdependencies are largely ignored (Roelich et al 2014b). Schemes designed to reduce end-use of energy, such as building regulations, the Green Deal and the Energy Company Obligation don’t address the end use of water. Fragmentation of regulatory authorities is acting as a constraint to extension of these schemes to end-use of other utilities and the resulting resource savings.

This separation of regulation has driven decision processes that are locked into consideration of separate infrastructure operation. This is the case for both infrastructure providers and national and local infrastructure planning. This is exacerbated by poor information sharing between sectors[30] which can reduce co-operation and lead to market failures as a result of information asymmetries (Hall et al. 2012).

There is limited evidence of the risk and benefits of integration of infrastructure operation despite the Council for Science and Technology's (CST) conclusion that operating infrastructure systems in silos leads to "*financial and operational inefficiencies, a poorer service to citizens and businesses, and unintended negative consequences"* (CST 2009b). There is increasing interest from Infrastructure UK in the risk of increasing interconnectivity but also the opportunities that this presents. Work has been commissioned to quantify the potential savings from integration during the construction of infrastructure, which concluded that significant savings could be made from sharing facilities and unlocking new investment and growth (Frontier Economics 2012). However, there is no equivalent evidence for infrastructure operation.

The separate regulation of infrastructure sectors and lack of evidence demonstrating its benefits is creating a series of constraints to more integrated modes of infrastructure operation.

8.5 Constraints from Regulatory State Paradigm

The unfavourable selection pressures for characteristics of alternative modes of operation and the constraints on scaling out or scaling up can predominantly be traced back to the current pro-market policy paradigm. Paradigms affect the goals of policy (cost efficiency), the kind of instruments that are used to attain them (market-based) but also the very nature of the problems they are meant to be addressing (infrastructure as a free market) (Hall 1993). If the paradigm is not aligned with the challenge of a sustainable infrastructure system, it hinders the ability of policy and regulation to enable transition towards a more sustainable infrastructure system (Kern & Mitchell 2010).

The effect of the current paradigm is so strong that one of goals of energy policy is to promote competitive markets: "*Vigorous competition in energy stimulates innovation and ensures the efficient allocation of resources, improving service quality and driving down price*" (DTI 2003: 95). Energy policy thus relies on market-based, entrepreneurial approaches, rather than intervening, because of the greater faith in markets than in political processes (Foxon, Pearson, et al. 2005). This is mirrored in water policy which persists in using proxy competition to drive innovation; despite evidence exposing the ineffectiveness of this approach (CST 2009b).

---

[30] Which is partly due to data protection limitations on sharing customer details

Paradigms restrict and shape policy makers' perspectives of what is feasible, possible and desirable (Kern & Mitchell 2010). Mitchell describes this effect as a band of iron holding together the current framework – some movement is possible within the framework "*but, in the end, this framework constrains certain actions or policies*"(Mitchell 2010: 2). The band of iron of the current policy paradigm constrains alternative modes of operation in a number of ways.

Firstly the *pro-market focus* of the paradigm views markets and competition as the most effective way of meeting society's choices. This has resulted in policy that mimics markets as far as possible (new policy must pass the 'market test') and prevented targeted support for specific modes of operation necessary to increase diversity in infrastructure provision and deliver sustainability goals (Gross et al. 2012; Smith 2009). Furthermore, making choices based solely on price in competitive markets implies a bias against long-term decision making, which is essential in enabling alternative modes of infrastructure operation (Mitchell & Woodman 2010).

Secondly, the *narrow definition of value* in purely economic terms overlooks the non-monetary benefits that end-users receive from more efficient and inclusive infrastructure operation, such as reduction in fuel poverty and local employment. This is exacerbated by an approach to regulation where decisions are based on pre-known quantitative costs and outcomes (Mitchell and Woodman 2010). Therefore, choices favour economic, over environmental and social goals.

Finally, the sole focus on *switching and price control* as a means to deliver customer protection. Switching is encouraged to deliver competition, but can actually exclude new entrants to the market who rely on a stable customer base; this has the effect of actually reducing competition from outside the six biggest energy supply companies. It also reinforces the exclusion of non-monetary benefits of infrastructure described above.

The paradigm has shaped the institutions which govern infrastructure operation, which is dominated by *economic* regulators. The primary remit of these regulators is to protect customer (through marginal cost reductions of units of energy and water) and maximise competitions (BIS 2011). This has taken priority over (and often conflicts with) their secondary remit to ensure sustainability and look after needs of future customers (Mitchell & Woodman 2010).

8.6 Summary

The water and energy sectors were privatised in the 1980s and some parts of each system have been opened up to competition since that time. It was considered by the then Conservative government that the market context was best suited to establish priorities for investment and operation (Roelich et al 2014b). The primary purpose of post privatisation governance was to introduce competition to deliver economic efficiency and protect consumer rights (Mitchell 2010). This market-led ideology (often called the pro-market policy paradigm (Kern et al. 2014)) has persisted since this time, despite the increasing importance of sustainability objectives.

The pro-market policy paradigm, and the policies and regulation which stem from it, constrain alternative modes of infrastructure operation in a number of ways. The narrow definition of value in purely monetary terms overlooks the significant benefits that might arise from alternative modes of infrastructure operation, such as fuel poverty alleviation, GHG emissions mitigation and local economic growth. This can mean that alternative modes are outcompeted because their competitiveness is measures on production of financial value alone. Many of the market-based mechanisms used to regulate the incumbent providers, such as the wholesale market and supplier licencing, create very high transaction costs. These systems have been designed for large operators with economies of scale in functions like market forecasting and customer support (Mitchell and Woodman 2010).

Many actors wanting to engage in alternative modes of operation haven’t been involved in infrastructure operation for a great deal of time, for example municipalities. This can leave them locked-in to their current institutional role or lacking in essential capabilities to deal with technical, procurement or regulatory aspects of operation (Hawkey 2013, Bale et al. 2012). The risks presented by engagement in infrastructure are often unacceptably high to organisations not experienced in risk management. This is particularly the case with the financial risks, especially during the early stages of project development. One particular constraint is that support intended to encourage access to finance (such as CESP) and reduce regulatory burden (such as Licence Lite) are administered through incumbent providers, who, perhaps unsurprisingly, show little inclination to engage fully with these mechanisms.

These constraints can present unassailable barriers to alternative modes of infrastructure operation and limit their contribution to isolated examples. The next section of this thesis discusses how these constraints might be alleviated to enable a diversity of modes of operation.

9 Responses to policy and regulatory constraints

This chapter responds to research question four and explores responses that might mitigate the regulatory and policy constraints to alternative modes of operation described in chapter 8. It is structured in three parts; the first describes potential policy proposals that might contribute to overcoming constraints to alternative modes of infrastructure operation. It is recognised that these proposals respond specifically to the constraints to alternative infrastructure operation identified in chapter 8. As such, they do not represent a comprehensive assessment of policy necessary to enable a low-carbon transition. Furthermore, they are presented as initial proposals and the full implications of their implementation have not been assessed.

These policy responses presented in section 9.1 all represent incremental change from the current system of governance and would not contribute to the necessary break in the strong link between governance and the mainstream mode of operation. In this way these proposals address specific constraints but are unlikely to result in wholesale transformation of infrastructure operation. To do this we need to rethink the system of governance, to break the lock-in to mainstream infrastructure operation. Therefore, the second section of this chapter draws on theories of governance described in chapter 2 to develop a framework that presents elements of an alternative governance system that might be supportive of alternative modes of operation with the characteristics described in chapter 7. However, the success of both specific policy proposals and governance system change will be affected by the political ideology; therefore, the final section discusses the implications of the policy paradigm for the proposals and analysis in the first two sections.

9.1 Actions to remove constraints: policy responses

This section describes a series of proposals for policy makers to mitigate the constraints from policy and regulation described in chapter 8. They are designed to reflect the two processes described by Smith et al, which will instigate regime change to enable alternative modes of operation to co-exist with currently dominant utility model more effectively (Smith et al. 2005):

1. Shifting selection pressures bearing on the regime to be more favourable to alternative modes of operation.
2. The co-ordination of resources available inside and outside the regime to enable alternative modes of operation to respond and adapt to these pressures.

Selection pressures are processes which exert pressure for change, for example competition, regulation and cultural attitudes. Policy can encourage a shift in selection pressures to ensure they are more coherently aligned with the goals of carbon mitigation and resource efficiency. This could be by creating conditions so that organisations are competing to deliver social and environmental value as well as economic value; ensuring regulation is co-ordinated and clearly articulates the urgent need for mitigation; or creating policy that aims to alter social norms (for example the change in attitudes towards driving without a seatbelt have been driven by strong regulation in this area). This requires changes in regulation but also a change in the decision processes behind policy development to ensure that selection pressures are coherent. Responses that have been identified include; encouraging more integrated policy to ensure consistent, cohesive selection pressures; targeted support for alternative modes of infrastructure operation, not just market-based instruments which favour incumbents; alternative approaches to valuation to capture social and environmental value created by alternative modes of operation; and new approaches to regulation that are better aligned with low-carbon goals. These responses to shift selection pressures are discussed in more detail below.

Responding to selection pressures requires certain knowledge, skills, competence and attributes, which are often described as human capital when considered in relation to their potential to increase economic development (Keeley 2009). In common with other work on local innovations this thesis takes a broader definition of these resources, as capabilities, which Sen defines as the "*ability of humans to lead lives they have reason to value and enhance the substantive choice they have"* (Sen 1997). The value in this definition can be direct (enriching individuals' lives) or indirect (contributing to a project or wider production). Using capabilities allows for social change, not just economic change (Sen 1997). Capabilities can be enhanced in individuals and organisations or transferred and co-ordinated through networks (Hargreaves, Hielscher, et al. 2013). Policy responses to enhance or transfer capabilities are discussed in more detail below.

9.1.1 A more integrated approach

The tendency of the current, fragmented policy and regulatory system to focus on individual goals, infrastructures, project stages and organisations in isolation constrains projects which have a broader range of motivations. For example, district heating networks in the Newcastle City Council case study will create low-carbon heat but will also make a significant contribution to reducing fuel poverty. However, increasing the adoption of heat networks and reducing fuel poverty are dealt with by different policy groups in DECC and the Council's application for funding in support of these networks will not consider the fuel poverty reduction benefits of the scheme (Jones 2013). This could mean that it is seen to deliver few benefits than more commercially viable schemes and is not selected for funding. A more integrated approach to all aspects of infrastructure operation is necessary to enable alternative and more sustainable infrastructure operation.

This should start with problem framing in policy which recognises the complex system dynamics of infrastructure operation and respects the multitude of values and priorities of organisations and individuals that could be engaged (Leach et al. 2010). There is a need to 'open up' beyond outcomes of 'cost efficient carbon emissions reductions' to embrace a range of alternative outcomes (Stirling 2008). Alternative outcomes don't reject the importance of climate change mitigation but do draw attention to complementary and additional benefits of alternative approaches. Incentives and funding criteria need to be adapted to encourage the wider benefits of alternative modes of operation, beyond carbon saving, such as fuel poverty reduction.

Selection pressures are often conflicting and need to be made more coherent and better aligned with the goals of sustainable development. For example, Water Companies are required by Ofwat to demonstrate water efficiency measures but in parallel to this, Ofwat regulates the price of each unit of water. In this situation demand reduction will reduce water company profits, and the regulation of the price per unit of water is poorly aligned with the goal of demand management (Water Company Manager 2013). The government is required to steer more deliberately away from unsustainable pathways by directing incentives and interests towards more innovative and sustainable modes of operation (Leach et al. 2012). Greater harmonisation is required to increase the effectiveness of policy. There is a particular need for better integration of innovation, industrial and infrastructure policy (Foxon, Pearson, et al. 2005; Mitchell 2010).

Policy needs to recognise the interconnectedness of infrastructure systems and enable the exploitation of the cost and resource efficiencies associated with a close integration of the operation of infrastructure systems. Many of the cases examined in this thesis were attempting to engage with more than one infrastructure sector but were overwhelmed by the complexity and barriers they faced in trying to do this (Bench et al. 2013). To reduce this complexity it will be necessary to introduce formal systems to increase integration and cooperation between infrastructure regulators. There are institutions that could enable this co-operation, such as the Joint Regulatory Group which includes representatives from all sector-specific economic regulators in the UK (Hall et al. 2012). However, the group's remit does not include cross-sectoral issues or integration. Infrastructure UK has an Engineering and Interdependency expert group with a specific remit to consider economic opportunities from interdependencies at the development stage but could be extended to include operational integration and governance.

Long-term strategic goals are necessary to develop a stable and consistent policy framework and to create directionality for policy in the short-term (J.-P. Voß et al. 2009; Leach et al. 2012). It could be argued that the Climate Change Act provided this strategic direction; however, targets remain limited to energy generation and associated targets for demand management have thus-far been resisted (perhaps as a result of fears that this would reduce economic growth). Inclusion of water and energy demand reduction targets would support long-term investment in modes of operation specifically designed to reduce resource consumption, such as energy performance contracts.

9.1.2 Targeted support

Proponents of the pro-market paradigm argue that infrastructure transition should be left to the market and that market-based instruments such as carbon pricing will drive adoption of efficient business models (Helm 2009). It is suggested that targeted policy distorts the function of carbon markets, ultimately hindering efforts at decarbonising the economy (Less 2012; Moselle & Moore 2011). However, neutral policy favours large, incumbent operators and is less supportive of alternative modes of operation and smaller operators, which have the potential to undertake more innovative activities. It could be argued that market-based instruments could slow innovation because incumbent companies are likely to concentrate on doing the same things more cheaply, in order to stay competitive, rather than making radical innovations in modes of operation (Mitchell & Woodman 2010). Furthermore, market-based mechanisms, such as carbon prices still include significant uncertainties which can make future conditions unpredictable and discourage investment (Gross et al. 2012).

There is an increasing body of evidence supporting a more targeted approach to policy that can reduce the risk and uncertainty associated with low-carbon investment and drive innovation (Gross et al. 2012). This is particularly important in the absence of a global carbon price, which undermines the effectiveness of market-based but local carbon trading schemes, such as the European Union Emissions Trading Scheme. These findings relate to support for renewable technologies, but is equally relevant to alternative modes of operation. There is an increase in targeted support for alternative modes of infrastructure operation, such as the strong support for energy performance contracts and ESCos in Europe. This support is increasing the prevalence and effectiveness of these modes of operation over and above those countries which do not offer such support (such as the UK) (Marino et al. 2010). Targeted support for alternative modes of operation is enshrined in the EU Energy Efficiency Directive and the UK must improve its implementation of these aspects of the directive (European Commission 2012a).

It had been argued that targeted policies are more suitable for risky projects, like Feed in Tariffs (FiTs) for renewable energy, than market-based instruments, like cap and trade (Gross et al. 2012). It is possible to extend FiTs to energy or water savings (Eyre 2013) or infrastructure services delivered by local authorities or communities, which could provide a more favourable selection environment for more risky modes of infrastructure operation until measures to improve capabilities have delivered results.

Support is particularly important to reduce financial risk and uncertainty in scaling up from small scale experiments to fully commercial business models; the so-called valley of death in technology innovation (Foxon, Gross, et al. 2005). Targeted financial support is particularly critical during the these stages, in particular, support that allows experimentation and (intermediate) failure and allows initiatives to adapt and diffuse learning (Seyfang & Smith 2007). Learning is also easier in stable social and financial arrangements so consistent financial support is essential (Nooteboom 2000).

9.1.3 New approaches to valuation

One of the most contradictory selection pressures of the current policy and regulation regime is the measurement of value based on the marginal cost efficiency of delivering each unit of utility product. This excludes the value of protecting future customers as well as the value of non-monetary benefits described in chapters 7 and 8, and disincentivises alternative modes of operation.

A more plural approach to valuation of infrastructure operation is needed to value the non-monetary benefits of schemes, so social benefit generated by more local schemes is captured and assessed on a more equal footing with financial benefit. There is a need to balance the goals of economic, environmental and social sustainability, whilst taking into account both price and non-price issues (Mitchell & Woodman 2010). Traditional, business-case approaches to social and environmental accounting, often called triple bottom line accounting, attempt to quantify, often monetise, the value of social and environmental benefits of activities for business and shareholders (Brown & Fraser 2006). However, this funder- and investor- driven approach can actually serve to reinforce the dominance of economic measures and so may not be relevant in support of an alternative approach to valuation (Gibbon & Dey 2011).

By contrast, stakeholder-accountability approaches attempt to evaluate an organisations performance with regard to the accomplishment of social objectives (Chen 1975) and thereby increase the accountability and transparency of organisations. Accounting methods exist, such as Social Accounting, that enable organisations to report on their social value, without attempting to monetise it (Gibbon & Dey 2011). This method accepts the plurality of social value and may promote a broader understanding of social value. However, these methods tend to report on existing activities, rather than develop evidence to support the case for new modes of operation and may need to be adapted further to enable this kind of analysis.

This new approach to accounting and valuation should also take into account the benefits derived by future users, for example, by avoiding dangerous climate change. This would require that costs and benefits would need to be assessed over a longer period of time. For example, the requirement for efficiency savings to occur 5-7 years after investment in the water industry encourages short-termism and precludes investment in schemes that would generate more significant benefit, but in the longer term. This would also need to include means to assess value for money of alternative of modes of operation whereby profits are derived by alternative mechanisms such as selling energy saving or infrastructure service. This would be a substantial move away from the commodification of energy and water and towards the valuation of infrastructure services (Patterson 2008; Roelich et al. 2014a).

9.1.4 New approaches to regulation

Regulation of infrastructure is necessary to manage effects, such as natural monopolies, which would not be controlled by the market alone. However the goals of economic regulators in particular need to be realigned with wider goals of transitioning to a low-carbon energy system. Sustainability goals need to be equal to, or take precedence over, economic goals. One way to do this would be to extend measures of efficiency and cost-benefit to include social and environmental goals, which could be done using methods described above. These proposals go beyond calls to internalise externalities by introducing, for example environmental tax reform, which shifts taxes from ‘goods’ such as labour to ‘bads’ such as carbon emissions (Ekins et al. 2012); they aim to recognise the plural values of infrastructure rather than attempting to reduce all social and environmental value to monetary units.

9.1.4.1 New approaches in energy system regulation

The narrow focus on switching suppliers as a means to deliver customer protection should be replaced by one which protects the needs of customers now and in the future but does not preclude new entrants. This could include a new model of customer charter, which includes service standards for a wider range of services beyond providing of units of utility products (such as energy saving or energy services), and is relevant to a range of organisations (DECC 2013c; Knoeri et al. 2014). This kind of approach was demonstrated in cases such as Byker Community Trust, which had a long-term agreement to reduce total energy costs for residents supplied by a district heat network. This incentivises system-wide efficiency to improve both the resource efficiency of each unit supplied and reduce the volume of supply necessary to provide warm homes. The regulatory regime must adapt to enable these outcome-oriented, long-term agreements between suppliers and end-users. This requires removal of restrictions of long-term contracts so that operators were able to attract and retain customers for longer to justify investment in generation capacity and retrofit.

This should be accompanied by simplification of supply licencing arrangements, including removal of the need for smaller operators to enter into agreement with large, incumbent operators. This would necessitate concomitant modifications to the wholesale market to reduce the transaction costs associated with interacting with this market. There has been a great deal of research into market arrangements to encourage investment in low-carbon technologies. Recommendations include introduction of capacity markets, using central procurement of generation capacity and single buyer models with central procurement of both capacity and energy (Baker et al. 2010). However, there has been less analysis of the most effective market arrangements to encourage adoption of alternative modes of infrastructure operation. This is beyond the scope of this thesis but would merit further attention.

It has been argued that the period over which network investments are evaluated should be extended and that future network users should be engaged in the development of these long-term investment plans (Bolton 2011). It is likely that these policies would reduce connection costs and encourage more network connections from smaller operators. This could substantially reduce the constraints to projects, such as Woking Borough Council, who were forced to develop a private wire network in order to avoid excessive network charges.

9.1.4.2 New approaches to water system regulation

In its review of market reform proposals, Defra limited reforms to the non-domestic sector only. It decided that legally separating water companies’ retail functions, in order to increase competition in all markets, presented *"risks to a successful model given the challenges we face in building the resilience of the sector"*(Defra 2011: 3). The eligibility threshold for the non-domestic retail market was significantly reduced to increase the number of potential customers who could switch suppliers. It proposed that this was supported by reform to the water supply licencing (WSL) regime; removing restrictions on company structure, including sewerage services in retail competition, amending pricing regimes and introducing statutory market codes to encourage new entrants and make switching easier (Defra 2011b).

However, encouraging new entrants to a very limited market could be somewhat self-defeating; the reduction in threshold increased the potential customer base from 2,180 to 26,000 (Defra 2011b). This still represents a very small proportion of the 1,146,200 non-domestic premises and is a smaller proportion still of all customers in the UK. It seems unlikely that this would encourage many alternative modes of operation without creating access to a larger number of customers. Perhaps one of the biggest challenges to opening up competition in small non-domestic and domestic customers is the need to ensure very stringent water quality (European Commission 1998) and wastewater treatment standards (European Commission 1991; European Commission 2000). However, drinking water standards need not be applied to non-potable water, which could open up a market for provision of water treated to a lower standard or recovered from grey or rainwater. This would also enable the extension of services from district heating systems to include hot water, further increasing the efficiency of such schemes. This is supported by the EU Energy Efficiency Directive which encourages cogeneration (European Commission 2012a).

The new WSL regime includes the right for companies to negotiate discounts for demand management measures by non-domestic customers, including investing in surface drainage systems and reducing water use during peak periods or drought. However, demand management on a wide scale would affect the revenue of water companies in the long-term unless they are allowed to charge for water services and water savings in a similar manner to energy performance contracts. There are precedents of gain-share arrangements in the water sector but these are the exception, rather than the norm (Business Stream 2013). Regulation needs to adapt to take into account for alternative approaches to revenue generation, which will require new measures of value for money.

9.1.5 Enhancing capabilities

Infrastructure services have historically been delivered by a small pool of specialist utility companies, and the capability to operate infrastructure and engage customers is concentrated in these organisations. If other actors are to engage in infrastructure delivery, it is important to build their capabilities in both technical and operational areas. This was exemplified in the Chale Green case study where, alongside retro-fit of social housing, a number of residents were trained as energy advisors. The advisors were available to provide training and support to other residents in the operation of these low-carbon technologies and energy meters. This was effective in improving uptake of technologies and also increased awareness of climate change and energy saving (DECC 2012b). Moreover, capabilities in the community (and confidence in these capabilities) were enhanced by the additional support to such an extent that the community group was considering taking on more ambitious and complex projects, such as energy storage. This type of support should be made more widely available to provide project-specific capabilities but also to increase the ambition and potential of alternative infrastructure providers.

Increasing the capabilities of individual organisations is time and resource intensive and capabilities are often specific to the type of organisation and the mode of operation employed. However, for organisations with similar goals and institutional structures, enabling partnerships, like the London Energy Partnership, could be formed. These partnerships can overcome this challenge by pooling resources to build capabilities in particular issues – such as the Licence Lite (Haringey Council 2012) and setting up ESCos (London Energy Partnership 2007). More formal networks, such as the Core Cities consortium (of which the Newcastle City Council case is part)[31], have started programmes of collective capability building, where lessons from individual participants are used to create guidance or advocate change at a national level. Additional support should be provided for this collective building of capabilities, where appropriate, and for dissemination of the results.

The cases identified that the lack of technical understanding could be one of the principal barriers to more active engagement of local authorities in energy provision, which is supported by the findings of Bale et al. (2012). In the case studies examined in this thesis, this was overcome by entering into partnerships with technology providers (as at Woking Borough Council) or by appointing an energy specialist internally (as at Newcastle City Council). If these options were not available as a result of financial constraints or aversion to engaging the private sectors this constraint could be overcome in some circumstances by the development of common technology standards, for example for district heating and cooling (Hawkey et al. 2013). Furthermore, provision of sources of information or decision support tools could

---

[31] http://corecities.com/

support identification of more suitable technologies. This service is currently provided with regard to small-scale technologies for individuals by the Energy Savings Trust; it could be extended to address new actors in infrastructure provision, such as local authorities.

The challenges of new forms of contracts, partnerships and funding require different capabilities, which would also benefit from information provision and tools to identify the most effective and appropriate forms of contract and sources of funding. Attempts have been made to develop standard contract terms; such as the RE:FIT scheme in London (Mayor of London 2013). The programme streamlines the procurement process for retrofit of public sector building by providing pre-negotiated, EU-regulation-compliant contracts that can be used with a group of pre-qualified ESCos. The roll-out of this scheme to other parts of the public sector is being explored and is encouraged by the EU Energy Efficiency Directive (European Commission 2012a). However, this ignores the domestic sector, where transaction costs are arguably higher, and is restrictive in its format so does not support diversity in modes of operation. A more flexible process providing a variety of best practice partnership agreements and contracts could offer greater benefit.

Capability development in community-led modes of operation presents a more significant challenge, as a result of the diversity of aims, motivations and modes of operation. It is not desirable to standardise community-led projects or to define the ‘ideal’ capabilities (Hargreaves, Hielscher, et al. 2013; Seyfang & Haxeltine 2012). However, additional support could reduce the pressure on capabilities and increase the number of projects that could succeed. Intermediaries have an important role in building participant confidence and capability; which can be more important than technical knowledge in community schemes (Hargreaves, Hielscher, et al. 2013). This kind of support requires a great deal of time and resource; it can’t necessarily be done with a toolkit, and has been cut back in the current economic climate. There has been increasing activity to provide generic guidance but consistent resource should be available to provide specific support.

Networks or partnerships are a good way to maximise the resources and capacity available to alternative modes of operation (Seyfang & Haxeltine 2012). This is particularly the case for networks including a range of stakeholders who have access to a wider variety of resources such as the SPV formed at Woking (Thorp 2011) or the proposed partnership between Byker Community Trust and Newcastle City Council to maintain the district heating system after handover (Beadle 2014). This can reduce the pressure on an individual or an individual organisation to provide all necessary resource or capabilities and more effectively match capabilities to outcomes. There is a key role for policy makers in facilitating these networks and partnerships. This could be through formal networking opportunities or developing databases of organisations and capacities to encourage identification of partnerships.

9.1.6 Knowledge transfer

Knowledge transfer plays an important role in building capabilities and sharing learning between projects can reduce the barriers that individual projects face. This was demonstrated in a number of the case studies examined in this thesis. For example Eco-Island built on the experiences of the Chale Green retrofit projects and in turn the experiences of the Eco-Island team were informing the ambitious approach proposed at the Welborne project. The importance of intermediaries in accelerating learning within and between organisations is reinforced by Hargreaves et al's work investigating development of community energy projects (Hargreaves, Hielscher, et al. 2013). Intermediaries connect specific projects with one another and with the wider world and identify common issues and problems encountered across multiple projects and can help accelerate learning (Hargreaves, Hielscher, et al. 2013; Howells 2006).

This kind of learning was facilitated through personal connection in the cases examined in this thesis; however, this individual connection would be difficult to scale up. More structured forms of knowledge transfer would be needed to scale up this learning process. Policy makers have an important role in creating and maintaining this structure and there should be greater support for knowledge transfer and specific intermediaries.

One of the greatest challenges the case studies faced in transferring knowledge was the very different context in which each case operated. Knowledge transfer systems are essential to aggregate knowledge across a broad range of local projects to identify context-free lessons that work on a wider scale (Seyfang & Smith 2007). This requires 'institutional infrastructure' to collect, store and exchange this knowledge as well as the provision of advice, guidelines or even templates for future projects. There should be a greater support from government for this institutional infrastructure and for intermediaries to accelerate knowledge transfer.

Learning isn't just needed within organisations operating infrastructure; it's important that policy learning is improved and policy feedback loops are integrated into policy making processes (Foxon, Pearson, et al. 2005). This improves the effectiveness of policy, in relation to achieving outcomes, but could also lead to bolder policy decisions. It is also important that policy makers learn from others' experimentation and even failure (Seyfang & Smith 2007).

Experimentation provides insight into what is possible and failure can identify specific areas where policy support would be most effective. Therefore; there is a role for intermediaries between infrastructure operators and policy makers and regulators. This could help to identify the most effective support that balances the development of policy system with the needs of organisations. For example, the move from grant funding towards encouraging models of enterprise in community projects means that business models become more important and; therefore, more specific help on enterprise and management is required, rather than on technology.

9.2 Beyond constraints: governance system change in support of alternative modes of operation

In section 9.1 specific actions were identified to remove particular constraints identified in chapter 8 either by changing selection pressures or through supporting capabilities that enabled operators to respond to selection pressures. However, the majority of recommendations represent incremental changes from the current system of policy and regulation. Current policy and regulation has evolved with and supports the current mode of operation; therefore it could be contributing to path dependency and lock-in. Incremental change may be insufficient to break this lock-in (Unruh 2000). The alternative modes of infrastructure operation analysed in this thesis show characteristics which differ greatly from those of the mainstream mode of operation, which is dominated by large, centralised utility companies and excludes end users and local actors (as described in chapter 7). Furthermore, the complexity of the infrastructure systems is such that the outcomes of isolated interventions are difficult to forecast. Feedback loops, time and spatial lags and path dependency lead to unstable and unpredictable responses. Without a fundamental change in the system of governance, beyond incremental policy change, it seems unlikely that transition in infrastructure operation of the necessary scale would occur. In this sense governance is defined broadly to include ”*the use of institutions, structures of authority and even collaboration to allocate resources and coordinate or control activity in society or the economy.*” (Bell 2002 quoted in Hall et al 2012). It is not limited to the actions of national governments, to which recommendations in section 9.1 are limited, but includes the rules and policy developed and implemented by a complex network of non-state actors at international and sub-national levels (Smith et al. 2005). Partnerships and institutions developed by these non-state actors have been central to the success of many of the cases examined in this thesis. Therefore, it is important that these aspects are included in analysis of possible future systems of governance.

This section aims to connect and build on theoretical approaches to governance of transitions described in Chapter 2 and includes governance by non-state actors. It presents a conceptual framework that identifies necessary elements of an alternative system of governance that could be more supportive of the characteristics of alternative modes of operation identified in this thesis. These include; alternative operators, with a broader range of motivations than profit, creating social and environmental value as well as infrastructure services and engaging in partnerships and creating institutions with the public and private sector and end users. A fundamentally different approach to governance, based on encouraging diversity, could reduce the chances of becoming locked in to another, possibly equally unsustainable mode of operation and is more likely to accelerate transition.

9.2.1 Analysing alternative governance systems: a conceptual framework.

The conceptual framework presented in this section organises important elements and relationships identified in the analysis of existing theories in chapter 2 and from empirical research in chapters 7 and 8. The framework is designed to help analysts examine a diversity of governance systems to identify elements or relationships that are absent from a system of interest. From this, alternative approaches to governance can be identified that are more supportive of the characteristics of alternative modes of operation. Frameworks can act as "*metatheoretical devices that help provide a general language for describing relationships at multiple levels and scales."* (Ostrom 2010: 659). The framework is presented in figure 9.1 and described in more detail below.

The key elements of the system are contrasted with those used to describe the current infrastructure governance system in Chapter 7. It is clear from this representation that the relationships between different parts of the system are substantially different from the current system. These relationships are described below.

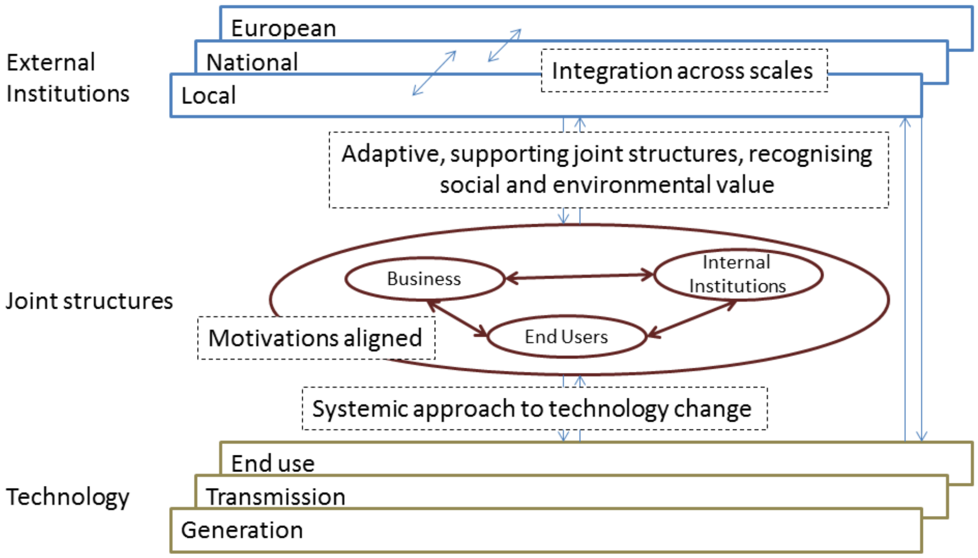


Figure 9.1: Governance analysis framework. Actors and assets are represented as boxes. The relationships between actors and assets are represented as directional arrows. Particular features of interest are highlighted in dashed boxes.

Business and end-users are no longer separate but combine to form joint structures, or partnerships, the stability of which are important to long-lasting transition (Foxon 2011). The creation and evolution of a joint structure of business, end users and often governmental organisations was a key factor of success identified in case study analysis. These joint structures align the interests and motivations of the heterogeneous actors, whilst still delivering benefits to all participants. Governance has a role to play not just in supporting the creation of these joint structure but in minimising disruption at the macro-level (for example by removing market regulation barriers to district heating systems) to support the longevity of joint structures.

The scope and role of actors has broadened, including businesses and end users. Many different types of organisations participate in the transition, including technology providers, property developers, local authorities and community groups as well as the traditional locus of utility companies. End users become active participants in infrastructure system transition. This diversity in participants in infrastructure needs to be better recognised by governmental institutions, such as policy makers and regulators.

The joint structures have internal institutions to set the rules of interaction and increase their resilience to change. These internal institutions maintain alignment of motivations and fair distribution of benefits between participants. Internal institutions are analogous to institutions necessary for self-governance described in the work of Ostrom and colleagues in relation to natural resources (Ostrom 2005). To be successful, internal institutions will need to interact with institutions at other levels and responsibility for governance should be distributed across these levels. The diversity of institutional structures would imply that a less structured and more problem focussed approach to governance would be relevant, which has the advantages of being more flexible and encouraging innovation. However this presents significant challenges in communication and monitoring progress, can reduce the stability of the policy environment and reduces certainty about future policy and represents a significant change from the current top-down approach (Marks & Hooghe 2003).

This flexible and dynamic approach requires that external institutions are more integrated and adaptive. The formation of joint structures is encouraged and enabled by institutions at all levels. The description of these institutions as ’external’ implies a division between internal and external actors. In reality the same organisation could operate in both systems; for example, DECC could partner with a private sector organisation in a scheme, which would have to operate under the policy set by DECC.

The relationships between external institutions and joint structures are reciprocal (rather than top-down) and dynamic to allow co-evolution of governance and joint structure operation. These relationships support the internal institutions of the joint structure by allowing them to, for example generate revenue and interact with other joint structures unimpeded. A significant challenge for external institutions is that the form and scale of joint structures will vary significantly.

Interaction between technology and end-users is more systemic and co-evolving. This is moderated by the joint structures who promote interaction with the infrastructure system in a more integrated and effective manner (for example, selecting more complementary supply and end-use technologies to reduce the need for storage).

Finally, the relationships between different parts of the system will co-evolve and both internal and external governance will need to adapt to this co-evolving system, whilst maintaining the focus on sustainability.

9.2.2 Application of conceptual framework to UK infrastructure governance

When the conceptual framework is applied to the current system of infrastructure policy and regulation in the UK, it becomes clear that there are some key deficiencies in the current system, over and above the constraints identified directly from case study analysis. These deficiencies are most directly related to diversity of modes of operation; adaptability of policy;

support of partnerships and self-governance. These deficiencies are discussed in more detail below.

9.2.2.1 Diversity

The portfolio effect, combining different technologies or markets, is widely recognised to hedge different risks and can mitigate the challenge of predicting which technologies will be most effective in the future (Watson 2008; Awerbuch 2005). More diverse approaches and forms of innovation are more likely to enable transition in complex system, such as infrastructure (Leach et al. 2012). Furthermore, a more diverse portfolio of technologies or modes of operation means that the infrastructure could be more resilient because a smaller part of the system is affected if a particular technology or mode of operation becomes obsolete. However, diversity in both technologies and modes of operation could require some compromise in other aspects of performance (Stirling 2007). For example, network infrastructure may be required to invest in smart technology to enable more intermittent generation technologies and better match supply and demand (DECC 2009). Furthermore, economies of scale may be lost by having a greater variety of bespoke modes of operation. It is important that these implications are considered and investigated in more detail.

In this thesis, it is argued that diversity of modes of infrastructure operation is equally important in breaking the lock-in to the dominant, centralised mode of operation and avoiding lock-in to another, equally unsustainable mode of operation in the future. Diversity should be encouraged in business models, in the same way as technological diversity. The current regulatory system, which has evolved support the monolithic, mainstream mode of operation, is incapable of achieving this diversity.

It is important to note that diversity is not simply a function of having lots of different options; to exploit the full benefits of diversity, it is important to have variety, balance and disparity (Stirling 1998). Variety is a measure of the different options that are supported within the portfolio. Balance is the profile of shares of these different options within the portfolio. Disparity represents the extent to which options are different from each other. Risks of lock-in are only mitigated if the portfolio of modes of infrastructure operation displays these features.

It is important that policy and regulation empowers a wider space for innovation by addressing the distortions and structural inequalities that exist in current policy and market conditions (Hargreaves, Hielscher, et al. 2013). The current system must fundamentally change to address these distortions to support a greater variety of modes of operation. Support is also required to develop a skills base commensurate with the promotion of option diversity (Foxon, Pearson, et al. 2005).

9.2.2.2 Self-governance

National policy and regulation also needs to recognise and respect the diversity of aims and approaches to alternative modes of operation, particularly in relation to local authority and community-led modes of operation. It is important to enable flexible and locally relevant institutional arrangements to deliver on these aims and distribute benefits locally. This means that the goals and execution of regulation, and in particular economic regulation, must change to allow this local definition and management of value. This presents a significant challenge for policy and regulation in monitoring the effectiveness of operators given the many different objectives they are set up to serve and the different contexts in which they emerge (Newell et al. 2012). The principles of self-governance described in chapter 2 could be used as a framework around which national governance could be modified to monitor and enable this self-governance.

The corollary to this is that local action must be held accountable for its contribution to pressing national goals (Leach et al. 2012). There is an understandable tendency for self-governance to reflect the needs and aims of the users in the locality. The prioritisation of these aims over nationally important goals, such as infrastructure transition, could undermine the contribution of local action and self-governance. Policy and regulation must create a framework of incentives and support that enables locally relevant action at scale, which contributes to national goals.

The design principles for supporting institutional survival described by Ostrom et al have a high degree of relevance to infrastructure (Ostrom 2010). The application of Ostrom's design principals to inform the design of internal institutional arrangements for infrastructure operation are discussed in more detail below. They could support the stability of joint structures and increase their resilience to external influences. This could also help to guide the structure of external governance to ensure that disruptions are minimised.

1A. *User Boundaries*: Clear and locally understood boundaries between legitimate users and nonusers are present.

Control of access to infrastructure systems is an essential precursor of efficient system management and fair distribution of costs and benefits. For example, illegitimate connections to sewer networks can cause flooding if design flows are exceeded or pollution if wastewater is connected to a sewer designed to carry surface water. Furthermore connection without contributing to the cost of the infrastructure, either through user charges or contribution to construction costs increases the cost to legitimate users. Management of access has presented challenges even for large operators with extensive resources to monitor and manage connections; mechanisms must be developed to enable smaller operators to regulate access.

1B. ***Resource Boundaries*****: Clear boundaries that separate a specific common-pool resource from a larger social-ecological system are present.**

It is very difficult to define clear boundaries of alternative modes of infrastructure operation as a result of the extensive interconnection of assets and between systems. Many modes of operation will rely on shared infrastructure assets to connect supply and demand, such as electricity distribution networks. Furthermore, they may rely on existing institutional structures such as national or local policy to supplement their own operating procedures. It is important that these interconnections are identified and the contribution of shared assets and institutions is clearly defined to ensure that the costs and benefits of such interconnections are reflected in the institutional arrangements.

2A. ***Congruence with Local Conditions*****: Appropriation and provision rules are congruent with local social and environmental conditions.**

It is important that the supplier-user arrangements reflect the needs and capabilities of local users. For example the charging system for heat at Byker Estate was a fixed monthly cost, because the majority of users had low and fixed levels of income and needed certainty about monthly costs. This could be extended to include consideration of congruence with global environmental conditions to address local contribution to national/global goals, such as climate change mitigation. This could encourage projects which not only develop sustainable infrastructure supply but also reduce demand for infrastructure products.

2B. ***Appropriation and Provision*****: Appropriation rules are congruent with provision rules; the distribution of costs is proportional to the distribution of benefits.**

Charging is likely to be very different in alternative modes of infrastructure operation as a result of delivery of alternative forms of value or delivery of an infrastructure service and demand management. Payment mechanisms are needed that capture the benefits of non-monetary value and demand management and compensate operators appropriately for delivery of these benefits. This might be a challenge when benefits accrue to those outside the defined user boundaries. In this case, some form of government subsidy might be required to support schemes contributing to national or long-term goals.

3. ***Collective Choice Arrangements*****: Most individuals affected by a resource regime are authorized to participate in making and modifying its rules.**

The smaller scale of operation and closer connection with end users presents an opportunity to better engage end-users in determining appropriate institutional arrangements. One way this could manifest itself is through increasing participation in accounting for social value and tariff setting (to include payment for social value). This is important to ensure that value delivered by the infrastructure scheme remains relevant to the end users. Participation of end-users has also been shown to increase acceptance and support for infrastructure project (Walker & Devine-Wright 2008).

4A. ***Monitoring Users*****: Individuals who are accountable to or are the users monitor the appropriation and provision levels of the users.**

Metering has traditionally been used to determine the appropriation of infrastructure products, based on consumption of units of energy or water, and set user charges. Modes of operation which deliver alternative services, such as thermal comfort, or deliver additional value may need to find alternative ways to monitor user consumption. Mechanisms may be required to encourage better matching of demand (appropriation) and supply (provision). This might include incentives for demand reduction at certain times (such as time of use tariffs) or contractual agreements to manage peak demand (similar to those currently used to curtail load in industry) (Ofgem 2010). An example of this kind of system at small-scale is the energy system on the Isle of Eigg. All users are subject to a 5kW cap per day and are provided with smart meters to monitor usage (Yadoo et al. 2011). Further voluntary measures are used to encourage users to reduce their demand when generation from renewable sources is running low. This includes a ‘traffic light’ system indicating when generation is low and suggesting that islanders should reduce their consumption. Some of these measures require collective action, but there is evidence from natural resource management that this is possible and effective.

4B. ***Monitoring the Resource*****: Individuals who are accountable to or are the users monitor the condition of the resource.**

This could have two elements; monitoring the condition and efficiency of the infrastructure to ensure that it continues to deliver value as effectively as possible. This could be done using some form of benchmark against which to compare performance. This could ensure that technology is maintained and upgraded regularly and that the costs of this maintenance and upgrade are included in user charges. Another aspect of this principle could be the monitoring of the contribution of individual projects to national goals, such as carbon emissions reduction, to ensure that local projects also monitor their effect on global resources. Mechanisms exist that allow mainstream operators to report carbon mitigation measures though schemes, such as ECO. However these would need to be adapted to be relevant to smaller, more diverse organisations.

5. *Graduated Sanctions*: Sanctions for rule violations start very low but become stronger if a user repeatedly violates a rule.

Sanctions will need to be in place in the event that either the supplier or user breaks agreements about provision or appropriation. These may have to be more comprehensive and specific than current supplier agreement which simply list the conditions for payment, metering, data management, supplier transfer and the conditions under which energy or water may be cut off (Ofgem 2014). These conditions would need to be adapted to account for the diversity in modes of operation and the different basis of agreement, including service agreements. This may require the development of additional metrics to measure a wider variety of services provided. For example, service metrics, such as usable floor area at 18C, might be required for service-oriented contracts rather than provision of electricity or gas (Knoeri et al. 2014). Standards of service are currently set out by the Government; including standard levels of compensation in the event that standards are breached. These standards and the associated compensations would also need to be adapted to reflect the variation in required standards and appropriate sanction measures.

6. *Conflict Resolution Mechanisms*: Rapid, low cost, local arenas exist for resolving conflicts among users or with officials.

Mainstream infrastructure operators have formal complaints management systems, which are monitored by economic regulators and a free Ombudsman Service is available to help settle customer complaints in a fair and unbiased way. Locally relevant versions of these complaints procedures will need to be developed and be monitored nationally in some way. This will obviously increase the burden on national regulators to monitor conflict resolution. However, it is possible that the creation of locally relevant procedures could prevent complaints escalating to the stage at which intervention is required.

7. *Minimal Recognition of Rights*: The rights of local users to make their own rules are recognized by the government.

This means that the goals and execution of regulation, and in particular economic regulation, must change to allow local definition and management of value along with local arrangements for customer protection. This would require adaptation of supplier licencing regimes and other mechanisms to enable alternative modes of operation to manage service provision and customer service in an appropriate manner whilst ensuring that customers and the environment were adequately protected.

8. *Nested Enterprises*: When a common-pool resource is closely connected to a larger social-ecological system, governance activities are organized in multiple nested layers.

The connection between governance of local system and rules for governance of wider infrastructure systems needs to be recognised and articulated. This is also true for the connection between local infrastructure and national infrastructure systems. The effect of alternative modes of operation on the national infrastructure system must be more clearly understood and the implications managed.

This analysis presents a first attempt to articulate some principles for self-governance of infrastructure, based on evidence gathered in this thesis and a limited review of literature. In order to develop more details and robust recommendations these initial principles would benefit from more detailed analysis and empirical evidence.

9.2.2.3 Policy adaptation

Perhaps one of the most significant challenges that governance of infrastructure faces is the uncertainty associated with both the scale of change necessary and the response of the infrastructure system to intervention. As discussed in the introduction to this thesis, climate change in driven by the cumulative quantity of carbon emissions and if action in the short-term is ineffective, the cumulative budget will be used up more quickly and the subsequent rate of emissions reductions must accelerate significantly (IPCC 2013). These uncertainties are not statistical in nature and cannot be reduced by gathering and analysing more information (Walker et al. 2013). This means that the future cannot be predicted and a static, “optimal plan” is unlikely to be successful (Haasnoot et al. 2013). Instead policy must be adaptive and able to respond to changing future conditions. This is a considerable challenge for policy-makers who must provide a stable policy framework but also plan to change over time.

Adaptive policy is developed in light of the uncertainty over the potential scale and speed of system change. It is designed to be changed over time as new information about the effectiveness of policy in the short-term and the scale of the challenge in the long-term becomes available. The history of adaptive policy can be traced back to 1927, when John Dewey proposed that “policies be treated as experiments, with the aim of promoting continual learning and adaptation in response to experience over time” (Dewey, 1927). A series of approaches and computational tools have been devised to support development of adaptive policy but most have been applied to climate change adaptation (Walker et al. 2013; Haasnoot et al. 2013; Foxon et al. 2009; van der Pas et al. 2013). The application of these approaches to the field of low-carbon infrastructure transition would merit further attention.

9.3 The Policy Paradigm – are these proposals possible?

Policy processes take place within a policy paradigm – a framework of ideas and standards which shape policy goals, instruments and setting of instruments. If the policy paradigm is not supportive this could prevent the implementation of the specific proposals and wider system reform suggested in this thesis. Some proposals require first or second order change, where the settings of instruments or the instruments themselves change – in this case it is important that proposals correspond with the current goals of policy (Hall 1993). However, many, in particular the proposals for wider system reform, represent third order change in the hierarchy of goals, which requires a corresponding change in the policy paradigm. In this section the potential for the policy paradigm to enable the recommendations in this thesis is discussed. The framework for analysis of policy paradigms developed by Kern et al will be used to structure this analysis (Kern et al. 2014). The framework describes the paradigm as being a function of interpretive frameworks, objectives, instruments and governance institutions.

Kern et al observe that "*Although the interpretive framework has altered to include climate change and geopolitical ideas, elements of belief in market ideas continue to persist alongside*" (Kern et al. 2013: 10). The presence of different perspectives, and in particular the persistence of pro-market beliefs, means that goals and policies are 'picked and mixed' between perspectives, regardless of the compatibility of those goals and policies (Keay 2012). A particular challenge is the contestation of the relative importance of different perspectives across different government departments, which further erode the coherence of resulting policy.

The introduction of goals carbon emissions reductions (HM Government 2008a), energy security and fuel poverty reduction (HM Government 2007) signals a significant divergence from a sole focus of policy and regulation on cost efficiency (Hall et al 2012). However, the integration of these goals into coherent policy remains a challenge. This is exemplified by economic regulation, which as it stands is unable to take a more integrated approach and achieve the balance between economic, environmental and social goals (Mitchell & Woodman 2010). This is clearly reflected in the government's Principles of Economic Regulation which are dominated by economic goals (BIS 2011). This is despite the economic regulators having an equal duty to contribute to the achievement of sustainable development (HM Government 2008b).

There are signs that policy instruments are also departing from being entirely market-based. The implementation of FiTs for certain renewable technologies, and the extension of the scheme to renewable heat, allows smaller producers to compete more effectively with large companies. The FiT specifically contradicts the principal that policies should be technology blind. This is further reflected in the Renewable Obligation for large producers, which is moving from purely market-based to more of a banded system to offer differential support to less mature technologies (R. Gross & Heptonstall 2010). This indicates an increase in support for targeted policy to overcome lock-out of renewable technologies and smaller producers. This targeted support has yet to extend to alternative modes of operation.

Specific institutions have been set up to support alternative infrastructure providers. These include a specific team at the Department for Energy and Climate Change, which has produced a Community Energy Strategy and a Community Energy Contact Group, and the Heat Network Development Unit, which provides support to local authorities looking to develop heat networks. However, the role of these institutions is predominantly to provide support and guidance. There is little indication that they will affect the mainstream system of incentives and regulation which favours the incumbent, private operators.

These changes suggest some movement away from a purely pro-market paradigm and less of a reliance on market-based policy instruments. However, there is no indication of any change in the definition of customer protection (as cost reduction now) or in the goal of promoting competition above all else. There would need to be a parallel movement away from these goals, in particular a *“shift to achieving sustainability (at the lowest cost but defined in a long-term sense which include qualitative factors) rather than establishing a cost and seeing how much sustainability comes out (as [the paradigm] is now)”* (Mitchell 2010: 212).

Many of the proposals in sections 9.1 and 9.2 require a more fundamental change in the way that costs and benefits are assessed and in the way that diversity is encouraged. The independence of economic regulators is ensured by detailed, quantitative analysis of proposals; however, these analyses use economic efficiency as the basic criterion of assessment of costs and benefits. This is still rooted in the original purpose of the privatisation and liberalisation of infrastructure sectors to improve cost efficiency and reduce monopolistic competition. However, recent concerns, such as climate change, energy security and access to energy are excluded from this approach to analysis and it could be argued that we will be locked into the current governance system until the appraisal metrics are revised.

There is evidence of change in the paradigm to incorporate new perspectives, goals, instruments and institutions which are more supportive of the proposals set out in this chapter. However, the persistence of the pro-market perspective alongside emerging perspective of sustainability limits the coherence of resulting policy and institutions (Kern et al 2014). This allows economic goals to dominate as a result of the hegemony of economic regulation and the Treasury. This favours large, centralised, private actors and limits the potential for both diversity in modes of infrastructure operation and self-governance of infrastructure by communities and local authorities. The rigid focus on short-term cost effectiveness limits the ability to foresee alternative paths to a low-carbon infrastructure system and mitigates against adaptive policy making.

9.4 Summary

This chapter begins by presenting policy proposals that respond specifically to constraints identified in chapter 8. These proposals represent incremental change from the status quo but are limited in their potential to address the lock-in of policy and regulation to the current mode of operation. Using an evolutionary analogy, the proposals described either attempt to change the selection pressures acting on alternative modes of infrastructure operation to be more favourable or co-ordinate resources to enable alternatives to respond or adapt to these pressures.

Proposals include a more integrated approach to all aspects of infrastructure operation, including the goals of policy, addressing the interdependence of infrastructure systems and addressing all project stages and organisations. There is a need for policy and regulation to better recognise the complex system dynamics of infrastructure operation and respect the multitude of values and priorities of organisations and individuals that could be engaged. Selection pressures should be more coherent and better aligned with sustainability. In particular, the goals of economic regulators need to be realigned with the transition to low carbon infrastructure. Support for alternative modes of operation needs to be more targeted to reduce uncertainty and risk, particularly during the early stages of project development. This support should also aim to build the capabilities of actors engaged in infrastructure operation and enable knowledge transfer. A more plural approach to valuation is needed to account for non-monetary value creation and to balance economic, environmental and social sustainability.

Beyond these incremental changes, a more fundamental shift in the system of governance is needed to break the lock-in with incumbents and to be more supportive of alternatives. A conceptual framework is proposed to help with the analysis of governance systems to identify elements or relationships that are absent from the system of interest. When applied to the UK infrastructure governance system it becomes clear that there are some key deficiencies in the current system. The current system of governance does not enable the diversity in modes of infrastructure operations necessary to hedge different risks, improve resilience and enable transition in complex systems. The current system must address the distortions and structural inequalities to enable a greater variety of modes of operation. National policy and regulation also needs to enable flexible and locally relevant institutional arrangements to deliver on a diversity of aims and approaches to engaging in infrastructure operation. In contrast, local action must be held accountable for its contribution to national goals. Finally, policy must be developed in the face of uncertainty over the potential scale and speed of system change and must be designed to enable change as new information about the effectiveness of policy in the short-term and the scale of the challenge in the long-term comes to light.

Both the incremental proposals and the systemic change presented in this chapter rely on change in the pro-market policy paradigm to succeed. There is some evidence that new perspectives, such as climate change, are being incorporated into the current paradigm but this change is not consistent or coherent, which allows economic goals to dominate still. Further consolidation of paradigm change is required before these proposals could be achieved.

10 Discussion: Analysing constraints to infrastructure transition

This chapter assesses the contribution of this thesis to the analysis of constraints to low-carbon infrastructure transition. It is structured in response to the research questions which were outlined in the introductory chapter. Importantly, the discussion compares findings from across both material and policy and regulatory constraints and considers whether there are common lessons from these seemingly distinct cases.

The chapter starts by reflecting on the importance of a clear conceptualisation of constraints and the influence this has over the selection of methods and form of knowledge created. It goes on to consider the methodological advancements made in this thesis and the limitations of the methods used. The key findings related to the nature and scale of constraints under analysis are summarised and the relative ease of interpretation of these findings is discussed. Following this, policy responses identified to mitigate constraints and accelerate low-carbon infrastructure are compared and responses common to both constraints are highlighted. The penultimate section of this chapter considers how the approaches to constraint analysis have helped to improve our understanding of low-carbon transitions and the role of governance in accelerating transition.

The chapter concludes with a discussion of future methodological and theoretical developments that might address some of the limitations of methods and analysis identified throughout the preceding sections.

10.1 Conceptualising constraints

The analysis undertaken in this thesis has demonstrated the importance of clear and transparent conceptualisation of constraints. Clear articulation of the constraint improves the interpretation of results and enables comparability with other, similar studies. The conceptualisation of constraints also influenced the analytical approach and choice of methods; ultimately affecting the form of knowledge that was created through analysis. This has implications for the relevance of knowledge to both policy analysis and overall system understanding; showing the importance of problem framing and conceptualisation to research in this area (Leach et al. 2010).

10.1.1 Conceptualising material criticality constraints

The importance of clarity was demonstrated strongly in the analysis of critical material constraints. Previous assessments have used somewhat opaque conceptualisations of criticality, which has made assessments difficult to interpret and compare (Erdmann & Graedel

2011). This thesis specifically conceptualises criticality as the potential for disruption of the supply of critical materials to slow down or halt the planned installation of low-carbon technologies in the UK necessary to achieve UK emissions reductions targets. This conceptualisation is analogous to risk, which is conventionally defined as the product of the probability of an event and the severity of harm resulting from that event (Renn 2008). Risk is a well-established process, familiar to policy-makers and industry alike so this conceptualisation could make results more accessible to these parties (Royal Society 1992).

Typically, risk-based approaches evaluate the risks of different options in quantified terms and identify preferred options and mitigation factors (Renn 2008). This lends itself to a quantitative approach which enables comparison of different pathways, which influenced strongly the indicator-based method identified as appropriate for this analysis. This also drove the externalisation of policy responses, such as recycling and substitution, which were represented as mitigating factors rather than incorporated in the assessment of criticality. To some extent, the conceptualisation of material criticality constraints also determined the relationship between metrics. For example supply disruption potential was represented as an underlying potential for imbalance between production and requirements and a series of factors (such as co-mining) which might exacerbate this imbalance. The conceptualisation also implicitly addressed the exposure of a particular technology or infrastructure goal, rather than simply identifying which materials were critical. This was an important dimension of the analysis and enabled a more systemic analysis of material criticality constraints.

10.1.2 Conceptualising policy and regulatory constraints

The conceptualisation of policy and regulatory constraints had two dimensions, which both provided essential structure for later analysis. The first drew parallels with evolutionary ideas whereby processes of policy and regulation (at the landscape and regime level) were seen to prevent or limit the development of alternative modes of operation (at the niche level). The second element was the articulation of three ways that development of alternative modes of operation could happen including scaling out, scaling up and translation to another sector or setting. Thus, the conceptualisation adopted provided more clarity over the nature of the constraint (factors affecting the success of an evolutionary process) and allowed for multiple ways in which that evolutionary process could happen, which provided more structure for analysis.

The centrality of evolutionary ideas to the conceptualisation drove methods selection in a number of ways. An analytical framework was developed that focussed on the interactions and relationships between actors (because evolutionary change emerges as a result of actor interaction) but also took into account the fact that these processes occurred at different levels in the system. The evolutionary conceptualisation also influenced the articulation of constraints as either unfavourable selection pressures or as a lack of capabilities to respond to these selection pressures. Furthermore, policy responses were described as the means by which to remove or re-align selection pressures or to enhance or co-ordinate capabilities to respond to these pressures (Smith et al. 2005). The necessary focus on actors and relationships demanded a more qualitative approach to analysis. Therefore, case study analysis was used to examine how actors and relationships in existing cases had resulted in emergence of new modes of infrastructure operation; an approach driven by the evolutionary conceptualisation.

10.1.3 Linking conceptualisation and knowledge

The conceptualisation of the constraints has a great deal of influence on the forms of knowledge produced by each line of enquiry. Material criticality was conceptualised as a predominantly physical constraint caused by disruption in supply of critical materials. Perhaps as a result, the quantitative, indicator-based method of analysis was concerned with events and processes which could explain and predict this disruption. Policy responses that resulted from this form of analysis were about controlling this disruption. This form of knowledge could be considered to fall within Habermas' empirical-analytic category, which is often more closely associated with natural sciences (Habermas 1972). The conceptualisation of criticality as a physical constraint could have driven this line of enquiry and resulted in analysis which was, to some extent, limited to explanation of causality and prediction and resulted in policy interventions focussed on control.

The more qualitative approach used for analysis of regulatory and policy constraints, aimed to understand the interactions between individuals and organisations which resulted in change (and constraints to those interactions). This focus on social interactions, as well as the institutions necessary to underpin these interactions, provided more detailed theoretical insights. The analysis was more time consuming but was more fruitful than just identifying drivers of physical constraints in isolation of system change. By contrast with material criticality, the conceptualisation of policy and regulation as an institutional constraint was more aligned with Habermas' historical-hermeneutic form of scientific enquiry (Habermas 1972). As such, it required research strategies that paid far closer attention to interactions between individuals and institutions and resulted in a greater understanding of a broader part of the system.

10.2 Methods for analysing constraints

The methods developed as a result of this conceptualisation make a number of contributions to the literature, which are summarised below.

10.2.1 Indicators for material criticality constraints

The indicator set developed in this thesis aimed to go beyond current criticality assessment based on static indicators by providing more quantitative analysis, systemic, goal oriented analysis, explicitly including dynamic metrics and externalising adaptive capacity. These advancements are discussed in more detail below.

All metrics used in the indicator set developed in this thesis are readily quantifiable, without the need for expert judgment. Furthermore, the supply disruption potential indicator is normalised to iron to which allows the expression of the relative criticality; the magnitude of the increase in criticality. This supports a consistent approach to analysis of a wide range of materials and technologies and allows comparison of relative criticality, which overcomes the arbitrariness of the individual indicators.

The systemic nature and goal orientation of the approach used in this thesis allows the analysis of the relative risk to different pathways to achieve the goal of low-carbon transition (demonstrated using DECC's Core and Renewables pathways). It moves beyond identifying materials which themselves are critical, to identifying how this criticality might constraint a particular goal. The methodology is used for the analysis of material criticality constraints to the goal of energy system decarbonisation but could readily be extended to other societal goals.

The temporally dynamic analysis of material criticality constraints, by projecting many indicators to 2050, allows analysis of the evolution of criticality with time. This provides more specific and relevant information to support decision making under uncertainty and may prevent sole reliance on pathways and technologies that could become highly critical in the future, avoiding 'lock-in'.

The identification of policy interventions is further improved by the externalisation of the adaptive capacity of the system (such as the potential to substitute materials). The effect of policy interventions, such as substitution (of materials or technologies) or recycling, on criticality could be explored through exogenous scenarios. This would support analysis of the effectiveness of different interventions to help shape policy approaches to reduce the criticality of particular goals.

10.2.2 Case study analysis of policy and regulatory constraints

The analytical framework developed for analysis of policy and regulatory constraints offers a number of enhancements that improve the application of current socio-technical and co-evolutionary theories to transitions in modes of infrastructure operation. These include; a clearer articulation of the connection between niches, regimes and the landscape as processes in one level which affect the ability of another to function or change; the exploration of processes that enable niches to co-exist with, not just overthrow, regimes; a more differentiated analysis of niche 'up-scaling'; and a focus on relationships and dynamic processes to better represent the infrastructure system as a complex adaptive system.

The analytical framework specifically accounts for the interactions between the landscape and regime and between the regime and the niche. This highlights the close connection between levels, and how the characteristics or processes at one level can strongly constrain action at another. This is exemplified by the effect of the policy paradigm (at the landscape level) on policy goals, instruments and settings (at the regime level) and by the constraints created by economic regulators (at the regime level) on the operation and evolution of alternative modes of operation. This calls into question the concept of protected space which is central to transitions management and in particular strategic niche management.

These interactions gain in relevance and importance, since the framework aims to identify processes and relationships that enable co-existence of the alternative mode of operation with the regime. The notion of co-existence is different to most interpretations of the MLP, which aim to identify how a niche might overthrow a regime. Considering co-existence emphasises the necessary changes to processes at both the niche and the regime level to enable niche evolution. This reflects the argument that a combination of top down and bottom up governance is necessary to deliver change in complex adaptive systems (Rijke et al. 2013; Huntjens et al. 2012; van de Meene et al. 2011).

The analytical framework includes the three different evolutionary pathways for increasing the contribution of alternative modes of infrastructure operation; scaling out, scaling up and translation. This provides greater resolution of analysis and reveals the different constraints faced by different evolutionary paths. This provides greater potential to distinguish between different modes of operation; their varying potential to contribute to future infrastructure operation and the differential support required to overcome specific constraints. This support is highly dependent on the nature of the mode of operation and the desirable evolutionary pathway; therefore this additional granularity in analysis is important.

The focus of analysis on relationships and dynamic processes, rather than structure, better reflects the evolution of complex adaptive systems, which occurs from the interaction of actors (Foxon 2011). This underlines the importance of collective action and formal and informal networks in disrupting existing patterns of behaviour, encouraging novelty, or creating windows of opportunity by creating structures, rules, interactions, interdependencies, tension and culture (Olsson et al. 2006). This is also a more effective way to identify the mutual causal influences that underpin co-evolution and increases the analytical framework's relevance for analysis of social innovations.

This focus on relationships and processes resulted in less reductive approach to analysis. The unit of analysis was taken to be the system relevant to the case study of interest, rather than the energy supply system alone, for example. Modes of operation, and the practices they support, often cross more than one system (for example, cooking draws on food, water and energy regimes) and the separation of these systems can overlook key processes and relationships crucial to understanding transition dynamics.

10.2.3 Methods, complexity and action

Despite having markedly different conceptualisations and methods, the overarching aspiration of both analyses was to be systemic in nature. This systemic analysis manifested itself in different ways, depending on the challenges of the methods used. The quantitative analysis of material criticality constraints used the goal of low-carbon transition as the focus of analysis to drive a more systemic and temporally dynamic approach, which is far more relevant to low-carbon transitions than previous approaches. However, it was necessary to reduce the complexity of the system to a series of quantifiable metrics that could define the most significant parts of the material criticality system. A great deal of resolution was lost in this process, in particular the influence of actors on the evolution of individual metrics and, to some extent, the interaction of metrics themselves. Nevertheless, even the current high-level systemic analysis provides valuable insights for policy makers and industrialists and presents a first attempt at capturing some of the complexity of this system. This type of analysis can be used to identify the most important drivers of criticality constraints requiring urgent action or further research.

The systemic nature of policy and regulatory constraints analysis was related to the examination of the interaction of institutional, technology and actor systems, which resulted in the emergence (or not) of alternative modes of operation. This allowed for a richer understanding of the system and, in particular the interaction between alternative modes of operation and policy and regulatory actors. However, this analysis was more time consuming and the descriptive nature of the results makes them less readily communicable to policy makers than the more simplistic, quantitative results generated by the criticality indicators. Despite these limitations the richer understanding derived from qualitative analysis enabled the identification of a wider range of more fundamental policy responses; it is still possible to identify recommendations for action if this is central to the analytical approach.

The two approaches represent contrasting approaches to systemic analysis, which have attempted to balance understanding complexity with enabling action in different ways. The qualitative approach to analysis provided greater understanding and addressed complexity more effectively, but the simplicity of the qualitative approach could be a more effective way to engage policy makers with the results of analysis of this nature.

10.3 Nature and scale of constraints

The different approaches to conceptualisation and analysis of constraints resulted in quite dramatically contrasting characterisations of the mechanisms by which low-carbon transitions could be constrained and the extent and severity of these mechanisms of constraint. Perhaps unsurprisingly, the more quantitative approach to analysing material criticality constraints better characterised the scale of the constraint. The qualitative approach used for analysis of policy and regulatory constraints had better explanatory powers and provided greater insight into the nature of constraints.

10.3.1 Material criticality constraints

The results of analysis of neodymium criticality constraints to low-carbon electricity showed a significant increase in criticality in both low-carbon electricity pathways examined. The ‘Core’ low-carbon electricity pathway showed a four-fold increase in criticality over the period from 2012 to 2050. The ‘Renewables’ pathway showed a more dramatic, almost ten-fold, increase as a result of its greater reliance on wind turbines. The temporally dynamic analysis also provided information about how the scale of material criticality constraints changed over time. The overall results for both pathways showed a steep increase after 2030 when there was a step-change in roll out of wind turbines.

The nature of material criticality constraints is explained, to some extent by the definition of and trends in individual metrics. The first important insight into the nature of material criticality is the observation that a likely decrease in supply disruption potential would be outweighed by an increase in exposure of the goal of low-carbon electricity to this supply disruption. The price sensitivity metric of exposure is held static; therefore the change in exposure is dominated by goal sensitivity, which increases in the case study analysed in this thesis as the electricity system becomes increasingly reliant on wind turbines. That is not to say that supply disruption potential is not important, but that it is not contributing to the increasing trend. In the case study analysed in this thesis the neodymium production-requirements imbalance was exacerbated by two principal factors in equal measures at the start of the period of analysis (access and environmental constraints) with co-mining contributing to a much lesser extent. However, the decreasing trend in supply disruption potential is driven by increasing diversification of neodymium producing countries and an associated decrease in the access exacerbating factor. The relative importance of different factors, and the associated nature of material criticality constraints, can be tracked over the period of analysis and used to inform a flexible approach to policy intervention and the nature changes.

10.3.2 Infrastructure policy and regulatory constraints

By contrast, the case study analysis used to investigate policy and regulatory constraints provided few concrete insights into the scale of constraints. It could be implied that constraints from fixed and long-term regulatory processes were more severe than those from policy initiatives which tend to change more regularly and can have less extensive effects. However no empirical evidence was collected to support these assumptions about scale. Perhaps the only definite statement that can be made about scale is the observation that the constraints faced in the operation of water infrastructure are more severe and extensive that in the operation of energy infrastructure. This is reflected by the fact that there were very few examples of alternative modes of operation of water infrastructure and these were principally confined to non-domestic customers.

However, the explanatory nature of the case study methodology and analytical framework provided far greater insight into the nature of the constraints. These insights were both theoretical and empirical. Theoretically, constraints were articulated (using evolutionary analogies) as unfavourable selection pressures, which meant that mainstream modes of operation were selected in favour of alternatives. Constraints can also arise from a lack of capabilities to respond to selection pressures. This provided structure for both identification of constraints and also informed recommendations for responses to these constraints.

Empirically, a wide variety of constraints were identified to scaling out (becoming more numerous), scaling up (becoming larger or widening scope) and translating (to other sectors or settings) of alternative modes of infrastructure operation. Many constraints stem from some fundamental differences in the motivations and organisation of mainstream and alternative modes of operation which were described in chapter 7. Infrastructure policy and regulation has evolved around the motivations and organisation of the mainstream mode of operation and; therefore, preferentially selects these modes of operation and locks out alternatives (Unruh 2000). One of the strongest and most unfavourable selection pressures is the concept of value that is central to policy and regulation; this is one-dimensional and is limited to a narrow view of economic value. It is characterised by commodification of water and energy and the promotion of marginal cost efficiency per unit of output in the short-term over all other measures of effectiveness and customer service (Patterson 2008). Alternative modes of operation, which may prioritise other dimensions of value, such as resilience, self-sufficiency or sustainability might produce less economic value using the above definition but produce these other forms of value in far greater quantities than mainstream operation. Other alternative modes may measure the economic efficiency not of each unit of infrastructure product, but of the overall service delivered. However, neither alternative forms of value nor service efficiency are considered in the development of policy, in criteria for funding and support or in the development and enactment of regulation.

Some of the principal regulatory mechanisms, such as the wholesale energy markets of supplier licencing procedures have been created around the capabilities of large, specialist energy companies and are too complex and resource intensive for smaller organisations to compete effectively. Many of these constraints can be traced back to the current pro-market paradigm, which places greater faith in market-based approaches than political processes (Mitchell 2010). It views markets and competition as the most effective way to meet society’s needs. This means that policy that might support those modes of operation is eschewed in favour of policy that mimics markets. Therefore, they continue to be out-competed by the mainstream, despite the fact that they might more effectively meet societal needs in the long-term. It has been argued that the narrow definition of value discussed above originates from the pro-market paradigm. While the current paradigm persists and acts like a ‘band of iron’ around policy and regulation it is likely that the prioritisation of unit cost efficiency and competition over sustainability will persist (Mitchell 2010).

10.4 Potential responses to constraints

One of the principal aims of this thesis was to identify potential responses to the constraints identified during analysis. This means that identification of policy is central to the analytical approach employed. This section summarises key policy insights and assesses the value of the two contrasting approaches to constraint analysis in the identification of policy responses.

10.4.1 Responses to material criticality constraints

The methodology for analysis of material criticality constraints developed in this thesis is purposefully transparent to support analysis of the dominant contributions to constraints and to allow responses to be targeted at the most important contributions. The structure of the indicator, and the ability to decompose results to show the contribution of individual metrics, helped to identify a number of policy interventions to reduce both the potential for disruption and the exposure to that disruption.

Responses to Supply Disruption Potential were primarily focussed on reducing demand for critical materials in the UK through encouraging material efficiency and supporting the development of recycling facilities. Policy responses to exposure material supply disruption centred on the need for diversity in energy technology scenarios. Importantly, this diversity is more than just a function of the number of different technologies in a decarbonisation scenario. For example a scenario where 80% of the goal is delivered by one exposed technology and 20% is delivered by a non-exposed technology is no less exposed that a scenario where the non-exposed 20% is delivered by five different technologies. A balance between the different technologies is needed to reduce the exposure of the goal to individual material criticalities. Additionally, technologies must rely on different materials in order to reduce exposure to individual material criticality constraints. For example, hybrid vehicles and electric vehicles both require neodymium in permanent magnets, a scenario with a balance of these two technologies would do nothing to reduce exposure; a greater disparity of technologies is required, which rely on different materials (Stirling 2007).

Dynamic analysis is particularly important in identifying effective policy responses. The case studies illustrated a dramatic increase in criticality over the period of analysis. If a snapshot of current criticality was taken the situation looks far less concerning. Having foresight of this potential future increase in material criticality constraints prevents reliance on technologies that may be subject to material criticality constraints in the future, creating ‘lock-in’. It also shows the importance of considering the nature of the change in criticality over time – the results showed a steep increase in criticality after 2030, when roll out of direct drive turbines is projected to increase dramatically. It will be more difficult to devise policy responses to such steep changes than to static high levels of criticality.

The results highlight the importance of aligning industrial policy with energy and climate change policy. Energy policy is concerned with reducing constraints to low-carbon technology roll-out; however many of the policy responses to reduce material criticality constraints fall under the remit other departments. This is particularly the case with responses that reduce supply disruption potential and involve using resources more effectively, which would be the responsibility of both BIS and Defra. Furthermore, many of the responses identified call for targeted support, to encourage recycling of specific materials for example, rather than market-based instruments. This is a significant change to previous policy intervention in the UK into this area, which has been limited to addressing specific market failures (Defra 2013). Further research is needed in this area to identify the specific mechanisms to provide this targeted support.

10.4.2 Responses to policy and regulatory constraints

The focus of the analytical framework on evolutionary processes helps to identify responses to constraints from policy and regulation in two ways. Firstly, policy responses can shift selection pressures bearing on the regime to be more favourable to alternative modes of infrastructure operation. Secondly, policy responses can co-ordinate resources available inside and outside the regime to enable alternative modes of operation to respond and adapt to these pressures (Smith et al. 2005). Policy can encourage a shift in selection pressures to ensure they are more coherently aligned with the goals of low-carbon transition and resource efficiency and enable a diversity of modes of operation to co-exist. This shift could be achieved by creating conditions which enable organisations to compete to deliver social and environmental value as well as economic value; ensuring regulation is co-ordinated and clearly articulates the urgent need for climate change mitigation; or creating policy that aims to alter social norms. This requires changes in regulation but also a change in the decision processes behind policy development to ensure that selection pressures are coherent. Responses to shift selection pressures include; encouraging more integrated policy to ensure consistent, cohesive selection pressures; targeted support for alternative modes of infrastructure operation, not just market-based instruments which favour incumbents; alternative approaches to valuation to capture social and environmental value created by alternative modes of operation; and new approaches to regulation that are better aligned with low-carbon goals.

Responding to selection pressures requires certain knowledge, skills, competence and attributes, which are defined in this thesis as capabilities (Sen 1997). Increasing actor and organisational capabilities and encouraging co-ordination are resource and time intensive.

Alternative modes of infrastructure operation will have a great deal of variety in nature and scope of operation and in the motivations and capabilities of actors involved. Some capabilities might benefit from common standards and tools; for example technical standards and procurement processes or decision support tools. However, building participant confidence and connection with networks requires bespoke support that cannot necessarily be done with a toolkit. The increasing trend away from specific support to generic guidance reduces the cost of support for capabilities but must be reversed to enhance and co-ordinate some of the capabilities fundamental to alternative modes of operation. Furthermore stable, flexible funding is required to support experimentation and increase the rate of learning between projects.

10.4.3 Common insights from contrasting constraint analysis

Despite very different approaches to analysis, there are some striking similarities in the potential responses to constraints. The results of both analyses emphasize the importance of diversity in the future infrastructure system, both in technology and in modes of operation. Furthermore, they highlighted the need for not just variety but also balance and disparity in solutions to increase resilience and avoid future lock-in. Technology diversity could be one of the most significant means to reduce material criticality constraints. Conversely, a lack of diversity in modes of infrastructure operation acts as a constraint and new approaches to governance are required to open up a space for diversity. The historical focus of policy and regulation on a small number of monolithic stakeholders will need to evolve to recognise the contribution of different actors and ensuring they respond to the needs of equally diverse end users. This creates further diversity; in the modes of operation that could potentially contribute to infrastructure provision, which policy and regulation will have to deal with. There are means of analysing diversity (Stirling 2010) and calls to ‘open up’ space for technology diversity (Stirling 2008); however, the UK government is wedded to deterministic notions of infrastructure progress, defining ‘pathways’ and has little experience of a more pluralistic approach to policy that encourages diversity whilst moving towards specific goals.

Both analyses resulted in the identification of the need for a more integrated approach to policy and stress the need for concerted action across policy areas. This goes beyond the observation that infrastructure systems are regulated separately; many of the responses are not related to traditional infrastructure governance. For instance waste management policy, traditionally considered by Defra and very separately to infrastructure, needs to encourage recycling of critical materials. With regard to infrastructure operation; engagement with communities and local government is required maximise the potential of infrastructure operation to improve wellbeing, which falls under the remit of the Department for Communities and Local Government.

Many of the policies identified in response to both constraints represent what could be termed targeted support, rather than market-based mechanisms. Both constraints are caused by a complex array of factors and market mechanisms reflect the non-monetary drivers and benefits poorly. Further research is needed to determine the most appropriate mechanisms through which to deliver this targeted support.

The complex nature of both constraints means that there is a great deal of uncertainty associated with the future evolution of the transitions studied and the effect of policy designed to reduce constraints. However, this complexity and uncertainty should not preclude action. Policy and governance processes must become more adaptive so it can respond as system understanding becomes more sophisticated. Neither the necessary changes in the infrastructure system nor the change in governance systems will be instantaneous. The non-linearity and uncertainty associated with infrastructure transitions presents a significant problem for policy development, particularly the top-down, deterministic approach that characterises the current policy paradigm. Governance needs to be more reflexive and adaptive to account for the gradual co-evolution of systems and the unpredictability of systemic change. This means that policy makers must reflect on what is experienced and improve policy learning (Foxon, Pearson, et al. 2005). In parallel to this, a more dynamic and adaptive approach to policy must be established. Research on adaptive governance stresses the importance of creating a strategic vision of the future and committing to short-term actions that create momentum towards this vision (Walker et al. 2013). This must be supported by a framework to guide future actions; so that when policy makers recognise that momentum towards the strategic vision has stalled, or that the vision is no longer relevant a process has been created to realign action with strategies or reframe the strategy (Ranger et al. 2010).

10.5 Understanding low-carbon transitions

The principal focus of this thesis was on the analysis of constraints and the identification of responses to those constraints. However, the final research question had a more reflective aspect which aimed to determine whether analysis intended to examine constraints to a transition could tell us much about the nature of the transition itself and the role of governance in accelerating desirable transitions.

10.5.1 Material criticality constraints and understanding transitions

The approaches to transitions analysis described in chapter 2 tend to analyse socio-technical or socio-ecological transitions. This research has highlighted the importance of physical resources to infrastructure transitions so suggests that both social and ecological systems are important in enabling technology transitions. Exclusion of both of these aspects in the current, predominantly techno-economic, analysis of infrastructure transitions excludes important interactions which are central to the transition process. The analysis in this thesis provides initial insights into the interaction between physical and technology systems but does not tell us how these interactions change over time. A better understanding of the dynamics of the physical system and its interaction with social and technical system would enhance our understanding of low-carbon transitions further.

The policy responses identified in this analysis included interventions directly into infrastructure systems but also into supporting physical systems of resource supply. This reinforces the fact that governance of transitions necessarily involves action across many systems.

10.5.2 Policy and regulatory constraints and understanding transitions

Despite its focus on constraints and responses, the analysis of policy and regulatory constraints provided insights into both the nature of infrastructure transition and also the role that governance might play in accelerating this transition. The majority of the policy responses identified in section 9.1 represent incremental change from the current system of policy and regulation. As a result, these responses are unlikely to break the lock-in to the current mode of infrastructure operation and accelerate the necessary scale and speed of transition. However, the qualitative analysis of relationships and dynamics that led to these responses provided deeper insights into the nature of infrastructure transition. These were valuable in identifying the fundamental change in the system of governance necessary to ensure that the mutually re-enforcing relationship between the current system of governance and the mainstream was broken. Analysis identified the considerable difference in characteristics of mainstream and alternative modes of operation and better reflected the complexity of the infrastructure system which current constrains these alternatives. This resulted in the identification of a conceptual framework to support the analysis of alternative systems of governance that might accelerate transition to alternative modes of infrastructure operation.

The smaller scale and closer engagement with end users, which characterises many alternative modes of operation, have resulted in institutions that are specific to each case but fundamental to their success. This suggests a greater role for self-governance and non-state actors in infrastructure transition. This signals the needs for an expansion of ‘bottom-up’ governance processes to operate in tandem with the more ‘top-down’ policy and regulation. This has two implications for policy; firstly, more support is needed to enable self-governance. This may include guidance or design rules to support the creation of bespoke institutions to manage the distribution of costs and benefits, particularly in partnerships. Secondly, top-down governance processes will have to adapt to enable the institutional diversity necessary to reflect the variety of modes of infrastructure operation and varying character of self-governance processes.

The challenge of deterministic governance is exacerbated by the mismatch between the focus of infrastructure regulation on cost efficiency now, which conflicts with the necessary transition towards a more sustainable infrastructure system (Mitchell and Woodman 2010). The goals of infrastructure policy and regulation need to be realigned to place a greater focus on sustainability and the needs of current and future users. This will have consequences for the definitions of value and the process of value analysis. The definition of value needs to take into account social and environmental benefits and value analysis needs to measure the effects of systemic change, not just appraise individual changes.

The conceptual framework developed in this thesis could be used to create a long-term vision for infrastructure governance but is currently disconnected to the short-term actions described in section 9.1. Articulation of processes to make this connection and determine how governance could accelerate change is perhaps where most potential exists to further develop the work in this thesis. The short-term responses outlined in section 9.1 identify the actions to realign current regulation and governance and policy with the goal of sustainable infrastructure. Section 9.2 describes a future infrastructure system and the role that governance might play in its supporting it. This can be used to identify what is lacking in current infrastructure governance systems but does not help to describe the pathway or pathways between the two.

The results highlighted the importance of features in the landscape, and in particular the policy paradigm, in creating constraints and limiting the potential responses to these constraints. Without a fundamental change in the policy paradigm, it seems unlikely that a low-carbon infrastructure transition is possible at the scale and speed necessary to avoid dangerous climate change. Analysis has uncovered evidence that the new paradigms are emerging to exist alongside the historic pro-market paradigm. However, this coexistence of paradigms can lead to contradictory pressures on policy makers and incoherence in the resulting policy and institutions (Kern et al. 2014). This lack of coherence in paradigm is fundamental to infrastructure transitions and must be addressed to release the potential of governance to acceleration transitions.

10.6 Limitations of analysis and methodological developments

The approaches described above and the results of analysis have a series of limitations in addressing the research questions in this thesis. These are discussed in more detail in this section along with methodological developments that might address these limitations to some extent.

10.6.1 Limitations of method for analysis of material criticality constraints

The limitations of composite indicators more generally are discussed in detail in Chapter 4; therefore this section summarizes the specific limitations associated with the indicators developed in this thesis.

A number of factors limit the ability of the indicators to comprehensively represent material criticality constraints. Some of the metrics are not entirely independent (for example the HHI is used to partly weight the ECR) and additive combination of indicators assumes this to be the case. Some metrics which might contribute to our understanding of criticality, such as political stability in mineral producing countries, have been excluded because robust data is not available or they are not widely supported as reliable indicators.

The assessment of supply disruption potential only considers disruption at the point of production. However, the supply chain of critical materials has additional down-stream stages where disruption may occur, which are excluded from the current analysis. This has been explored qualitatively (IEA-RETD 2012) but a quantitative analysis will require significant further data gathering, investigation and analysis.

A number of the indicators are forecast in order to demonstrate that a more dynamic analysis of criticality can potentially be useful in providing policy makers with information to reduce the probability of ‘locking-in’ to currently attractive but potentially future-critical technologies. It has been necessary to make a series of assumptions to support this forecasting that are of course contestable. The method used to forecast production relies on the continuation of historic trends to forecast future change. Future production will in fact be driven and constrained by a multitude of interrelated factors, many of which cannot be predicted (e.g. disruptive new technologies). Nonetheless, forecasts are considered to be appropriate for their intended purpose of providing an indication of the potential for an imbalance between production and requirements in the future.

The indicators developed are dynamic, to the extent that individual metrics change over time; however, this does not take into account the interdependencies between indicators, particularly the relations between production and requirements. For example, policy developed to reduce exposure (by encouraging the reduction in material requirements per unit of technology) will also affect supply disruption potential (by reducing global requirements for material and mitigating production-requirements imbalance). Furthermore, policy analysis is static and external to the analysis of material criticality constraints so it is not possible to track inadvertent negative effects of policy interventions. For example, the substitution of a critical metal for another (to reduce exposure to the critical metal) could affect the production-requirements imbalance of the replacement material and create material criticality constraints from a different material. The analysis would benefit from better integration of actors allowing these dynamic interactions of decisions and material criticality constraints (Knoeri et al. 2013).

No discussion of a threshold of criticality has been included, because the indicators are intended to be used to compare the relative criticality of different pathways to low-carbon infrastructure, rather than defining the point at which criticality becomes unacceptable. This is not to say that it is not possible to define a threshold of this nature, but rather that it is not the intention of this thesis; criticality thresholds will need to be informed by a combination of political and economic factors as well as a technical analysis of criticality.

10.6.2 Limitations of method for analysis of policy and regulatory constraints

The limitations of case study research in general are discussed in chapter 4. This section considers the specific limitations associated with the analytical framework and the case studies analysed in this thesis.

The conceptualisation of policy and regulatory constraints as a predominantly institutional constraint focuses analysis on interactions between institutions and individuals. This provided a great deal of insight into how these interactions enable constrained organisations to operate in an alternative way. However, these alternative modes of operation also have a significant influence over infrastructure technology, which was excluded from analysis. The alternative providers analysed in this thesis relied on more distributed technologies and placed greater importance on demand management. These developments could have significant implications for the current system of technologies and connecting infrastructure, which could constrain the roll out of alternative modes of operation, but are not considered in this thesis.

The qualitative nature of analysis of regulatory constraints does not allow for assessment of the relative severity of different constraints, which could make it difficult to prioritise efforts to reduce constraints. Moreover, the constraints are felt to different extents by different actors and modes of operation; a point which is lost to some extent through the process of theory building and generalisation. In this way, the richness of case study findings is diminished by the need to generalise findings.

The individual cases were deliberately selected to incorporate a wide range of participants in infrastructure operation to ensure that the potential diversity was fully represented. This diversity made generalisation of results challenging. This was particularly difficult for comparing findings of case studies in different regulatory environments, such as water and energy; however, exploration of different regulatory environments gave a better impression of the underlying policy paradigm.

The analytical framework focuses its attention on the evolution of novelty and the specific constraints presented by lack of resources or unfavourable selection environments. However, it doesn't adequately account for the processes and structures that define and reinforce everyday life (referred to as social practices) which may also constrain this evolution (Shove & Walker 2010; Hargreaves, Longhurst, et al. 2013). Shove and Pantzar describe practices as being made up of 'images' (meanings, symbols), 'skills' (know-how, forms of competence), and 'materials' (artefacts, technologies) that are actively and iteratively integrated through circuits of reproduction (Shove & Pantzar 2005). This integration and reinforcement of different practice elements is poorly reflected in the current analytical framework.

The analytical framework supported the identification of necessary changes to policy and regulation in the short-term and new roles for government and regulators in the long-term. However, it does not help identify a pathway or pathways of change towards the long term vision, nor how to deal with the uncertainty associated with these pathways. Additional work is needed to establish a framework to guide future action to ensure short-term action contributes to long-term change and allows for dynamic adaptation over time to meet changing circumstances.

10.6.3 Methodological developments

The limitations above suggest some improvements that could be made to the methods used in this thesis. These are principally limited to the choices made with regard to the approach to systemic analysis and complexity, described in section 10.2.3. The inclusion of actors and their interactions in analysis of material criticality constraints has been discussed in chapter 5 as a means to improve the representation of complexity. It has been suggested that this could be done by incorporating agent-based modelling into the analytical approach (Knoeri et al. 2013). This would address the absence of relationship-based dynamics in the temporally dynamic analysis and better represent the interactions of social, physical and technology systems. This might also enable analysis of the interactions between indicators, which would address the dependence of indicators and might help strengthen assumptions used in forecasts of indicators.

The reverse is true for analysis of policy and regulatory constraints, which would benefit from more considered analysis of the implications of alternative modes of infrastructure operation and the policy responses identified in this thesis for the infrastructure technology system. Changes in operation and governance will require an associated change in the supporting technology and connecting infrastructure. This would imply that some form of complexity analysis or modelling is necessary, which is capable of improving our understanding of interaction between institutions and the technology system.

Both developments imply that increasing the complexity of analysis is beneficial to systemic understanding and identification of relevant responses. This is supported by Habermas' work, which argues that both causal explanation and interpretation of social interactions are needed in order to create the necessary system change (emancipation) (Habermas 1972). Approaches exist which can bring together understanding of the technological, natural and social system, drawing from systems analysis and complexity science. However, it is important that the sophistication of analysis is balanced with the purpose of the analysis; to enable change.

11 Conclusions

The central question posed at the outset of this thesis asked: *how can constraints to low carbon infrastructure transition be characterised with sufficient detail to enable action to mitigate disruption?* The different analytical approaches applied in this thesis revealed constraints from critical material supply disruption and from infrastructure policy and regulation that were very different in nature. Despite this, some common insights emerged relating to the methods used and the potential policy responses the constraints identified. This chapter draws together the key findings of this analysis relating to the conceptualisation of constraints, methods of analysis, nature of constraints and policy responses. It summarizes the contribution that this thesis has made to specific fields of academic literature and describes future work that might build on this contribution.

11.1 Key findings

Many of the conclusions from this research were found to be independent of specific constraints, despite the selection of two very different constraints; one principally physical and one principally institutional. The key findings of the research are presented below as statements followed by a more detailed discussion of the significance of the finding. The majority of the key findings relate to analysis of constraints in general, however, and perhaps unsurprisingly, findings relating to the nature of the constraint varied significantly. In these cases the statement includes mention of the specific constraint to which it applies.

The conceptualisation of the constraint (as a physical or institutional one) influences the methods used for analysis, which in turn influence the type of knowledge created

Material criticality was conceptualised as a predominantly physical constraint caused by disruption in supply of critical materials. As a result, the quantitative, indicator-based method of analysis was concerned with events and processes which could explain and predict this disruption and its effects on low-carbon infrastructure transition. The knowledge generated was empirical in nature and policy responses that resulted from this form of analysis were concerned with controlling this disruption. There was no consideration of human interaction with the physical system, which limited system understanding.

Policy and regulation was conceptualised as an institutional constraint; processes of policy and regulation were thought to affect the evolution of alternative modes of operation. This placed the focus of analysis more squarely on understand the interactions between individuals and organisations and resulted in a more qualitative approach to analysis. This focus on social interactions resulted in a greater understanding of a broader part of the infrastructure system. The policy responses identified as a result of this analysis covered a correspondingly broader scope and called for systemic change. The analysis of policy paradigms provided some insight into the power relationships underpinning infrastructure policy and regulation. This improved both the system understanding and enabled a more effective analysis of the potential for change.

Dynamic analysis is important in the analysis of constraints to transitions

Transitions are dynamic in two senses; the transition occurs over an extended period of time; and the transition occurs as a result of the interaction of actors and assets. As a result, analysis of constraints must be able to take into account these dimensions of dynamic analysis. The analysis of material criticality constraints was able to undertake temporally dynamic analysis but excluded interaction of actors from the analysis. As such, knowledge was limited to prediction of criticality, with very little insight to help understand the issue. This is one of the principal weaknesses of the analytical approach employed for examination of material criticality constraints and one of the key areas for future development. However, the temporally dynamic approach supported analysis of the evolution of criticality over time and may help to prevent reliance on technologies that could become highly critical in the future. In this way, the approach helps decision makers to account for some elements of uncertainty and complexity and make more robust decisions than they would in the absence of temporally dynamic analysis.

The more qualitative approach used to analyse constraints from infrastructure policy and regulation focussed more specifically on the second dimension of dynamic analysis; the relationships and interactions between actors and between actors and assets. This yielded greater insights into the process of transition and better reflected the complexity of the infrastructure system. The emergence of alternatives through a process of interactions illuminated a more varied role for governance in not just encouraging specific modes of operation but in creating selection pressures that would enable a diversity of modes of operation.

Constraints from disruption in the supply of critical material are likely to increase dramatically in even the most conservative scenarios of electricity system decarbonisation

Material criticality constraints in the case study analysis of neodymium and wind turbines increased over the period from 2012 to 2050. The roll out of wind turbines in the Core

Pathway causes an increase in constraints of four-fold over the period, with a step-change occurring in 2030, as shown with reference to 2012 values. This trend is even more dramatic in the Renewables scenario with an almost ten-fold increase. Despite the significant increase in criticality, the Supply Disruption Potential of neodymium (normalized to iron) decreases over the period of analysis by nearly 30%. This reduction is due to increasing competition in the supply of neodymium, as a result of new mining activities outside China.

The reducing Supply Disruption Potential indicates that the overall trend in criticality of UK electricity system decarbonisation is driven strongly by the UK's increasing exposure to supply disruption and that intervention would be most effective if focused on reducing this exposure.

Many constraints from policy and regulation arise because governance systems have evolved around the mainstream mode of operation and are locked-in to this mode of operation

Infrastructure governance systems have evolved with the mainstream mode of operation. As a result, policy and regulation have created selection pressures that favour mainstream operators. Alternative modes of infrastructure operation, which have very different characteristics to the mainstream, are locked out. For example, the wholesale market was created to avoid monopoly price control but is complex and presents high transaction costs and risks for small operators. This means that larger mainstream operators can out-compete alternatives that don't have capabilities to interact with such complex systems.

The commodification of infrastructure actively deters modes of operation which prioritise non-monetary value or measure efficiency of a service, rather than unit of infrastructure product. The promotion of marginal cost efficiency per unit of output over the short-term over all other measures of effectiveness and customer service means that alternative modes of operation are further limited in their ability to compete with the mainstream.

Diversity is a good way to mitigate current constraints and avoid future lock-in

Diversifying the technology or modes of operation contributing to low carbon transition could reduce exposure to material criticality constraints and enable more favourable selection pressures, reducing constraints from infrastructure policy and regulation. Furthermore, diversity can mitigate future lock-in, hedges ignorance and also offers a means to promote innovation (Stirling 2007). A more diverse portfolio of technologies or modes of operation means that the infrastructure system could be more resilient because a smaller part of the system is affected if a particular technology or mode of operation becomes obsolete.

Diversity, and the ensuing resilience under shock and robustness under stress, is more than just the presence of many technologies or modes of operation: *"diversity is generally a state under which an observed system is seen to display: (1)* ***even balance*** *across (2) a* ***variety*** *of (3)* ***mutually disparate*** *categories"* (Stirling 2011). In the case of constraints from critical materials, mutually disparate categories could be taken to be technologies with similar functions (low-carbon energy generation, for example) which do not rely on the same critical material. A corresponding definition of diversity in modes of infrastructure operation would be modes of operation by different instigators and with different motives.

Responses need to be co-ordinated and integrated

The tendency of the current, fragmented policy and regulatory system to focus on individual goals, infrastructures and project stages and organisations in isolation can hinder effective policy responses. The responses to constraints described in this thesis require co-ordination and integration of policy across government departments and across the infrastructure lifecycle. Responsibility for mitigating material criticality is shared between Defra and BIS, who are responsible for waste management and business resilience. To some extent, there is a precedent for this cross-department action. However, in order to address material criticality systematically, this remit must be expanded to include DECC and Treasury to address incentives which affect technology selection.

Selection pressures in infrastructure operation are often conflicting and need to be made more coherent and better aligned with the goals of sustainable development. For example, Water companies are required by Ofwat to demonstrate water efficiency measures but in parallel to this, Ofwat regulates the price of each unit of water. In this situation demand reduction will reduce water company profits, and the regulation of prices is poorly aligned with the goal of demand management. The government is required to steer more deliberately away from unsustainable pathways by directing incentives and interests towards more innovative and sustainable modes of operation (Leach et al. 2012). Greater harmonisation is required to increase the effectiveness of policy. There is a particular need for better integration of innovation, industrial and infrastructure policy (Foxon, Pearson, et al. 2005; Mitchell 2010).

Policy needs to recognise the interconnectedness of infrastructure systems and enable the exploitation of the cost and resource efficiencies associated with a close integration of the operation of infrastructure systems. To reduce this complexity it will be necessary to introduce formal systems to increase integration and cooperation between infrastructure regulators.

Policy needs to be targeted – the market will not fix criticality in the short-term and alternative modes of operation are out-competed

It is argued that infrastructure transition should be left to the market and that market-based instruments will encourage material substitution (avoiding material criticality) and drive adoption of efficient business models. However, neutral policy favours large, incumbent operators and is less supportive of alternative modes of operation and smaller operators, which might undertake more innovative activities. It could be argued that market-based instruments could slow innovation because incumbent companies are likely to concentrate on doing the same things more cheaply, in order to stay competitive, rather than making radical innovations in modes of operation (Mitchell & Woodman 2010). Furthermore, substitution of critical materials is complex and challenging and market signals may not reduce Supply Disruption Potential or exposure at the necessary scale.

There is an increasing body of evidence supporting a more targeted approach to policy that can reduce the risk and uncertainty associated with low-carbon investment and drive innovation (Gross et al. 2012). These findings relate to targeted policy for renewable technologies, but there is an increase in targeted support for alternative modes of infrastructure operation, such as the inclusion of energy performance contracts and ESCos in the Energy Efficiency Directive (European Commission 2012a). This support is increasing the prevalence and effectiveness of ESCos in those countries which offer such support (Marino et al. 2010). There has been less research into targeted support for materials criticality constraints. The EU Raw Material Initiative has identified targeted support through trade initiatives at a European level (European Commission 2011a), however, it is less clear what specific mechanisms are needed in the UK to provide this targeted support.

A new approach to valuation is needed in infrastructure operation that captures the long-term value of system change and captures social and environmental value

One of the most counter-productive selection pressures of the current policy and regulation regime is the measurement of value for money based on the marginal cost efficiency of delivering each unit of utility product. This metric focus prioritises marginal efficiency over total efficiency: for example, in most performance contracting schemes, the user would pay more per unit energy, for fewer energy units, making the whole transaction cheaper. But these types of total contract cost savings are disincentivised by narrow focus on marginal cost efficiency. As a result, a wide range of benefits of alternative modes of operation are not included in a standard cost benefit analysis. These include protecting future customers as well as the value of non-monetary benefits such as mitigation of climate change, resilience and fuel poverty reduction. This could mean that projects are uncompetitive and fall by the wayside because they can’t access finance or that they are delivered in a different way that maximises revenue, rather than wider benefits.

A more integrated approach to valuation of infrastructure operation is needed to value the non-monetary benefits of schemes, so social benefit generated by more local schemes is captured and assessed on a more equal footing with financial benefit. There is a need to balance the goals of economic, environmental and social sustainability, whilst taking into account both price and non-price issues (Mitchell & Woodman 2010). This approach could build on existing methods, such as Social Accounting but should not monetise social and environmental value; a more pluralistic approach to value is needed.

Iterative change in policy and regulation will not be enough; radical change in the means of governance is necessary to overcome policy and regulatory constraints.

Current policy and regulation evolved with and supports the current mode of operation; therefore it could be contributing to path dependency and lock-in. Incremental change may be insufficient to break this lock-in (Unruh 2000). The alternative modes of infrastructure operation analysed in this thesis show characteristics which differ greatly from those of the mainstream mode of operation, which is dominated by large, centralised utility companies and excludes end users and local actors. Furthermore, the complexity of the infrastructure systems is such that the outcomes of isolated interventions are difficult to forecast. Feedback loops, time and spatial lags and path dependency lead to unstable and unpredictable responses. Without a fundamental change in the system of governance system, beyond incremental policy change, it seems unlikely that transition in infrastructure operation of the necessary scale would occur.

This thesis presents an initial conceptual framework which identifies characteristics of a governance system that is more supportive of a diversity of alternative modes of infrastructure operation. This is used to identify important deficiencies of the system but does not identify how to correct these deficiencies. These deficiencies include lack of support for diversity, lack of adaptability of policy and lack of support for self-governance and partnerships. Each part of the framework needs to be analysed in more detail along with pathways to correct systemic deficiencies.

Self-governance of infrastructure operation will become increasingly important but is poorly supported by the current system of policy and regulation

The alternative modes of infrastructure operation examined in this thesis worked at a far smaller scale and were far less uniform in structure and organisation than the incumbent mode of operation. Furthermore, many modes of operation, such as those that deliver infrastructure services, require a closer relationship between the provider and the end-user. The nature of governance arrangements necessary to manage the distribution of costs and benefits varied significantly between cases. This would imply that a monolithic, entirely top-down approach to regulation, which closely controls the relationship between providers and users, would be inappropriate for these modes of operation. This signals the need for a greater role for bottom up, self-governance, which is a step change for infrastructure regulation.

There is a long history of successful self-governance of natural resources and a significant body of literature which examines the institutions necessary to enable self-governance. This literature has a great deal of potential to inform necessary changes to national policy and regulation to better support self-governance. This includes the creation of infrastructure policy to support development of institutions for self-governance of infrastructure and the adaptation of regulation to enable institutional diversity.

The pro-market policy paradigm constrains alternative modes of operation in many ways but could also prevent effective implementation of responses

Policy processes take place within a policy paradigm – a framework of ideas and standards which shape policy goals, instruments and setting of instruments. If the policy paradigm is not supportive this creates constraints and prevents the implementation of the specific proposals and wider system reform suggested in this thesis. The constraints on scaling out or scaling up alternative modes of infrastructure operation can predominantly be traced back to the current pro-market policy paradigm. Paradigms affect the goals of policy (cost efficiency), the kind of instruments that are used to attain them (market-based) but also the very nature of the problems they are meant to be addressing (infrastructure as a free market) (Hall 1993). If the paradigm is not aligned with the challenge of a sustainable infrastructure system, it hinders the ability of policy and regulation to enable transition towards a more sustainable infrastructure system.

There is evidence of change in the paradigm to incorporate new perspectives, goals, instruments and institutions which are more supportive of the proposals set out in this chapter

(Kern et al. 2014). However, the persistence of the pro-market perspective alongside the emerging perspective of sustainability limits the coherence of resulting policy and institutions. This allows economic goals to dominate as a result of hegemony of economic regulation and Treasury priorities. This favours large, centralised, private actors and limits the potential for both diversity in modes of infrastructure operation and self-governance of infrastructure by communities and local authorities. The rigid focus on short-term cost effectiveness limits the ability to foresee alternative paths to a low carbon infrastructure system and hinders adaptive policy making.

11.2 Contribution to literature

This thesis is necessarily interdisciplinary, but has made contributions to a number of specific fields of research. These contributions are summarised in the following sections.

11.2.1 Industrial Ecology

Industrial ecology is the study of material and energy flows through societies' systems. It aims to replicate the propensity of natural systems to re-use materials and employ closed loop cycling and reduce the impact of material use on the natural environment. There is a nascent body of work exploring the use of critical materials to which this thesis makes a contribution. The indicator set developed in chapter 5 contributes to literature on criticality assessment by improving the conceptualisation of criticality, providing more quantitative analysis, externalising adaptive capacity, explicitly including dynamic metrics and supporting analysis of a technology portfolio. The systemic nature and goal orientation of this approach enables analysis of the relative risk to different pathways to achieve the goal of low carbon transition. This is a significant advancement on current approaches, which separately analyse the criticality of an economy, company or technology, and underemphasize the systemic nature of criticality.

11.2.2 Institutional Economics

Institutional economics focuses on understanding the role of human-made institutions in shaping economic behaviour. This thesis makes a particular contribution, adding to the work of Ostrom on institutional dynamics in the management of socio-ecological systems. Infrastructure represents a crucial interface between social and ecological systems influencing the level and composition of society's resource demand and the institutional and social organisation of society. Despite this, infrastructure governance is rarely addressed in most institutional analyses of socio-ecological systems. There has been a great deal of work of the role of rural communities in governing natural resources, but less on community governance of infrastructure. The research in this thesis has taken the first steps towards exploring the evolution of institutions for collective action on infrastructure management. It has made an initial attempt to extend Ostrom's design principles for self-governance of natural resources for application to management of infrastructure by resource users, not just management of resources directly.

11.2.3 Transitions Studies

The analytical framework developed in this thesis offers a number of enhancements that improve the application of current socio-technical and co-evolutionary theories to transitions. These include; a clearer articulation of the connection between niches, regimes and the landscape; the exploration of processes that enable niches to co-exist with, not just overthrow, regimes; a more differentiated analysis of niche 'up-scaling'; and a focus on relationships and dynamic processes to better represent the infrastructure system as a complex adaptive system. It applies these to transition in modes of infrastructure operation and perhaps uniquely, it considers transition to a portfolio of modes, rather than focusing on a single mode of operation. The benefit of this focus is the creation of diversity as a means of reducing future lock-in.

The results of analysis show the importance of including the policy paradigm, at the landscape level, in the analysis of constraints. The paradigm is the source of many constraints to transitions in mode of operation, and could also reduce the effectiveness of potential policy responses.

11.3 Future research

The methods used in this thesis took a systemic approach to analysis but often sacrificed detail and appreciation of complexity in favour of systemic analysis. This resulted in a number of limitations which could be addressed by methodological developments. The analysis in this thesis has also identified a number of emerging findings that would benefit from more detailed analysis for future research. Both are described below.

11.3.1 Adding complexity to analysis – bringing institutions, technology and resources together (without losing focus on action)

One of the most significant weaknesses of the approach to analysis of material criticality constraints was the exclusion of actors from the system of analysis, which falsely separated the social and natural systems. The method described in this thesis could be extended to incorporate actors, their interactions and decisions into analysis of criticality. Knoeri et al propose a framework that couples an agent-based behaviour model with a dynamic material flow model (Knoeri et al. 2013). This enables the analysis of the effects of decisions (for example to substitute a critical material for a less critical material) on the production of or requirements for a material. This could help to account for the interdependencies between indicators and between decisions and criticality. It would also enable analysis of the effects of policy interventions and institutional change on material criticality more effectively than is possible with the current approach. Agent-based models are able to better reflect society-nature interactions and thus the complexity inherent in the system. However, it is possible to start with a low level of resolution to identify the most significant systemic interactions then iteratively add detail and complexity to manage the balance between dealing with complexity and enabling action.

The institutional focus of analysis of policy and governance constraints overlooked the effect that alternative modes of infrastructure operation would have on infrastructure technologies and the connecting network infrastructure. Many modes of operation examined in this thesis relied on more distributed technologies and on demand management, which could have significant implications for the operation and reliability of current technologies. This issue would benefit from more integrated analysis of the co-evolution of institutions and technology. The integration of actor interaction and energy models has been explored qualitatively (Hughes 2013) but would benefit from more quantitative approaches, again using agent-based models.

11.3.2 Adaptive, dynamic decision making – pathways, not policies

The complexity of infrastructure systems makes it hard to imagine how changes in the short-term will lead to the necessary system reconfiguration in the long-term. This thesis has described short-term actions and some of the characteristics of a future system of governance; however it has not gone as far as to describe what that future system might be, nor how short-term action is connected to the necessary long-term change. Regulatory and policy reform must be planned to move towards this alternative governance system, without knowing what the alternative might be. This challenge is exacerbated by the fact that the speed of technological change is moderated by society, politics and the economy; the desirable technology scenarios we are adept at describing won’t happen without associated social, political and economic change. If action in the short-term is ineffective in the face of these moderating factors it may be necessary to shift to more a more radical approach to intervention in order to accelerate system change and avoid dangerous climate change.

Tools or approaches are needed to support and inform decisions in the face of real uncertainty about the nature of a reconfigured system, how parts of the system might change and about how effective interventions might be. There are tools and approaches which aim to achieve a balance between the understanding of uncertainty and system complexity and the articulation of policy pathways that are both adaptive and dynamic (Walker, Haasnoot, & Kwakkel, 2013). Such approaches are predominantly used in water resource management and adaptation to a changing climate. A good example of this is the adaptive planning approach used for Thames Estuary 2100 project (Reeder & Ranger 2010), which provides an adaptive routemap of options for long-lived infrastructure with high sunk-costs under high uncertainty. Further work is needed to explore the application of adaptive and dynamic decision making approaches to infrastructure system reconfiguration.

11.3.3 Valuing (not monetising) non-monetary benefits – un-commodification of infrastructure

Many alternative modes of infrastructure operation have the potential to reduce GHG emissions and resource consumption, improve quality of life and deliver social value such as skills development or poverty alleviation. However, it’s very hard to robustly quantify this social and environmental value and use it to support the case for system change. There is strong support in national and local government for innovations that deliver social and environmental, as well as economic benefit; however, assessment of the case for support is most often limited to economic criteria. This can mean that many promising projects are out-competed by those that focus on maximisation of economic value.

There has been a great deal of research on the valuation of environmental impacts and benefits but less on the valuation of social benefits. Accounting methods exist that enable organisations to honestly report on their social impacts, including Social Accounting (SA) and Social Return on Investment (SROI) (Gibbon & Dey 2011). However, these methods tend to report on existing activities, rather than develop evidence to support the case for system change. Further work is needed to adapt existing social accounting methods to enable the inclusion of social criteria in decision making processes; especially in areas of policy which might support infrastructure operation. This will also help alternative providers to assess potential social impact and identify how innovations could be adapted to maximise social benefit.

Furthermore , new approaches are needed that value investment in the reliability of coupled, complex systems, the public good generated by infrastructure, to enable generation of revenue from means other than direct user charges. This will help to move infrastructure away from commodification towards an approach based on systemic investment.

11.3.4 Self-governance and partnerships – beyond public-private-partnerships

Further work is needed to develop the first steps taken in this thesis to apply existing work on institutions for self-governance of natural resources to self-governance of infrastructure. This could explore the benefits of self-governance of infrastructure for society (including the infrastructure users) and ecological systems (including preservation of energy resources). The initial attempt in this thesis to develop design principles for self-governance of infrastructure by resource users is based on a limited sample of cases and the focus of analysis was on the constraints faced by the cases, not on self-governance. The emerging principles could be refined by undertaking detailed empirical analysis of successful cases of community governance of infrastructure. This analysis should also be supported by exploration of the technical, social and environmental implications of a more active role for communities in infrastructure governance. The assumption in this thesis that more control by communities is better for social and environmental outcomes is based on the observed outcomes of a small number of cases and there are some who argue that community management of infrastructure would be a burden, not a benefit. The implications of community management should be explored in more detail to examine this argument in more detail.

11.4 Final conclusions

The principal aim of this thesis was to characterise two contrasting constraints to infrastructure transitions *with sufficient detail to enable action to mitigate* these constraints. This aim had a number of implications for the current and future research, which are discussed in this section. Firstly, the dual focus on characterising constraints and enabling action forces a balance between trying to understand a phenomenon and trying to identify how it might be changed. As a result, policy responses were explicitly addressed in one of the research questions, and carried equal weight in the results chapters, rather than being confined to a short discussion of policy relevance. This affected the approaches to research significantly and it could be argued that this increased the policy relevance of the research.

The focus on *enabling action* forced a more systemic and dynamic approach to analysis. Change in complex systems, such as infrastructure, emerges as a result of system interactions rather than from the characteristics of particular parts of the system. Therefore; whole system analysis can provide more useful insights into how the wider system might be enabled to change than detailed analysis of one part of the system. Consequently, a choice had to be made about the level of detail and complexity that could be addressed within the constraints of this thesis. Analysis was limited to examining infrastructure systems at a high level and detail was added where this would support identification of specific policy recommendations.

Equally weighting the importance of policy responses also encouraged the examination of the dynamics of constraints, and in particular the interaction between different parts of the infrastructure system. The resulting methodological advancements are an important contribution of this thesis. Moreover, dynamic analysis not only helped to identify the causes of constraints but also revealed some of the mechanisms by which these constraints could be reduced, increasing its relevance.

Nevertheless, this policy-driven analysis was not without its challenges. In a number of places detail was sacrificed in favour of taking a more systemic approach to analysis. Consequently, the results are subject to a series of limitations which increase the uncertainty associated with both the understanding of constraints and the effectiveness of potential responses. This uncertainty can make it harder to defend results, which could be considered to be a high risk strategy for a PhD thesis. However, intervention into a complex system, such as infrastructure, carries with it inherent uncertainty. The high-level systemic analysis applied in this thesis could be a useful way to acknowledge and manage this uncertainty. In effect, it acts as a sensitivity analysis, identifying the areas of the system where intervention could be most effective and where there is most uncertainty. In this way, it can be used to prioritise action in terms of both policy intervention and more detailed research. As such, it forms a useful basis for a more adaptive approach to policy; with initial work identifying important and urgent areas to intervene. The mechanism for intervention can then be adapted as system understanding develops and the response to initial policy becomes clearer. In this way, action is not precluded until the system is understood in detail; moreover, it could raise actors' awareness of the potentials wider implications of individual actions.

The choice of two contrasting constraints required that both quantitative and qualitative analysis was used. This multi-methods approach provided a useful chance to reflect on the relative advantages of different methodological approaches and, in particular, their relevance to policy analysis. The quantitative work provided a great deal of comfort; it generated absolute results that are easier to interpret and communicate and can instil the user with confidence when prioritising action. Furthermore, it is easier to see how quantitative data can be integrated with existing analysis of infrastructure systems, much of which is based around energy system modelling. However, quantification is necessarily reductive and can give a false feeling of certainty in results. By contrast, the qualitative work provided more discursive and less definite feeling results but led to a better system understanding and a more sophisticated appreciation of the complexity of the infrastructure system. Perhaps one of the most significant challenges to application of this work will be integrating the qualitative data into existing work on infrastructure transitions. Contrasting methods did make comparative analysis quite challenging so perhaps some of the cross-case lessons have been lost because different forms of knowledge were produced and different system boundaries used.

Finally, the topic examined in this thesis is an example of *problem-oriented interdisciplinary* in the sense that it was driven by a complex issue that cannot be assigned to a given discipline (Froderman et al. 2010). This means that it has drawn on and attempted to integrate diverse range of literatures, which has enabled to a great degree the systemic approach central to the thesis. Despite this aim, the research has not been able to take a true interdisciplinary approach to analysis; analysis of material criticality constraints underplayed social interactions and policy and regulatory constraints underplayed the role of technology. The research in this thesis has found that effective analysis of constraints to infrastructure transitions must reflect the social, environmental and technical aspects of transitions together and particularly the interactions between systems. Furthermore, interdisciplinarity itself is not without its pitfalls; it can be more time consuming and frustrating (Longhurst & Chilvers 2012) and research disciplines are important foundations for academic careers (demonstrated in the recent Research Excellence Framework assessment and journal ranking system (Rafols et al. 2012)). Despite this, the author feels that interdisciplinary research has a greater potential to address complex challenges, such as responding to climate change.

References

AEA Technology, 2011. *Pathways to 2050 - Key Results*,

AEA Technology, 2010. *Review of the Future Resource Risks Faced by UK Business and an Assessment of Future Viability. A report for the Department for Environment, Food and Rural Affairs*, London.

Allwood, J. & Cullen, J., 2012. *Sustainable materials with both eyes open*, UIT Cambridge.

Alonso, E., Field, F.R. & Kirchain, R.E., 2008. A dynamic simulation model for examining the implications of limited materials availability : The case of platinum. In *16th MIT-UAlbany-WPI System Dynamics Ph.D. Colloquium*.

Alonso, E. et al., 2012. Evaluating Rare Earth Element Availability: A Case with Revolutionary Demand from Clean Technologies. *Environmental Science & Technology*, 46, pp.3406-3414.

Anandarajah, G. et al., 2009. *Pathways to a Low Carbon Economy: Energy Systems Modelling*,

Anderson, K. & Bows, A., 2012. A new paradigm for climate change. *Nature Climate Change*, 2(9), pp.639-640.

Andersson, B.A., 2001a. *Material Constraints on Technology Evolution : The Case of Scarce Metals and Emerging Energy Technologies*. Chalmers University of Technology and Goteborg University.

Andersson, B.A, 2001b. Metal resource constraints for electric-vehicle batteries. *Transportation Research Part D: Transport and Environment*, 6(5), pp.297-324.

Andersson, B.A., 2000. Materials availability for large-scale thin-film photovoltaics. *Progress in Photovoltaics*, 8, pp.61-76.

Angerer, G. et al., 2009. *Raw materials for emerging technologies (English summary)*,

Arthur, W.B., 1989. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *The Economic Journal*, 99(394), pp.116-131.

Arup, 2011. *Decentralised Energy Masterplanning A manual for local authorities*.

Awerbuch, S., 2005. The Role of Wind in Enhancing UK Energy Diversity and Security.

Ayres, R.U. & Peiró, L.T., 2013. Material efficiency: rare and critical metals. *Philosophical Transactions of the Royal Society A*, 371.

BGS, 2011. *Rare Earth Elements*.

BIS, 2011. *Principles for Economic Regulation*.

Baer, P. et al., 2008. *The Greenhouse Development Rights Framework* Second Edi., Berlin: Heinrich Bool Foundation.

Baker, P.E., Mitchell, C. & Woodman, B., 2010. *Electricity Market Design for a Low-carbon Future*.

Bakker, K., 2005. Neoliberalising Nature? Market Environmentalism in water supply in England and Wales. *Annals of the Association of Americal Geographers*, 95(3), pp.542-565.

Bale, C.S.E. et al., 2012. Strategic energy planning within local authorities in the UK: A study of the city of Leeds. *Energy Policy*, 48, pp.242-251.

Balica, S.F., Wright, N.G. & van der Meulen, F., 2012. A flood vulnerability index for coastal cities and its use in assessing climate change impacts. *Natural Hazards*, 64, pp.73-105.

Barrett, J. & Scott, K., 2012. Link between climate change mitigation and resource efficiency: A UK case study. *Global Environmental Change*, 22, pp.229-307.

Beadle, R., 2014. Interview about Byker Community Trust and Byker heat network.

Beinhocker, E., 2006. *The origin of wealth: Evolution, complexity, and the radical remaking of economics*, Random House.

Bellaby, P., 2010. Uncertainties and risks in transitions to sustainable energy, and the part “trust” might play in managing them: a comparison with the current pension crisis. *Energy Policy*, 38(6), pp.2624-2630.

van den Bergh, J.C.J.M., Truffer, B. & Kallis, G., 2011. Environmental innovation and societal transitions: Introduction and overview. *Environmental Innovation and Societal Transitions*, 1, pp.1-23.

Berkhout, F., 2002. Technological regimes , path dependency and the environment. *Global Environmental Change*, 12, pp.1-4.

Berkhout, F. et al., 2010. Sustainability experiments in Asia : innovations shaping alternative development pathways ? *Environmental Science and Policy*, 13(4), pp.261-271.

Bijker, W.E., Hughes, T.P. & Pinch, T., 2012. *The social construction of technological systems. New directions in the sociology and history of technology*, Cambridge, Mass: MIT Press.

Le Billon, P., 2008. *The geopolitics of resource wars: resource dependence, governance and violence*, London: Routledge.

Binnemans, K. et al., 2013. Recycling of rare earths : a critical review. *Journal of Cleaner Production*, 51, pp.1-22.

Blaikie, N., 2008. *Approaches to social enquiry* 2nd ed., Cambridge, United Kingdom: Polity Press.

Bolton, R., 2011. *Socio-Technical Transitions and Infrastructure Networks: the cases of electricity and heat distribution in the UK*. University of Leeds.

Bolton, R. & Foxon, T.J., 2011. Governing infrastructure networks for a Low Carbon Economy: Co-evolution of technologies and institutions in UK electricity distribution networks. *Competition and Regulation in Network Industries*, 12, pp.2-26.

Bolton, R. & Foxon, T.J., 2013. Urban infrastructure dynamics: market regulation and the shaping of district energy in UK cities. *Environment and Planning A*, 45(9), pp.2194 - 2211.

Bows, A. & Anderson, K., 2008. Contraction and convergence: an assessment of the CCOptions model. *Climatic Change*, 91(3-4), pp.275-290.

Bradsher, K., 2010. Challenging China in rare earth mining. *New York Times*. Available at: http://www.nytimes.com/2010/04/22/business/energy-environment/22rare.html?pagewanted=all&_r=0.

Bradsher, K., 2012. US called vulnerable to rare earth shortages. *New York Times*. Available at: http://www.nytimes.com/2010/12/15/business/global/15rare.html?_r=3&hp&

Brown R. & Farrelly, M.A., 2009. Delivering sustainable urban water management: a review of the hurdles we face. *Water Science and Technology*,59(5) pp.839-846.

Brown R. & Fraser, M., 2006. Approaches and perspectives in social and environmental accounting: an overview of the conceptual landscape. *Business Strategy and the Environment*, 15, pp.103-117.

Busch, J. et al., 2014. Managing Critical Materials with a Technology-Specific Stocks and Flows Model. *Environmental Science & Technology*, 48 pp.1298-1305

Business Stream, 2013. Glasgow City Council water saving case study.

CCC, 2008. *Building a low-carbon economy – the UK's contribution to tackling climate change*,

CCC, 2013. *Meeting Carbon Budgets – 2013 Progress Report to Parliament*,

CCC, 2010. *The Fourth Carbon Budget. Reducing emissions through the 2020s*,

CST, 2009a. *A National Infrastructure for the 21st Century*, London.

CST, 2009b. Improving innovation in the water industry : 21st century challenges and opportunities.

CST, 2005. *Policy through dialogue: informing policies based on science and technology*,

Candelise, C., Speirs, J. & Gross, R., 2012. Materials availability for thin film (TF) PV technologies development: A real concern? *Renewable and Sustainable Energy Reviews*, 15(9), pp.4972-4981.

Carley, S. et al., 2011. Energy-based economic development. *Renewable and Sustainable Energy Reviews*, 15(1), pp.282-295.

Carlsson, B. & Stankiewicz, R., 1991. On the nature, function and composition of technological systems. *Journal of Evolutionary Economics*, 1(2), pp.93-118.

Carrillo-Hermosilla, J., 2006. A policy approach to the environmental impacts of technological lock-in. *Ecological Economics*, 58, pp.717 - 742.

Cashin, P. & McDermott, C.J., 2002. The Long-Run Behavior of Commodity Prices: Small Trends and Big Variability. *IMF Staff Papers*, 49(2), pp.175-199.

Cave, M., 2009. *Independent Review of Competition and Innovation in Water Markets: Final report*,

Centre for Alternative Technology, 2013a. *Zero Carbon Britain: Rethinking the Future*, CAT Publications.

Centre for Alternative Technology, 2013b. *Zero Carbon Britain: Rethinking the Future. Summary Report*,

Chakravarty, S. et al., 2009. Sharing global CO2 emission reductions among one billion high emitters. *Proceedings of the National Academy of Sciences of the United States of America*, 106(29), pp.11884-8.

Chen, R.S., 1975. Social and Financial Stewardship. *The Accounting Review*, 50(3), pp.533-543.

Chilvers, J., 2010. *Sustainable participation ?*

Cooper, T., 2005. Slower Consumption: Reflections on Product Life Spans and the "Throwaway Society." *Journal of Industrial Ecology*, 9(1-2), pp.51-67.

Corbin, J. & Strauss, A., 2008. *Basics of qualitative research: techniques and procedures for developinggrounded theory*, London: Sage.

Corner, A. et al., 2013. Messing with nature? Exploring public perceptions of geoengineering in the UK. *Global Environmental Change*, 23, pp.938-947.

Costanza, R. et al., 1993. Modeling Complex Ecological Economic Systems. Toward an evolutionary, dynamic understanding of people and nature. *BioScience*, 43(8), pp.545-555.

Cox, M., Arnold, G. & Tomás, S.V., 2010. A Review of Design Principles for Community-based Natural Resource. *Ecology And Society*, 15(4).

DECC, 2013a. *Digest of United Kingdom energy statistics (DUKES)*,

DECC, 2013b. *GHG Inventory Overview*,

DECC, 2013c. *Research into barriers to deployment of district heating networks*,

DECC, 2013d. *UK Renewable Energy Roadmap Update 2013*,

DECC, 2012a. *Domestic energy consumption in the UK since 1970*

DECC, 2012b. *Low Carbon Communities Challenge Evaluation Report*,

DECC, 2009. *Smarter Grids: The Opportunity*,

DTI, 2003. *Our energy future - creating a low carbon economy*,

DWI, About Us - Drinking Water Inspectorate. Available at: http://dwi.defra.gov.uk/about/index.htm.

Danlu, T., 2012. China vows stricter rare earth regulation. *Xinhuanet*. Available at: http://news.xinhuanet.com/english/china/2012-06/20/c_131666265.htm.

Darby, M., 2014. No man is an island. *Utility Week*.

Dawson, R et al, 2013. Infrastructure Definition for the iBuild Centre.

Dawson D. A, et al., 2014. Low carbon technology performance vs. infrastructure vulnerability: analysis through the local & global properties. *Environmental Science & Technology*. Under review.

Defra, 2013. *Review of Defra's funding for WRAP in England – Opportunity to comment.*

Defra, 2012. *Resource Security Action Plan: Making the most of valuable materials*.

Defra, 2011a. *Review of Ofwat and consumer representation in the water sector*.

Defra, 2011b. *Water for Life: Market reform proposals*.

Defra, 2008. *Future Water. The Government's water strategy for England*.

Du, X. & Graedel, T.E. 2011a. Global Rare Earth In-Use Stocks in NdFeB Permanent Magnets. *Journal of Industrial Ecology*, 15(6), pp.836-843.

Du, X. & Graedel, T.E. 2011b. Global in-use stocks of the rare Earth elements: a first estimate. *Environmental Science & Technology*, 45(9), pp.4096-101.

Égert, B., Kozluk, T. & Sutherland, D., 2009. Infrastructure investment: Links to growth and the role of public policies.

Eisenhardt, K.M., 1989. Building Theories from Case Study Research. *Academy of Management Review*, 14(4), pp.532-550.

Ekins, P. et al., 2013. *The UK energy system in 2050: Comparing Low-Carbon, Resilient Scenarios*.

Ekins, P. et al., 2012. Increasing carbon and material productivity through environmental tax reform. *Energy Policy*, 42, pp.365-376.

Electric Vehicle News, 2011. Toyota Developing Induction Motors to Break Reliance on China. Available at: http://www.electric-vehiclenews.com/2011/01/toyota-developing-induction-motors-to.html.

Emerson, J.W. et al., 2012. *Environmental Performance Index and pilot trend*, New Haven.

Environment Agency, 2013a. *Annual review of water resources management plans guidance*.

Environment Agency, 2013b. *Environmental Permitting Guidance Core guidance for the Environmental Permitting (England and Wales) Regulation 2010*.

Environment Agency, 2009. *A Low Carbon Water Industry in 2050*.

Erdmann, L. & Graedel, 2011. Criticality of non-fuel minerals: a review of major approaches and analyses. *Environmental Science & Technology*, 45(18), pp.7620-30.

European Commission, 2012a. Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC.

European Commission, 2012b. *EU Trade Policy for Raw Materials. Second Activity Report*,

European Commission, 2011a. COM (2011) 0025 Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Tackling the challenges in commodity markets and on raw materials.

European Commission, 2011b. *On an effective raw materials strategy for Europe*.

European Commission, 2010a. *EC Guidance on: Undertaking non-energy extractive activities in accordance with Natura 2000 requirements*.

European Commission, 2010b. *Report of the Ad-hoc Working Group on defining critical raw materials*,

European Commission, 2009. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources.

European Commission, 2008a. COM 2008 30 Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: 20 20 by 2020 Europe’s climate change opportunity.

European Commission, 2008b. *COM(2008) 699 Communication from the Commissions to the European Parliament and the Council. The raw materials initiative - meeting our critical needs for growth and jobs in Europe*.

European Commission, 2006. Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality.

European Commission, 2000. Establishing a framework for Community action in the field of water policy.

European Commission, 1998. Council Directive 98/83/EC on the quality of water intended for human consumption.

European Commission, 1992. Council Directive 92/43/EEC on the conservation of natural habitatis and of wild fauna and flora.

European Commission, 1991. Council Directive concerning urban waste water treatment.

Eurostat, 2009. *Economy Wide Material Flow Accounts : Compilation Guidelines for reporting to the 2009 Eurostat questionnaire*,

Eurostat, 2001. *Economy-wide material flow accounts*.

Evans, R. & Kotchetkova, I., 2009. Qualitative research and delivberative methods: promise or peril? *Qualitative Research*, 9(625-643).

Eyre, N., 2013. Energy saving in energy market reform — The feed-in tariffs option. *Energy Policy*, 52, pp.190-198.

Ezzy, D., 2002. *Qualitative Analysis: Practice and Innovation*, London: Routledge.

Farrelly, M.A. & R. Brown, 2011. Rethinking urban water management: Experimentation as a way forward? *Global Environmental Change*, 21(2), pp.721-732.

Fell, D. et al., 2009. *The diffusion of environmental behaviours; the role of influential individuals in social networks Report 1: Key findings. A research report completed for the Department for Environment, food and Rural Affairs*, London.

Findon, R., 2013a. Council calls for police probe into Ecoisland cash. *Isle of Wight County Press*. Available at: http://www.iwcp.co.uk/news/news/council-calls-for-police-probe-into-ecoisland-cash-51184.aspx.

Findon, R., 2013b. Ecoisland goes into liquidation. *Isle of Wight County Press*. Available at: http://www.iwcp.co.uk/news/news/ecoisland-goes-into-liquidation-51149.aspx.

Fischer-Kowalski, M. & Rotmans, J., 2009. Conceptualizing, Observing, and Influencing Social–Ecological Transitions. *Ecology and Society*, 14(2).

Fischer-Kowalski, M. & Weisz, H., 1999. Society as a Hybrid Between Material and Symbolic Realms . Toward a Theoretical Framework of Society-Nature Interaction. *Advances in Human Ecology*, 8, pp.215-251.

Fontenergy, 2009. License “light” - is this the change DE has been waiting for?.

Fouquet, R. & Pearson, P.J.G., 1999. A thousand years of energy use in the United Kingdom. *Energy Journal*, 19(4), pp.1-41.

Foxon, T. J., Gross, R. et al., 2005. UK innovation systems for new and renewable energy technologies: drivers, barriers and systems failures. *Energy Policy*, 33, pp.2123-2137.

Foxon, T.J., Pearson, P. et al., 2005. *Transforming policy processes to promote sustainable innovation: some guiding principles*.

Foxon, T.J., 2011. A coevolutionary framework for analysing a transition to a sustainable low carbon economy. *Ecological Economics*, 70(12), pp.2258-2267.

Foxon, T.J., 2012. Transition pathways for a UK low carbon electricity future. *Energy Policy*, 52, pp.10-24.

Foxon, T.J., Reed, M.S. & Stringer, L.C., 2009. Governing Long-Term Social – Ecological Change: What Can the Adaptive Management and Transition Management Approaches Learn from Each Other ? *Environmental Policy and Governance*, 19, pp.3-20.

Froderman, R. et al., 2010. *The Oxford handbook of interdisciplinarity*, Oxford: Oxford University Press.

Fromer, N., Eggert, R.G. & Lifton, J., 2011. *Critical materials for sustainable energy applications*.

Frontier Economics, 2012. *Systemic risks and opportunities in UK infrastructure: A report prepared for HM Treasury and Infrastructure UK.*

Fthenakis, V., 2009. Sustainability of photovoltaics: The case for thin-film solar cells. *Renewable and Sustainable Energy Reviews*, 13(9), pp.2746-2750.

Geels, F.W., 2007. Analysing the breakthrough of rock “n” roll (1930 – 1970). Multi-regime interaction and reconfiguration in the multi-level perspective. *Technological Forecasting & Social Change*, 74(8), pp.1411-1431.

Geels, F.W., 2004. From sectoral systems of innovation to socio-technical systems Insights about dynamics and change from sociology and institutional theory. *Research Policy*, 33, pp.897-920.

Geels, F.W., 2002. Technological transitions as evolutionary reconfiguration processes : a multi-level perspective and a case-study. *Research Policy*, 31, pp.1257-1274.

Geels, F.W., 2011. The multi-level perspective on sustainability transitions: Responses to seven criticisms. *Environmental Innovation and Societal Transitions*, 1, pp.24-40.

Geels, F.W. & Raven, R., 2006. Non-linearity and expectations in niche-development trajectories: Ups and downs in Dutch biogas development (1973-2003). *Technology Analysis and Strategic Management*, 18(3/4), pp.375-392.

Geels, F.W. & Schot, J., 2007. Typology of sociotechnical transition pathways. *Research Policy*, 36, pp.399-417.

Gibbon, J. & Dey, C., 2011. Developments in Social Impact Measurement in the Third Sector: Scaling Up or Dumbing Down? *Social and Environmental Accountability Journal*, 31(1), pp.63-72.

Glaser, B. & Strauss, A., 1967. *The discovery of grounded theory: stratgies for qualitative research*, New York: Aldine de Gruyter.

Graedel T.E. et al., 2012. Methodology of Metal Criticality Determination. *Environmental Science & Technology*, 46, pp.1063-1070.

Graedel, T.E., 2011. On the Future Availability of the Energy Metals. *Annual Review of Materials Research*, pp.323-337.

Greater London Authority, 2013. Licence Lite. *london.gov.uk*. Available at: http://www.london.gov.uk/priorities/environment/tackling-climate-change/energy-supply/licence-lite [Accessed December 31, 2013].

Green, R., 2010. Are the British electricity trading and transmission arrangements future-proof? *Utilities Policy*, 18, pp.186-194.

Gross, R. & Heptonstall, P., 2010. Time to stop experimenting with UK renewable energy policy.

Gross, R. et al., 2012. On picking winners: The need for targeted support for renewable energy.

Group of Eight (G8), 2009. Responsible leadership for a sustainable future, G8 Leader's Declaration, L'Aquila, Italy.

Grubler, A., 1998. *Technology and global change*, Cambridge University Press.

HM Government, 2014a. gov.uk Department for Environment, Food and Rural Affairs. Available at: https://www.gov.uk/government/organisations/department-for-environment-food-rural-affairs.

HM Government, 2014b. gov.uk Department of Energy and Climate Change. Available at: https://www.gov.uk/government/organisations/department-of-energy-climate-change.

HM Government, 2014c. gov.uk Environment Agency. Available at: https://www.gov.uk/government/organisations/environment-agency.

HM Government, 2011a. *The Carbon Plan: Delivering our low carbon future*.

HM Government, 2011b. *Water for Life*.

HM Government, 2009. *The UK Low Carbon Transition Plan*.

HM Government, 2008a. *Climate Change Act - Chapter 27*.

HM Government, 2008b. Energy Act.

HM Government, 2007. *Meeting the energy challenge: A White Paper on Energy*.

HM Government, 2004. Energy Act 2004.

HM Government, 2003. Water Act 2003.

HM Government, 2000. Utilities Act 2000.

HM Government, 1999. Water Industry Act 1999.

HM Government, 1989. Water Act.

HM Government, 1963. Water Resources Act, 1963.

HM Government, 1947. Electricity Act, 1947.

Haasnoot, M., Kwakkel, J. H & Walker, W. E, 2013. Dynamic adaptive policy pathways : A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change*, 23(2), pp.485-498.

Haberl, H. et al., 2011. A socio-metabolic transition towards sustainability? Challenges for another great transformation. *Sustainable Development*, 14, pp.1-14.

Haberl, H. et al., 2004. Progress towards sustainability? What the conceptual framework of material and energy flow accounting (MEFA) can offer. *Land Use Policy*, 21, pp.199-213.

Habermas, J., 1972. *Knowledge and human interests*, London: Heinemann Educational.

Hall, 1993. Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain. *Comparative Politics*, 25(3), pp.275-296.

Hall, J.W., Henriques, J.J., Hickford, J.J., et al., 2012. *A Fast Track Analysis of strategies for infrastructure provision in Great Britain: Technical report*, Oxford.

Hall, J.W., Sayers, P.B. & Dawson, R.J., 2005. National-scale Assessment of Current and Future Flood Risk in England and Wales. *Civil Engineering*, 36, pp.147-164.

Hannon, M.J., 2012. *Co-evolution of innovative business models and sustainability transitions: The case of the Energy Service Company (ESCo) model and the UK energy system*. University of Leeds.

Hannon, M.J., Foxon, T.J. & Gale, W.F., 2013. The co-evolutionary relationship between Energy Service Companies and the UK energy system: Implications for a low-carbon transition. *Energy Policy*, 61, pp.1031-1045.

Hargreaves, T., Hielscher, S., et al., 2013. Grassroots innovations in community energy: The role of intermediaries in niche development. *Global Environmental Change*, 23(5), pp.868-880.

Hargreaves, T., Longhurst, N. & Seyfang, G., 2013. Up, down, round and round: connecting regimes and practices in innovation for sustainability. *Environment and Planning A*, 45, pp.402-420.

Haringey Council, 2012. *Licence Lite - Strategic and policy issues*,

Harper, E.M., Kavlak, G & Graedel, T.E., 2012. Tracking the Metal of the Goblins: Cobalt's Cycle of Use. *Environmental Science & Technology*, 46(2), pp.1079-1086.

Hawkey, D., 2012. District heating in the UK: A Technological Innovation Systems analysis. *Environmental Innovation and Societal Transitions*, 5, pp.19-32.

Hawkey, D., Webb, J. & Winskel, M., 2013. Organisation and governance of urban energy systems: district heating and cooling in the UK. *Journal of Cleaner Production*, 50, pp.22-31.

Hekkert, M.P. et al., 2007. Functions of innovation systems : A new approach for analysing technological change. *Technological Forecasting and Social Change*, 74, pp.413 - 432.

Helm, D., 2009. Infrastructure investment, the cost of capital, and regulation: an assessment. *Oxford Review of Economic Policy*, 25(3), pp.307-326.

Helm, D., 2005. The assessment: The new energy paradigm. *Oxford Review of Economic Policy*, 21(1), pp.1-18.

Hildyard, L., 2011. *Power and money: How local economies can benefit from renewable energy*,

Hommels, A., Peters, P. & Bijker, W.E., 2007. Techno therapy or nurtured niches ? Technology studies and the evaluation of radical innovations. *Research Policy*, 36, pp.1088-1099.

Houari, Y. et al., 2014. A system dynamics model of tellurium availability for CdTe PV. *Progress in Photovoltaics: Research and Applications*, 22, pp.129-146.

House of Lords Science and Technology Committee, 2007. *Water Management Follow-up.*

House of Lords Science and Technology Committee, 2006. *Water Management Volume I: Report*.

Howells, J., 2006. Intermediation and the role of intermediaries in innovation. *Research Policy*, 35, pp.715-728.

Hughes, 1983. *Networks of power: electrification in Western society, 1880-1930*, Baltimore: Johns Hopkins University Press.

Hughes, N. 2013. Towards improving the relevance of scenarios for public policy questions: A proposed methodological framework for policy relevant low carbon scenarios. *Technological Forecasting & Social Change*, 80(4), pp.687-698.

Hughes N. & Strachan, N., 2010. Methodological review of UK and international low carbon scenarios. *Energy Policy*, 38(10), pp.6056-6065.

Hughes, N. Strachan, N. & Gross, R., 2012. The structure of uncertainty in future low carbon pathways. *Energy Policy*, 52, pp.45-54.

Huntjens, P. et al., 2012. Institutional design propositions for the governance of adaptation to climate change in the water sector. *Global Environmental Change*, 22, pp.67-81.

IBM, 2010. *Switching perspectives: Creating new business models for a changing world of energy*,

ICE, 2009. *A National Infrastructure Investment Bank*,

IEA, 2012. *Energy Technology Perspectives 2012 Pathways to a Clean Energy System Executive Summary*.

IEA, 2010. *Energy Technology Perspectives: Scenarios and strategies to 2050*.

IEA-RETD, 2012. *Securing the supply chain for wind and solar energy (RE-SUPPLY) Final report*.

IPCC, 2014. *Climate Change 2014: Working Group III - Mitigation of Climate Change: Technical Summary*.

IPCC, 2013. *Climate change 2013: The Physical Science Basis - Summary for Policymakers*.

IPCC, 2007. *Climate Change 2007: The physical science basis. Contribution of working Group 1 to the fourth assessment report of the Intergovernmental Panel on Climate Change*, Cambridge, United Kingdom: Cambridge University press.

IPCC, 2001. *Third Assessment Report: the scientific basis*.

Infrastructure UK, 2011. *National Infrastructure Plan 2011*.

Infrastructure UK, 2010. *National Infrastructure Plan 2010*.

Island Echo, 2013. Eco Island founder found dead. *Island Echo*. Available at: http://www.islandecho.co.uk/news/eco-island-founder-found-dead/.

Jacobsson, S. & Bergek, A., 2011. Innovation system analyses and sustainability transitions: Contributions and suggestions for research. *Environmental Innovation and Societal Transitions*, 1, pp.41-57.

Jones, A., 2013. Interview about Newcastle City Council Energy Masterplan.

Jones, A., 2012. *Newcastle Energy Masterplan*.

Kallis, G. & Norgaard, R.B., 2010. Coevolutionary ecological economics. *Ecological Economics*, 69(4), pp.690-699.

Kara, H. et al., 2010. *Lanthanide Resources and Alternatives*,

Keat, R. & Urry, J., 1982. *Social theory as science* 2nd ed., London: Routledge.

Keay, M., 2012. Death by a thousand regulations: the New Energy Bill.

Keeley & Scoones, I., 2003. *Understanding environmental policy processes: cases from Africa*, London: Earthscan.

Keeley, B., 2009. The Value of People. In *Human Capital: How what you know shapes your life*. OECD Publishing.

Kemp, R. & van Lente, H., 2011. The dual challenge of sustainability transitions. *Environmental Innovation and Societal Transitions*, 1, pp.121-124.

Kemp, R. & Loorbach, D., 2006. Transition management: a reflexive governance approach. In J.-P. Voß, D. Bauknecht, & R. Kemp, eds. *Reflexive governance for sustainable development*. Edward Elgar, pp. 103-130.

Kern, F., 2011. Ideas, institutions, and interests: explaining policy divergence in fostering`system innovations’ towards sustainability. *Environment and Planning C*, 29, pp.1116-1134.

Kern, F. & Mitchell, C., 2010. Policy Paradigms as part of the landscape: How do policy paradigms influence attempts to govern transitions? In *14th IRSPM Conference, Berne, Switzerland, 7-9th April 2010*. pp. 1-18.

Kern, F. & Smith, 2008. Restructuring energy systems for sustainability? Energy transition policy in the Netherlands. *Energy Policy*, 36, pp.4093- 4103.

Kern, F., Kuzemko, C. & Mitchell, C., 2014. Measuring and explaining policy paradigm change: The case of UK energy policy. *Policy & Politics*, In Press.

Kleijn, R. & van der Voet, E., 2010. Resource constraints in a hydrogen economy based on renewable energy sources: An exploration. *Renewable and Sustainable Energy Reviews*, 14(9), pp.2784-2795.

Kleijn, R. et al., 2011. Metal requirements of low-carbon power generation. *Energy*, 36, pp.5640-5648.

Knoeri, C., Steinberger, J.K. & Roelich, K., 2014. End-user centred infrastructure operation : Towards integrated infrastructure service delivery. *Journal of Cleaner Production* (In review).

Knoeri, C. et al., 2013. Towards a dynamic assessment of raw materials criticality: Linking agent-based demand - with material flow supply modelling approaches. *Science of the Total Environment*, 461-462, pp.808-812.

Koppenjan, J.F.M. & Bert, E., 2009. Public-Private Partnerships in Urban Infrastructures: Reconciling Private Sector Participation and Sustainability. *Public Administration Review*, pp.284-296.

Krausmann, F., Schandl, H. & Sieferle, R.P., 2008. Socio-ecological regime transitions in Austria and the United Kingdom. *Ecological Economics*, 65, pp.187-201.

Kwon, T.-hyeong, 2012. Technological Forecasting & Social Change Strategic niche management of alternative fuel vehicles : A system dynamics model of the policy effect. *Technological Forecasting & Social Change*, 79(9), pp.1672-1680.

Lachman, D.A., 2013. A survey and review of approaches to study transitions. *Energy Policy*, 58, pp.269-276.

Langbein, L. & Knack, S., 2010. The Worldwide Governance Indicators: Six, One, or None? *Development*, 46(2), pp.350-370.

Leach, M. et al., 2012. Transforming Innovation for Sustainability. *Ecology and Society*, 17(2).

Leach, M., Scoones, I. & Stirling, A. 2010. Governing epidemics in an age of complexity: Narratives, politics and pathways to sustainability. *Global Environmental Change*, 20, pp.369-377.

Lee, P., 2013. Only full-scale reform of our energy market will prevent endless price rises. *The Observer*. Available at: http://www.theguardian.com/commentisfree/2013/oct/26/energy-price-rises-market-reform.

Less, S., 2012. *Greening the economy - not “green economy,”*

Leventon, J. & Antypas, A., 2012. Multi-level Governance, Multi-level Deficits: The Case of Drinking Water Management in Hungary. *Water Management*, 267, pp.253-267.

Little, R.G., 2005. Tending the infrastructure commons: ensuring the sustainability of our vital public systems. *Structure and Infrastructure Engineering*, 1(4), pp.263 - 270.

Liu, J. et al., 2007. Coupled Human and Natural Systems. *Ambio*, 36(8), pp.639-649.

London Energy Partnership, 2007. *Making ESCOs Work: Guidance and Advice on Setting Up & Delivering an ESCO*,

Longhurst, N. & Chilvers, J., 2012. *Interdisciplinarity in transition? A Technical Report on the interdisciplinarity of the Transition Pathways to a Low Carbon Economy consortium*, Norwich.

van der Loos, F. & Loorbach, D., 2012. The Dutch Energy Transition Project (2000-2009). In G. Verbong & D. Loorbach, eds. *Governing the Energy Transition. Reality, illusion or necessity?* Routledge, pp. 220-250.

Lovins, A.B. 1996. Negawatts: Twelve transitions, eight improvements and one distraction. *Energy Policy*, 24(4), pp.331-343.

Lovins A.B. & Lovins, L.H. 1991. Least-cost climatic stabilisation. *Annual Review of Energy and Environment*, 16, pp.433-431.

Malerba, F., 2001. *Sectoral Systems of Innovation: Concepts, issues and analyses of six major sectors in Europe*, Cambridge University Press.

Marino, A., Bertoldi, P. & Rezessy, S., 2010. *Energy Service Companies Market in Europe*.

Marks, G. & Hooghe, L., 2003. Contrasting Visions of Multi-Level Governance. In *Multi-Level Governance: Interdisciplinary Perspectives*. Oxford University Press.

Marshall, T., 2010. Working Paper – Infrastructure and Spatial Planning Planning major infrastructure sectors in the UK during recent decades.

Mason, L. et al., 2011. Availability, addiction and alternatives: three criteria for assessing the impact of peak minerals on society. *Journal of Cleaner Production*, 19(9-10).

Mayor of London, 2013. RE:FIT Briefing Note.

McNichol, T., 2006. *AC/DC: The savage tale of the first standards war*, Wiley.

van de Meene, S.J., Brown, R.R. & Farrelly, M.A., 2011. Towards understanding governance for sustainable urban water management. *Global Environmental Change*, 21, pp.1117-1127.

Mitchell, V.G. 2006. Applying integrated urban water management concepts: A review of Australian experience. *Environmental Management*, 37(5), pp.589-605.

Mitchell, C., 2014. Change and Inertia in the UK Energy System – getting our institutions and governance right. iGov Working Paper.

Mitchell, C., 2010. *The political economy of sustainable energy*, Palgrave Macmillan.

Mitchell, C. & Woodman, B., 2010. Towards trust in regulation - moving to a public value regulation. *Energy Policy*, 38, pp.2644-2651.

Moe, E., 2010. Energy, industry and politics: Energy, vested interests, and long-term economic growth and development. *Energy*, 35, pp.1730-1740.

Moldan, B. & Dahl, A.L., 2007. Challenges to sustainability indicators. In T. Hak, B. Moldan, & A. L. Dahl, eds. *Sustainability Indicators - A Scientific Assessment*. Island Press, pp. 1-26.

Mont, O., 2004. Institutionalisation of sustainable consumption patterns based on shared use. *Ecological Economics*, 50(1-2), pp.135-153.

Mont, O. & Tukker, A., 2006. Product-Service Systems: reviewing achievements and refining the research agenda. *Journal of Cleaner Production*, 14(17), pp.1451-1454.

Moran, M., 2003. *The British regulatory state: High modernism and hyper-innovation*, Oxford: Oxford University Press.

Moriguchi, Y., 2010. Mineral resources: Quantitative and qualitative aspects of sustainability. In Thomas E Graedel & E. van der Voet, eds. *Linkages of Sustainability*.

Morley, N. & Eatherley, D., 2008. *Material Security: Ensuring resource availability for the UK economy*.

Mortimer, D., 2012. *Complex adaptive systems methodologies for infrastructure modelling*.

Moselle, B. & Moore, S., 2011. *Climate Change Policy – Time for Plan B*.

Moss, R.L. et al., 2011. *Critical Metals in Strategic Energy Technologies Assessing Rare Metals as Supply-Chain Bottlenecks in Low-Carbon Energy Technologies*.

Moss, R.L. et al., 2013. *Critical Metals in the Path towards the Decarbonisation of the EU Energy Sector*.

Mumford, J. & Gray, D., 2010. Consumer engagement in alternative energy - Can the regulators and suppliers be trusted? *Energy Policy*, 38(6), pp.2664-2671.

Munda, G., 2005. Measuring sustainability: a multi-criterion framework. *Environment, Development and Sustainability*, 7, pp.117-134.

Nassar, N.T. et al., 2012. Criticality of the Geological Copper Family. *Environmental Science & Technology*, 46, pp.1071-1078.

National Grid, 2005. The National Gas Archives - Nationalisation. Available at: http://www.gasarchive.org/Nationalisation.htm [Accessed March 8, 2014].

National Research Council, 2008. *Minerals, Critical Minerals, and the U.S. Economy*, Washington, DC: The National Academies Press.

Newell, P., Pattberg, P. & Schroeder, H., 2012. Multiactor Governance and the Environment. *Annual Review of Environment and Resources*, 37, pp.365-387.

Nooteboom, B., 2000. Institutions and forms of co-ordination in innovation systems. *Organization Studies*, 21(5), pp.915-939.

Norgaard, R.B., 1994. *Development betrayed: The end of progress and a co-evolutionary revisioning of the future*, London: Routledge.

Norgate, T., 2010. Deteriorating ore reserves: Energy and water impacts. In Thomas E. Graedel & E. van der Voet, eds. *Linkages of Sustainability*.

North, D., 1993. Economic performance through time. Lecture to the memory of Alfred Nobel.

Nye, M., Whitmarsh, L. & Foxon, T.J., 2010. Sociopsychological perspectives on the active roles of domestic actors in transition to a lower carbon electricity economy. *Environment and Planning A*, 42(3), pp.697-714.

OECD, 2008. *Handbook on Constructing Composite Indicators. Methodology and user guide*.

OECD, 1999. *Managing National Innovation Systems*.

Office of Fair Trading, 2010. *Infrastructure Ownership and Control Stock-take*.

Ofgem, 2014. *Standard conditions of electricity supply licence*.

Ofgem, 2013a. *Making it easier for independent suppliers and generators to compete*.

Ofgem, 2013b. *Wholesale power market liquidity: final proposals for a "Secure and Promote" licence condition*.

Ofgem, 2013c. Why Licence Lite?

Ofgem, 2011. *The Retail Market Review - Findings and initial proposals*.

Ofgem, 2010. *Demand Side Reponse: A discussion paper*.

Olivier, J.G.J., Janssens-Maenhout, G. & Peters, J., 2012. *Trends in global CO2 emissions: 2012 report*, The Hague: PBL Netherlands Environmental Assessment Agency; Ispra: Joint Research Centre.

Olsson, P. et al., 2006. Shooting the Rapids : Navigating Transitions to Adaptive Governance of Social-Ecological Systems. *Ecology And Society*, 11(1).

Ostrom, E., 2012. Nested externalities and polycentric institutions : must we wait for global solutions to climate change before taking actions at other scales ? *Symposium A Quarterly Journal In Modern Foreign Literatures*, pp.353-369.

Ostrom, E., 2010. Beyond Markets and States: Polycentric Governance of Complex Economic Systems. *American Economic Review*, 100, pp.641-672.

Ostrom, E., 2007. A diagnostic approach for going beyond panaceas. *Proceedings of the National Academy of Sciences*, 104(39), pp.15181-15187.

Ostrom, E., 2005. *Understanding institutional diversity*, Princeton University Press.

PWC Economics, 2013. *World in 2050 The BRICs and beyond : prospects , challenges and opportunities*.

Paavola, J. & Adger, W.N., 2005. Institutional ecological economics. *Ecological Economics*, 53, pp.353 - 368.

van der Pas, J.W.G.M. et al., 2013. Operationalizing adaptive policymaking. *Futures*, 52, pp.12-26.

Patterson, W., 2013. Beyond The Fire Age. , pp.1-13.

Patterson, W., 2008. Managing Energy: Rethinking the Fundamentals. Managing Energy Wrong.

Patterson, W., 1999. *Transforming electricity: the coming generation of change*, Earthscan.

Pihl, E. et al., 2012. Material constraints for concentrating solar thermal power. *Energy*, 44, pp.944-954.

Polinder, H. et al., 2005. Comparison of Direct-Drive and Geared Generator Concepts for Wind Turbines. *IEEE Control Systems Magazine*, pp.543-550.

RCEP, 2000. *Twenty Second Report: Energy - the changing climate*.

Rademaker, J.H., Kleijn, R. & Yang, Y., 2013. Recycling as a Strategy against Rare Earth Element Criticality : A Systemic Evaluation of the Potential Yield of NdFeB Magnet Recycling. *Environmental Science and Technology*, 47, pp.10129-10136.

Rafols, I. et al., 2012. How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management. *Research Policy*, 41(7), pp.1262-1282.

Ranger, N. et al., 2010. *Adaptation in the UK: a decision-making process*.

Raven, R., 2012. Analyzing emerging sustainable energy niches in Europe. A strategic Niche Management Perspective. In G. Verbong & D. Loorbacj, eds. *Governing the Energy Transition. Reality, illusion or necessity?* Routledge, pp. 125-151.

Reck, B.K. & Graedel, T.E. 2012. Challenges in Metal Recycling. *Science*, 337, pp.690-695.

Reeder, T. & Ranger, N., 2010. *How do you adapt in an uncertain world ? Lessons from the Thames Estuary 2100 project*.

Renn, O., 2008. Concepts of Risk: An Interdisciplinary Review. Part 1: Disciplinary risk concepts. *Gaia*, 17(1), pp.50 - 66.

Rijke, J. et al., 2013. Configuring transformative governance to enhance resilient urban water systems. *Environmental Science and Policy*, 25, pp.62-72.

Rinaldi, S.M., Peerenboom, J.P. & Kelly, T.K., 2001. Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies. *IEEE Control Systems Magazine*, pp.11-25.

del Rio, P. & Unruh, G.C., 2007. Overcoming the lock-out of renewable energy technologies in Spain : The cases of wind and solar electricity. *Renewable and Sustainable Energy Reviews*, 11, pp.1498-1513.

Rockström, J. et al., 2009. A safe operating space for humanity. *Nature*, 461(September), pp.472-475.

Roelich, K. et al., 2014. Towards resource-efficient and service-oriented integrated infrastructure operation. *Technological Forecasting and Social Change* (accepted).

Rotmans, J. & Kemp, R., 2008. Detour ahead: a response to Shove and Walker about the perilous road of transition management. *Environment and Planning A*, 40(4), pp.1006-1012.

Rotmans, J. & Loorbach, D., 2009. Complexity and Transition Management. *Journal of Industrial Ecology*, 13(2), pp.184-196.

Rotmans, J. & Loorbach, D., 2010. Research into the governance of transitions: A framework for Transitions Management. In J. Grin, J. Rotmans, & J. Schot, eds. *Transitions to sustainable development. New directios in the study of long term transformative change*. Routledge.

Royal Society, 1992. *Risk: analysis, perception and management: a report of a Royal Society Study Group*,

Safarzyńska, K. & van den Bergh, J.C.J.M., 2010. Demand-supply coevolution with multiple increasing returns: Policy analysis for unlocking and system transitions. *Technological Forecasting & Social Change*, 77(2), pp.297-317.

Saunders, R.W., Gross, R.J.K. & Wade, J., 2012. Can premium tariffs for micro-generation and small scale renewable heat help the fuel poor, and if so , how ? Case studies of innovative finance for community energy schemes in the UK. *Energy Policy*, 42, pp.78-88.

Schot, J. & Geels, F.W., 2008. Strategic niche management and sustainable innovation journeys : theory , findings , research agenda , and policy. *Technology Analysis & Strategic Management*, 20(5), pp.537-554.

Schüler, D. et al., 2011. *Study on Rare Earths and their recycling. Final report for the Greens/EFA Group in the European Parliament*.

Sen, A., 1997. Editorial: Human Capital and Human Capability. *World Development*, 25(12), pp.1959-1961.

Seyfang, G. & Haxeltine, A., 2012. Growing grassroots innovations: exploring the role of community-based initiatives in governing sustainable energy transitions. *Environment and Planning C*, 30, pp.381 - 400.

Seyfang, G. & Smith, 2007. Grassroots innovations for sustainable development: Towards a new research and policy agenda. *Environmental Politics*, 16(4), pp.584-603.

Seyfang, G., Jin, J. & Smith, A., 2013. A thousand flowers blooming? An examination of community energy in the UK. *Energy Policy*, 61, pp.977-989.

Sherry, A., 1984. The power game - the development of conventional power stations 1948-1983. *Proceedings of the Institution of Mechanical Enginees, Part A: Journal of Power and Energy*, 198(4), pp.257-280.

Shove, E., 2003. *Comfort, cleanliness and convenience: The social organisation of normality*, Oxford: Berg.

Shove, E. & Pantzar, M., 2005. Consumers, Producers and Practices: Understanding the invention and reinvention of Nordic walking. *Journal of Consumer Culture*, 5(1), pp.43-64.

Shove, E. & Walker, G. 2010. Governing transitions in the sustainability of everyday life. *Research Policy*, 39(4), pp.471-476.

Shove, E. & Walker, G. 2007. CAUTION! Transitions ahead: politics, practice, and sustainable transition management. *Environment and Planning A*, 39, pp.763-770.

Shove, E. et al., 2008. Comfort in a lower carbon society. *Building Research & Information*, 36(4), pp.307-311.

Skea, J., Ekins, P. & Winskel, M. eds., 2010. *Energy 2050: making the transition to a secure low carbon energy system*.

Smith, A. 2009. Energy Governance: The Challenges of Sustainability. In J. I. Scrase & G. MacKerron, eds. *Energy for the Future: A new Agenda*. Palgrave Macmillan, pp. 54-75.

Smith, A. 2007. Emerging in between: The multi-level governance of renewable energy in the English regions. *Energy Policy*, 35, pp.6266-6280.

Smith, A. & Raven, R., 2012. What is protective space? Reconsidering niches in transitions to sustainability. *Research Policy*, 41(6), pp.1025-1036.

Smith, A. & Stirling, A. 2006. Moving Inside or Outside? Positioning the Governance of Sociotechnical Systems. *Governance An International Journal of Policy and Administration*, 44(148), pp.0-21.

Smith, A. et al., 2009. Assessing dangerous climate change through an update of the Intergovernmental Panel on Climate Change (IPCC) “reasons for concern.” *PNAS*, 106(11), pp.4133-4137.

Smith, A. Stirling, A. & Berkhout, F., 2005. The governance of sustainable socio-technical transitions. *Research Policy*, 34, pp.1491-1510.

Smith, A. Voß, J.P. & Grin, J., 2010. Innovation studies and sustainability transitions : The allure of the multi-level perspective and its challenges. *Research Policy*, 39, pp.435-448.

Sorrell, S., 2005. *The contribution of energy service contracting to a low carbon economy*.

Sorrell, S., 2007. The economics of energy service contracts. *Energy Policy*, 35(1), pp.507-521.

Speirs, J. et al., 2011. Materials Availability: Potential constraints to the future low-carbon Working Paper.

Speirs, J., Gross, R., et al., 2013. *Energy Materials Availability Handbook*.

Speirs, J., Houari, Y., et al., 2013. *Materials Availability: Potential constraints to the future low-carbon economy Working Paper II: Batteries, Magnets and Materials*.

Stanford Clark, A., 2013a. Interview about IBM’s involvement in Chale Green.

Stanford Clark, A., 2013b. Interview about IBM’s involvement in Eco-Island.

Steinberger, J.K., van Niel, J. & Bourg, D., 2009. Profiting from negawatts: Reducing absolute consumption and emissions through a performance-based energy economy. *Energy Policy*, 37(1), pp.361-370.

Stirling, A. 2011. Pluralising progress : From integrative transitions to transformative diversity. *Environmental Innovation and Societal Transitions*, 1(1), pp.82-88.

Stirling, A. 2010. Multicriteria diversity analysis A novel heuristic framework for appraising energy portfolios. *Energy Policy*, 38, pp.1622-1634.

Stirling, A. 2009. Direction, Distribution and Diversity! Pluralising progress in innovation, sustainability and development. STEPS Working Paper 32.

Stirling, A. 2008. “Opening up” and “Closing down”: Power, participation and pluralism in the social appraisal of technology. *Science, Technology and Human Values*, 33(2), pp.262-294.

Stirling, A. 2007. A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4, pp.707-719.

Stirling, A. 1998. On the economics and analysis of diversity.

Taylor, A., Cocklin, C. & Brown, Rebekah, R. 2012. Fostering environmental champions : A process to build their capacity to drive change. *Journal of Environmental Management*, 98, pp.84-97.

Thomas, S. 2006. The British Model in Britain: Failing slowly. *Energy Policy*, 34, pp.583-600.

Thomas, M.A., 2010. What Do the Worldwide Governance Indicators Measure? *European Journal of Development Research*, 22, pp.31-54.

Thorp, J.P., 2011. Delivering affordable and sustainable energy: the results of innovative approaches by Woking Borough Council, UK. In S. Lahlou, ed. *System Innovation for Sustainability 4: Case studies in sustainable consumption and production - Energy use and the built environment*. Greenleaf Publishing Limited.

Thorp, J.P., 2007. Woking: Driving Sustainable Innovation. In *Sustainable Innovation: Building & Construction Technologies*.

Tilton, J.E. & Lagos, G., 2007. Assessing the long-run availability of copper. *Resources Policy*, 32, pp.19-23.

Tutton, T., 2009. *The (potentially) topsy-turvy world of energy regulation: RPI-X@2- and Project Discovery*,

UKERC & ETI, 2010. *Marine Energy Technology Roadmap.*

UNEP, 2011. *Recycling Rates of Metals*.

UNEP, 2010. *The Emissions Gap Report. Are the Copenhagen Accord pledges sufficient to limits global warming to 2C or 1.5C?*

UNFCCC, 2011. Decision 1/CP.16 The Cancun Agreements: Outcome of the work of the ad hoc working group on Long-term Cooperative Action under the Convention.

UNFCCC, 1998. Kyoto Protocol to the United Nations Framework Convention on Climate Change.

UNFCCC, 1992. *United Nations Framework Convention on Climate Change*.

UNISDR, 2009. *Terminology on Disaster Risk Reduction*.

US Department of Energy, 2011. *Critical Materials Strategy*.

US Geological Survey, 2013. *Mineral Commodity Summaries 2013*.

US Geological Survey, 2010. *Mineral Commodity Summaries 2010*.

United Nations General Assembly, 2010. Resolution 64/292. The human right to water and sanitation. , 1249(20378).

Unruh, G.C., 2002. Escaping carbon lock-in. *Energy Policy*, 30, pp.317-325.

Unruh, G.C., 2000. Understanding carbon lock-in. *Energy Policy*, 28, pp.817-830.

Unruh, G.C. & Carrillo-Hermosilla, J., 2006. Globalizing carbon lock-in. *Energy Policy*, 34, pp.1185-1197.

Usher, W. & Strachan, N., 2013. An expert elicitation of climate, energy and economic uncertainties $. *Energy Policy*, 61, pp.811-821.

Vatn, A., 2005. Rationality, institutions and environmental policy. *Ecological Economics*, 55, pp.203 - 217.

Verbong, G. & Loorbach, D., 2012. Introduction. In G. Verbong & D. Loorbach, eds. *Governing the Energy Transition. Reality, illusion or necessity?* Routledge, pp. 1-23.

Vestas, http://www.vestas.com/en/about-vestas/sustainability/sustainable-products/life-cycle-assessment/rare-earth-elements-1.aspx.

Voß, J.-P., Smith, A. & Grin, J., 2009. Designing long-term policy: rethinking transition management. *Policy Sciences*, 42, pp.275-302.

de Vries, E., 2011. Magnet price hikes hit direct drive's future. *Wind Power Monthly*. Available at: http://www.windpowermonthly.com/article/1101036/magnet-price-hikes-hit-direct-drives-future [Accessed March 7, 2014].

WRAP, 2011. *A methodology for quantifying the environmental and economic impacts of reuse.*

Walker, G. & Devine-Wright, P., 2008. Community renewable energy: What should it mean? *Energy Policy*, 36, pp.497-500.

Walker, G. et al., 2010. Trust and community: Exploring the meanings , contexts and dynamics of community renewable energy. *Energy Policy*, 38(6), pp.2655-2663.

Walker, G., 2011. The role for "community" in carbon governance. *Environment*, pp.777-782.

Walker, W., Haasnoot, M. & Kwakkel, J.H., 2013. Adapt or Perish: A Review of Planning Approaches for Adaptation under Deep Uncertainty. *Sustainability*, 5, pp.955-979.

Water Company Manager, 2013. Interview about Water Company demand management pilot trials.

Water UK, 2014. UK water and wastewater companies. Available at: http://www.water.org.uk/home/our-members/list-of-companies/water-and-sewerage [Accessed March 7, 2014].

Water UK, 2009. How the water industry is managing its contribution to climate change. Available at: http://www.water.org.uk/home/policy/climate-change/mitigation#curr [Accessed September 29, 2013].

Waterwise, 2011. *Evidence Base for Large-scale Water Efficiency Phase II Final Report April 2011*.

Watson, J., 2008. Setting Priorities in Energy Innovation Policy: Lessons for the UK..

Wellmer, F.W. & Dalheimer, M., 2012. The feedback control cycle as regulator of past and future mineral supply. *Mineralium Deposita*, 47, pp.713-729.

Whitmarsh, L., 2012. How useful is the Multi-Level Perspective for transport and sustainability research? *Journal of Transport Geography*.

Wilkinson, P., Smith, K.R. & Haines, A., 2007. Energy and Health 1: A global perspective on energy: health effects and injustices. *The Lancet*, 370, pp.965-78.

Willis et al., 2002. Water Companies' Service Performance and Environmental Trade-offs. *Journal of Environmental Planning and Management*, 45(3), pp.363-379.

Wilson, C. et al., 2012. Marginalization of end-use technologies in energy innovation for climate protection. *Nature Climate Change*, 2(11), pp.780-788.

Wyatt, A., 2013. Interview about Kimberley Clarke water recycling.

Yadoo, A., Gormally, A. & Cruickshank, H., 2011. Low-carbon off-grid electrification for rural areas in the United Kingdom: Lessons from the developing world. *Energy Policy*, 39(10), pp.6400-6407.

Yang, 2012. Situation and policies for China's rare earth industry. *Xinhuanet*. Available at: http://news.xinhuanet.com/english/bilingual/2012-06/20/c_131665890.htm.

Yin, R.K., 2009. *Case study research: design and methods* 4th ed., London: Sage.

Yu, C., Brown, R & Morison, P. 2012. Co-governing decentralised water systems: an analytical framework. , 66(12), pp.2731-2736.

Zegeye, A., 2000. US Public Infrastructure and Its Contribution to Private Sector Productivity.

Appendix 1: Supporting data for case study analysis

Appendix 1.1: Case study interviews

Details of individual case interviews a provided in table A1.1

Table A1.1.1 Summary of individual case interviews

Individual case	**Interviewee**	**Job title**	**Date**
Olympic park energy centre	Mike Carr	Director of Engineering	15.03.12
Water company Water Saving Trials	Not disclosed	Water company manager	22.05.12
Kimberley Clark Water Recycling	Alan Wyatt	Technical Director	22.06.12
Eco-Island	David Green[32]	Eco-Island Director	11.02.13
	Andy Stanford Clark	Master Inventor and consultant to Eco-Island	11.02.13
	Chris Cooper	Smarter cities architect	11.02.13
Welborne	Glynn Bench	Landowner	12.02.13
	David Pepper	Senior Partner	12.02.13
	Simon Ward	Chartered Surveyor	12.02.13
	Chris Cooper	Smarter cities architect	12.02.13
Chale Green Housing Trust	Andy Stanford Clark	Master Inventor and consultant to Chale Green	11.02.13
Newcastle City Council	Allen Jones	Energy Masterplanner	13.05.13
Byker Community Trust	Richard Beedle	District heating project manager	31.01.14
	Allen Jones	Energy Masterplanner	29.01.14

---

[32] Interview data not used from this interviewee

Appendix 1.2: Individual case write-up example: Water Demand Management Pilot Trial

A1.2.1 Sector Summary

The UK water industry is a fully privatised utility sector but is highly regulated by Ofwat, the UK water regulation body, with tight control over water prices and investment spending. The price for water is regularly reviewed and fixed for a five year period, making the profit that water companies generate highly depending on the amount of water sold. Investment in asset maintenance and construction is planned on the same five-yearly cycle, with only expenditure approved by OFWAT being allowed in any period.

The water sector is also highly connected to other sectors at the end user – with 18% of household energy use being for water heating (DECC 2012a) – and at the infrastructure level – it is the fourth most energy intensive sector (CST 2009b) and energy intensity is rising as water quality and waste water treatment standards become increasingly stringent.

A1.2.2 Individual case analysis

Some of the key drivers and characteristics of the water demand management case are described below, grouped according to the five systems in Foxon's co-evolutionary framework. The case consisted of a water saving trial carried out by one of the major water companies in response to Ofwat requirements. A typical trial consisted of identifying a population of metered houses and approaching residents offering free installation of water-saving measures including toilet cistern inserts, shower-flow reducing devices, tap inserts, shower timers, tooth-brushing timers. Water saving was monitored for an extended period after retrofitting to demonstrate saving potential. Average reductions of up to 21 litres per property per day were demonstrated as well as carbon emissions reductions of between 0.031 and 0.187 kgCO_2e per property per day (Waterwise 2011).

Institutions

One of the key drivers for the demand management pilot trial was a duty, recently introduced by Ofwat, to promote water conservation and efficiency. This has resulted in a number of water companies undertaking domestic water saving trials to demonstrate compliance with this duty and determine the effectiveness of a range on interventions for use in estimates of future water saving potential (Waterwise 2011). However, such innovation and research spending is currently completely voluntary for water businesses and regulatory intervention only occurs if certain efficiency targets, such as leakage, are failed and the company is put on special measures. This is part of Ofwat's light-touch approach where it aims to provide a policy framework for companies to operate within.

While Ofwat promotes water efficiency, the current licensing system itself creates a barrier to water conservation. At the moment license holders, e.g. water companies or farmers, can take an amount of water regardless of the current groundwater level or societal requirements (HM Government 2011b)). Licensees are not encouraged to reduce the amount of water that they abstract or stop to consider whether their own requirements should be given priority over others. "It is a very insensitive system" (Water Company Manager 2013). Already at the point of abstraction, demand management is difficult to enforce. Defra are currently aiming to attach a value to abstraction and thereby a value to water itself. In its White Paper, Water for Life, Defra advocated the idea of changing the licensing system for water abstraction and effectively creating a market for water abstraction (HM Government 2011b). "Once there is a value attached to the amount of water that is actually abstracted that will encourage demand management" (Water Company manager 2013). The missing value on water is a significant barrier for water efficiency. It is also the reason why metering of water usage is currently not supported by a viable business case. Although water companies have the statutory authority to put meters into properties at a change of occupancy this power is hardly executed. Alternatively, occupants may ask for a meter to be installed in their property. Again, the economic considerations act as a barrier to what could be a water efficiency driver as it would make consumption measurable and thereby make end-users accountable for their demand. It would change the pricing from "having a licence with a flat, meaningless low figure attached to it" (Water Company Manager 2013) to a precise measuring exercise for individual properties.

While the regulator supervises the water companies rather than controlling them, Ofwat does not currently allow water companies to extend their business pursuits beyond water-related operations. Historically, Ofwat left more liberties which enabled some companies to start integrating alternative forms of energy generation outside of hydro into their business models: "But [...] suddenly they got a sniff of people or companies which started to put loads and loads of wind turbines into their business plans and they probably got frightened a little bit and they thought, well we'll be an energy regulator now and not a water regulator. We have got to put a stop to it" (Water Company Manager 2013).

Technology

The technology used to reduce water usage tends to be basic efficiency devices. Retrofitting devices come in the form of add-ons to existing white goods or appliances and do not tend to alter the way water is used. “Although what the water efficiency alternatives are, like water tank on the toilet or something to screw on your shower head or onto your tap – I don’t think they come a lot more sophisticated than that. I think they are cheap retrofit type things” (Water Company Manager 2013). Generally, the technology has been described as unsophisticated have little effect on water consumption. Investment into innovation and research in required to yield better efficiency results in the future and make water efficiency worthwhile for the water businesses.

Furthermore, the technology to save water through leakage reduction is more cost effective and generates a more reliable water saving than investment in end-user efficiency. Water companies are more familiar with this type of capital-intensive, fit and forget technology, than with end-user technology, which require not only distribution but also end-user engagement.

Business strategies

Many water companies have ended pro-active retrofitting, and replaced it with passive advertisement of free water saving devices, self-audits and general water savings advice. When compared with other water companies, the Water Company interviewed offered the most comprehensive selection of technologies but this approach is nevertheless reactive rather than proactive and a direct result of satisfying the regulator. The Water Company is not directly approaching and contacting its customers in order to encourage water conservation. Rather, it provides information and access to devices and advice on its own and partner websites.

From a co-evolutionary point of view, it is evident that water utility businesses evolve in reaction to regulatory changes and requirements: “Ofwat is key or king rather, and we respond to what the regulator wants us to do” (Water Company Manager 2013). This strong interdependency between the institutional dimension and business model leaves little opportunity for UK water companies to implement changes based on ecological, technological or societal requirements or considerations. “At the end of the day, especially with the march towards more universal metering, it is not exactly in their interest to promote water efficiency”; “They are selling water at the end of the day; the more they sell the more profits they make. It is a little bit of a catch 22 for them” (Water Company Manager 2013). The only other driver that a water company can have in promoting conservation is to get this balance between supply and demand right from an economic perspective. “Let’s pretend that it is cheaper to promote water conservation, to send out packs and all the rest of it to get people to choose water consumption in the home than it is to develop new resources. That balance is not there at the moment" (Water Company Manager 2013).

The ownership of the Water Company had a significant influence over decisions; "And their approach was very much hands-off. They saw what they thought was a successful business model, providing a return. And they wanted the return. It was only when [the Water Company] started failing things, like leakage targets and stuff like that, that they actually started to get more involved in things"; "I would say that since it has been bought out by investment banks it has been delisted, that has started to change a little bit. Innovation is no longer seen as a long-term priority. The business is all about short-term, quick-fix solutions"; "A lot of companies invest very little in innovation. And that is mainly because of their ownership. If you are in it for the short-term, then you are really only interested in the returns. So why would you want to invest into long-term R&D and innovation. And a lot of companies are taking that route" (Water Company Manager 2013).

The Water Company interviewed, however, did demonstrate some innovations in its business model within these tight regulatory and financial constraints:

- It employs two different water conservation strategies for domestic and business customers in line with its business priorities: reactive provision of retrofit water efficiency devices for domestic customers to protect base water sales; and proactive provision of consultancy services to business customers to protect existing business relations;
- It focuses on core business activities to reduce water wastage through leakage but simultaneously explores additional opportunities to stabilise or increase revenue streams;
- It carefully balances short-term financial gains with long-term innovation benefits through strategic investment and funding programmes;
- It favours demand management and leakage prevention over the development of new water sources, but not to the point where it starts to damage the business.

User practices

Of course water consumption cannot be tackled through technologies and business strategies alone. The end-users need to be deeply involved as well. Customer's perceptions significantly impact the use or conservation of water. Unfortunately, living in a rainy and often wet country does not inspire end-users to save water and more information and education is necessary to raise awareness of the scarcity of water and to allow end-users "to understand this water situation seriously" (Water Company Manager 2013).

In addition to that, research in the water industry has proven that conservation psychology plays against continuous and long-term water conservation (Waterwise 2011). It showed that immediately after the installation of water saving devices or water meters consumption levels tend to falls. “But then over a longer period, it actually returns back to the level where it was previously” (Water Company Manager 2013). Keeping consumption at a low level would take continuous encouragement, incentivisation and interaction between the Water Company and end-users.

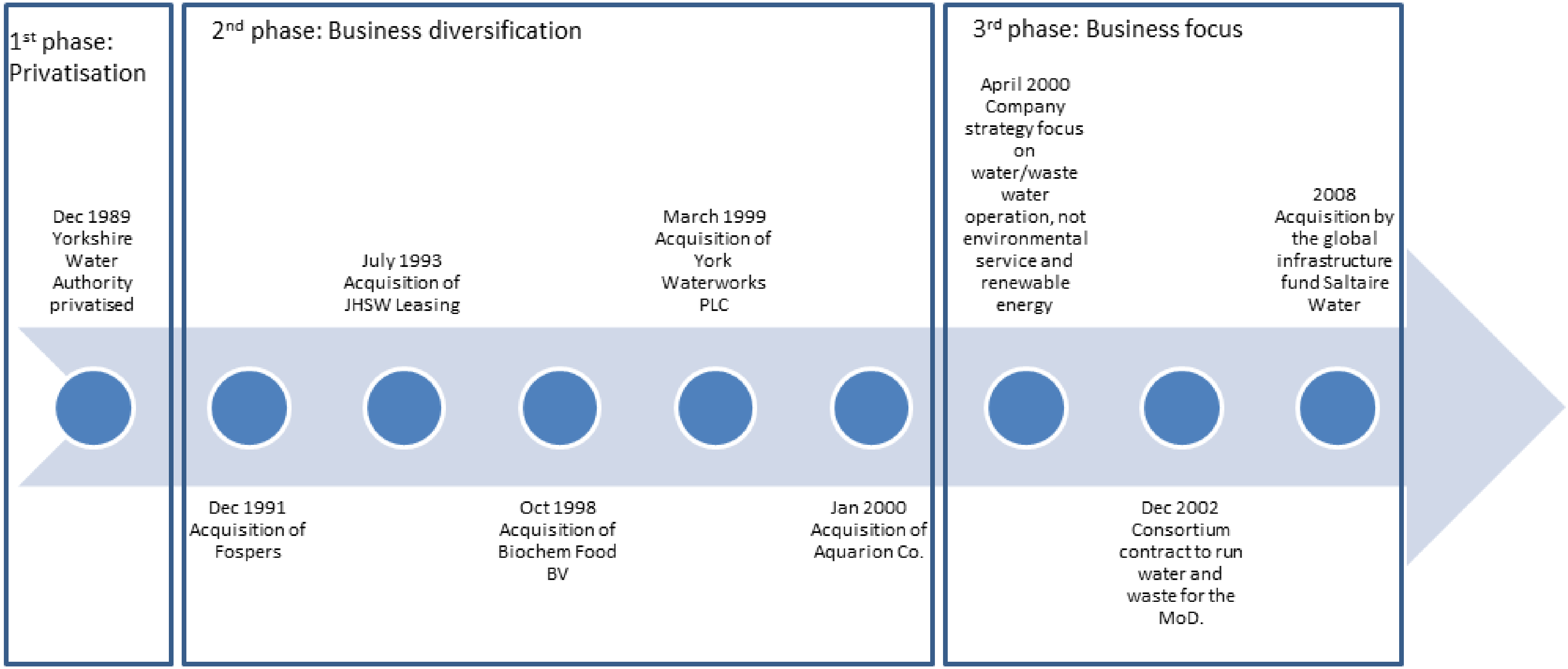
Environment

The duty to conserve water has been placed on water companies because of recent extreme weather and the pressure on water resources, even in historically wet regions. As a result, there has been increasing focus on the balance between supply and demand. It is forecast that this extreme weather will become increasingly prevalent (IPCC 2013) and thus a strong driver of demand management in the future.

A1.2.3 Case timeline

The historical context of the case study is summarised in figure A1.2.1

Figure A1.2.1



A1.2.4 Barriers to scaling up water demand management

The analysis above identifies a series of barriers to increasing the scale of water demand management activities. Despite the increasing pressure from Ofwat water companies need to overcome numerous barriers to make water conservation a business priority. Demand management is unattractive in the current regulatory environment because it is a costly exercise to provide water efficiency devices and reduces income from water sales (in metered customers), which would reduce a water company's profits substantially. It is not possible to create alternative revenue streams though increasing prices or offering complementary services as a result of regulatory controls.

The economics of the supply and demand balance are also not in favour of water efficiency efforts. If faced with a choice to promote water efficiency or alternatively to reduce leakage or develop new water sources, the latter two are currently the cheapest and most reliable options for water companies to provide reduce operating costs and improve reliability of water utility services.

As already mentioned, investment plans are reviewed and fixed on a five-year rota by Ofwat. This limits the ability of water companies to invest in schemes, such as demand management off-setting the need to invest in new resources, which pay back over a period longer than this 5 year rotation. In many cases this means that business areas like innovation or R&D receive little or no financial support in the water sector.

The ownership of water companies has significant implications for investment strategies, innovation and diversification. After privatisation, a number of companies expanded operations into waste management and energy generation, exploiting the significant interconnectivity between these sectors. However, as investment banks have taken over ownership, attitudes to risk have changed significantly (in addition to negative experiences of diversification) and businesses refocused on core assets to reduce risk and deliver steady profits.

Water conservation technologies tend to be unsophisticated add-ons to existing white goods or appliances and do not alter the way water is used. Rather, they squeeze a few litres off the overall consumption. Pilot trials show that water savings reduce over time (Waterwise 2011), reducing the confidence in the scale of water savings available and the associated investment justifications.

Water conservation relies heavily on changes in end-user behaviour. This poses a significant challenge because users do not perceive there is a need to reduce water use as a result of the apparent abundance of water. This is compounded by the fact that the majority of water customers are not metered and water prices are relatively cheap (compared to, say energy prices).

The regulator's intervention into water innovation and investment is limited. Innovation and research spending is currently completely voluntary for water businesses and regulatory intervention only occurs if certain efficiency targets, such as leakage, are failed and the company is put on special measures. This is part of Ofwat's light-touch approach where it aims to provide a policy framework for companies to operate within.

While Ofwat promotes water efficiency, the current licensing system itself creates a barrier to water conservation. At the moment license holders, e.g. water companies or farmers, can take any amount of water regardless of the current groundwater level or societal requirements. Licensees are not encouraged to reduce the amount of water that they abstract or stop to consider whether their own requirements should be given priority over others. Defra is currently aiming to attach a value to abstraction and thereby a value to water itself. In its White Paper, Water for Life, Defra is advocating the idea of changing the licensing system for water abstraction and effectively creating a market for water abstraction (HM Government 2011b). The missing value on water is a significant barrier for water efficiency. It is also the reason why metering of water usage is currently not supported by a viable business case. Although water companies have the statutory authority to put meters into properties at a change of occupancy this power is hardly executed. Again, the economic considerations act as a barrier to what could be a water efficiency driver as it would make consumption measurable and thereby make end-users accountable for their demand.

A1.2.5 Existing responses to barriers

The Water Company analysed has managed to work around investment payback restrictions by designing and implementing a 25 year innovation programme which is structured in two parts: "The first was aiming to deliver short-term quick-win benefits that provided business return [...]. And then the other part of it was to create a longer term innovation programme or strategy that started to do the more fundamental type of work and the more fundamental research that will ultimately feed through into business benefits" (Water Company Manager 2013). This approach convinced the internal business managers as well as external investors and the regulator Ofwat of the viability of utility innovation.

In contrast to its reactive approach to domestic users, the Water Company supports business customers through its proactive consultancy services provided by its specialised business customer department. Large business customers who consume large quantities of water can receive a service whereby water experts from the Water Company visit their offices and sites and help reduce the customer's water consumption and save them money. While this reduces the Water Company's own profit, they generate an additional income stream through water consulting. At the same time, they deepen customer relationships, create business opportunities for additional service provision and, most importantly, they create long-term business contracts. The Water Company thereby reduces its own risk which derives from business competition through other water companies. Business customers are entitled by the Competition Act to choose their water provider freely. Therefore, it is in the interest of the Water Company to maintain long-term business relationships. The Water Company benefits because its supply infrastructure is optimised around the water demand of large-scale business users. Long-term business interactions allow it to tailor consultancy advice and water services to the unique requirements of the business customer and turn themselves into a valuable and indispensable business partner. The customers in turn save on their water bills and increase their environmental credentials through increasing resource efficiency.

A1.2.6 Emerging insights

The analysis above has resulted in a series of case-specific insights in relation to the characteristics which differ from the mainstream and constraints from policy and regulation that might constrain the roll out of demand management. These case-specific insights will be compared to other individual cases to identify case-study-wide characteristics and constraints.

Characteristics

- The value of demand management is derived from off-setting investment in new supply; in some cases this can be sufficient to motivate significant investment in demand management is water scarce regions.
- Demand management is promoted when it can be used to improve business customer relations and retain important customers.
- Targeted, area-based demand management programmes can be successful but require significant resource to engage end-users.
- It is proposed that partnering water efficient retrofit with energy efficient retrofit could improve the uptake rates of water demand management (Waterwise 2011).
- Alternative financial models, bundling investments to create overall payback time acceptable to regulators, have been a successful means of delivering innovations and research.

Constraints

- The water industry is highly regulated by Ofwat, controlling water prices and investment spending so demand management reduces revenue.
- Technology fix is ineffective without support to change end-user behaviour.
- Cheap water prices, lack of metering and perception of abundance further limit behaviour change.
- Ofwat intervenes directly in water efficiency only if leakage targets are significantly failed, there are no fines if demand management is not implemented.
- The implementation of water efficiency regulation as a <u>duty to promote</u> water conservation, rather than formal water saving targets, created passive and reactive availability of water saving devices and advice for domestic customers, which have limited effect.

Appendix 2: Material Criticality Analysis Data

Appendix 2: Material Criticality Analysis Data

A2.1 Forecast of production and requirements

A2.1.1 Future requirements forecast (R(t)) for neodymium – use in low-carbon technology

The future requirements for neodymium in low-carbon technologies is estimated using scenarios from the International Energy Agency's Technology Roadmap scenario to 2050 (IEA 2012). The technology roll out estimates are presented in table A2.1 and the associated neodymium requirements are presented in table A2.2.

Table A2.1: Technology roll out scenario (from International Energy Agency Technology Roadmap scenario to 2050 (IEA 2012)**)**

	EV/PHEV		ICE		Onshore (GW)		Offshore (GW)	
Year	Market Share (%)	Sales	Market Share (%)	Sales	Capacity	Annual Installed	Capacity	Annual Installed
2012	2.2	1,880,833	97.8	83611604	260	38	24.3	10.6
2013	2.8	254456	97.2	88332747	298	38	34.9	10.6
2014	3.4	3,272,914	96.6	92,989,271	336	38	45.5	10.6
2015	4	4,065,882	96	97,581,177	373	38	56.1	10.6
2016	4.6	4,923,468	95.4	102,108,464	411	38	66.6	10.6
2017	5.2	5,845,673	94.8	106,571,133	449	38	77.2	10.6
2018	5.8	6,832,497	94.2	110,969,183	486	38	87.8	10.6
2019	6.4	7,883,939	93.6	115,302,615	524	38	98.4	10.6
2020	7	9,000,000	93	119,571,429	562	38	109	10.6
2021	9.3	11,691,428	90.7	114,022,857	589	26.8	117	8.5
2022	11.6	14,251,428	88.4	108,605,714	616	26.8	126	8.5
2023	13.9	16,680,000	86.1	103,320,000	642	26.8	134.5	8.5
2024	16.2	18,977,142	83.8	98,165,714	669	26.8	143	8.5
2025	18.5	21,142,857	81.5	93,142,857	696	26.8	151.5	8.5
2026	20.8	23,177,142	79.2	88,251,428	723	26.8	160	8.5
2027	23.1	25,080,000	76.9	83,491,428	750	26.8	168.5	8.5
2028	25.4	26,851,428	74.6	78,862,857	776	26.8	177	8.5
2029	27.7	28,491,428	72.3	74,365,714	803	26.8	185.5	8.5
2030	30	30,000,000	70	70,000,000	830	26.8	194	8.5
2031	31.8	33,257,500	68.2	71,325,833	868	37.6	211.2	17.2
2032	33.6	36,680,000	66.4	72,486,666	905	37.6	228.4	17.2
2033	35.4	40,267,500	64.6	73,482,500	943	37.6	245.6	17.2
2034	37.2	44,020,000	62.8	74,313,333	980	37.6	262.8	17.2
2035	39	47,937,500	61	74,979,166	1018	37.6	280	17.2
2036	40.8	52,020,000	59.2	75,480,000	1056	37.6	297.2	17.2

	EV/PHEV		ICE		Onshore (GW)		Offshore (GW)	
Year	Market Share (%)	Sales	Market Share (%)	Sales	Capacity	Annual Installed	Capacity	Annual Installed
2037	42.6	56,267,500	57.4	75,815,833	1093	37.6	314.4	17.2
2038	44.4	60,680,000	55.6	75,986,666	1131	37.6	331.6	17.2
2039	46.2	65,257,500	53.8	75,992,500	1168	37.6	348.8	17.2
2040	48	70000,000	52	75,833,333	1206	37.6	366	17.2
2041	49.2	72,775,000	50.8	75,141,666	1222	15.8	394.6	28.6
2042	50.4	75,600,000	49.6	74,400,000	1286	15.8	423.2	28.6
2043	51.6	78,475,000	48.4	73,608,333	1254	15.8	451.8	28.6
2044	52.8	81,400,000	47.2	72,766,666	1269	15.8	480.4	28.6
2045	54	84,375,000	46	71,875,000	1,285	15.8	509	28.6
2046	55.2	87,400,000	44.8	70,933,333	1,301	15.8	537.6	28.6
2047	56.4	90,475,000	43.6	69,941,666	1,317	15.8	566.2	28.6
2048	57.6	93,600,000	42.4	68,900,000	1,332	15.8	594.8	28.6
2049	58.8	96,775,000	41.2	67,808,333	1,348	15.8	623.4	28.6
2050	60	100,000,000	40	66,666,666	1,364	15.8	652	28.6

Table A2.2: Neodymium requirements for global low carbon technology roll out and other uses 2010-2050 (from International Energy Agency Technology Roadmap scenario to 2050 (IEA 2012)**)**

Future Demand Projection (tons Nd)

Year	**EV/PHEV**	**Onshore Wind**	**Offshore Wind**	**ICE Vehicles**	**Other**	**TOTAL**
2012	2,821	750	211	5,179	24,758	33,719
2013	3,817	812	229	5,471	25,501	35,829
2014	4,909	875	246	5,760	26,266	38,056
2015	6,099	937	264	6,044	27,054	40,397
2016	7,385	1,000	281	6,325	27,865	42,856
2017	8,769	1,062	299	6,601	28,701	45,432
2018	10,249	1,124	316	6,873	29,562	48,125
2019	11,826	1,187	334	7,142	30,449	50,938
2020	13,500	1,249	352	7,406	31,363	53,870
2021	17,537	912	289	7,063	32,304	58,104
2022	21,377	934	296	6,727	33,273	62,607
2023	25,020	956	303	6,400	34,271	66,950
2024	28,466	979	310	6,080	35,299	71,134
2025	31,714	1,001	317	5,769	36,358	75,160
2026	34,766	1,023	325	5,466	37,449	79,028
2027	37,620	1,045	332	5,171	38,572	82,741
2028	40,277	1,068	339	4,885	39,729	86,298

Future Demand Projection (tons Nd)

Year	EV/PHEV	Onshore Wind	Offshore Wind	ICE Vehicles	Other	TOTAL
2029	42,737	1,090	346	4,606	40,921	89,700
2030	45,000	1,112	353	4,336	42,149	92,950
2031	49,886	1,654	757	4,418	43,413	100,128
2032	55,020	1,748	799	4,490	44,716	106,773
2033	60,401	1,841	842	4,551	46,057	113,693
2034	66,030	1,935	885	4,603	47,439	120,892
2035	71,906	2,029	928	4,644	48,862	128,369
2036	78,030	2,122	971	4,675	50,328	136,126
2037	84,401	2,216	1,014	4,696	51,838	144,164
2038	91,020	2,309	1,056	4,707	53,393	152,485
2039	97,886	2,403	1,099	4,707	54,995	161,090
2040	105,000	2,497	1,142	4,697	56,644	169,980
2041	109,163	1,075	1,947	4,654	58,344	175,182
2042	113,400	1,102	1,994	4,608	60,094	181,198
2043	117,713	1,128	2,041	4,559	61,897	187,338
2044	122,100	1,154	2,089	4,507	63,754	193,604
2045	126,563	1,180	2,136	4,452	65,666	199,998
2046	131,100	1,206	2,184	4,394	67,636	206,520
2047	135,713	1,233	2,231	4,332	69,666	213,174
2048	140,400	1,259	2,279	4,268	71,756	219,961
2049	145,163	1,285	2,326	4,200	73,908	226,882
2050	150,000	1,311	2,374	4,129	76,125	233,940

A2.1.1 Future production forecast (M(t)) for neodymium

The future production of neodymium is forecast by projecting the historical, linear, trend in production growth from 1990 (US Geological Survey 2012).

Table A2.3: Forecast REO and neodymium production

Year	REO (tons)	Nd (tons)
2012	137993	24839
2013	142406	25633
2014	146819	26427
2015	151232	27222
2016	155645	28016
2017	160058	28810
2018	164471	29605
2019	168884	30399

2020	173297	31193
2021	177710	31988
2022	182123	32782
2023	186536	33576
2024	190949	34371
2025	195362	35165
2026	199775	35960
2027	204188	36754
2028	208601	37548
2029	213014	38343
2030	217427	39137
2031	221840	39931
2032	226253	40726
2033	230666	41520
2034	235079	42314
2035	239492	43109
2036	243905	43903
2037	248318	44697
2038	252731	45492
2039	257144	46286
2040	261557	47080
2041	265970	47875
2042	270383	48669
2043	274796	49463
2044	279209	50258
2045	283622	51052
2046	288035	51846
2047	292448	52641
2048	296861	53435
2049	301274	54229
2050	305687	55024

.

A2.2 Forecasting production distribution

Table A2.4 presents the results of interpolating the distribution of neodymium production from current production (US Geological Survey 2012) to the distribution of neodymium reserves. The distribution of production of neodymium is assumed to be the same as the distribution of rare earth elements.

Table A2.4: Forecast production distribution for neodymium

	2011	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**	**2025**
Australia	0	9	17	26	35	43	52	61	69	78	87	95	104	113	121
Brazil	99	97	94	92	90	88	85	83	81	78	76	74	72	69	67
China	23,400	23,098	22,796	22,494	22,192	21,890	21,588	21,286	20,984	20,682	20,380	20,078	19,776	19,474	19,172
CIS	0	103	206	309	412	515	618	721	824	927	1,030	1,132	1,235	1,338	1,441
India	540	543	546	549	552	555	558	561	564	567	570	572	575	578	581
Malaysia	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6
USA	0	70	141	211	282	352	423	493	564	634	704	775	845	916	986
Canada	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Greenland	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Namibia	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Kenya	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Tanzania	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Angola	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Mauritania	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Burundi	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Malawi	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Vietnam	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Thailand	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Indonesia	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Finland	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Sweden	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111
Turkey	0	8	16	24	32	40	48	56	64	72	79	87	95	103	111

Table A2.4: Forecasting production distribution (cont.)

	2026	**2027**	**2028**	**2029**	**2030**	**2031**	**2032**	**2033**	**2034**	**2035**	**2036**	**2037**	**2038**	**2039**	**2040**
Australia	130	139	147	156	165	173	182	191	199	208	217	225	234	243	251
Brazil	65	63	60	58	56	53	51	49	47	44	42	40	37	35	33
China	18,870	18,568	18,266	17,964	17,662	17,361	17,059	16,757	16,455	16,153	15,851	15,549	15,247	14,945	14,643
CIS	1,544	1,647	1,750	1,853	1,956	2,059	2,162	2,265	2,368	2,471	2,574	2,677	2,780	2,883	2,986
India	584	587	590	593	596	599	602	605	608	611	614	617	620	623	626
Malaysia	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
USA	1,057	1,127	1,198	1,268	1,338	1,409	1,479	1,550	1,620	1,691	1,761	1,832	1,902	1,972	2,043
Canada	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Greenland	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Namibia	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Kenya	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Tanzania	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Angola	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Mauritania	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Burundi	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Malawi	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Vietnam	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Thailand	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Indonesia	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Finland	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Sweden	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230
Turkey	119	127	135	143	151	159	167	175	183	191	199	207	215	223	230

Table A2.4: Forecasting production distribution (cont.)

	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050
Australia	260	269	277	286	295	303	312	321	329	338
Brazil	31	28	26	24	22	19	17	15	12	10
China	14,341	14,039	13,737	13,435	13,133	12,831	12,529	12,227	11,925	11,623
CIS	3,089	3,192	3,295	3,397	3,500	3,603	3,706	3,809	3,912	4,015
India	629	632	634	637	640	643	646	649	652	655
Malaysia	6	6	6	6	6	6	6	6	6	6
USA	2,113	2,184	2,254	2,325	2,395	2,465	2,536	2,606	2,677	2,747
Canada	238	246	254	262	270	278	286	294	302	310
Greenland	238	246	254	262	270	278	286	294	302	310
Namibia	238	246	254	262	270	278	286	294	302	310
Kenya	238	246	254	262	270	278	286	294	302	310
Tanzania	238	246	254	262	270	278	286	294	302	310
Angola	238	246	254	262	270	278	286	294	302	310
Mauritania	238	246	254	262	270	278	286	294	302	310
Burundi	238	246	254	262	270	278	286	294	302	310
Malawi	238	246	254	262	270	278	286	294	302	310
Vietnam	238	246	254	262	270	278	286	294	302	310
Thailand	238	246	254	262	270	278	286	294	302	310
Indonesia	238	246	254	262	270	278	286	294	302	310
Finland	238	246	254	262	270	278	286	294	302	310
Sweden	238	246	254	262	270	278	286	294	302	310
Turkey	238	246	254	262	270	278	286	294	302	310