Governance, Innovation and the Transition to a Sustainable Energy System: Perspectives from Economic Theory

Matthew Lockwood

EPG Working Paper: 1305

Abstract:
This paper reviews basic ideas in economic theory about the governance of innovation, with applications to debates in innovation in sustainable energy. The aim is to extract a few broad issues to frame comparative analysis and the analysis of change. The review is motivated by the relevance of economic theory to innovation policy, by the fact that economics remains the dominant language of policy makers, and by gaps between the economic treatment of innovation in policy and that in economic theory. Much of the debate about low-carbon energy innovation focuses quite narrowly on innovation policy. This is important, and is discussed in the first part of the review. However, some of the more interesting insights about innovation in economics come from consideration of the wider context for innovation. In particular, there are important literatures on market structure and innovation, on increasing returns and technology adoption, and on institutions and innovation. The paper concludes with a research agenda for understanding differential comparative performance in innovation in sustainable energy, and for how to accelerate such innovation in the UK.

Keywords: Economic theory, sustainability, institutions, innovation, governance, lock-in

Contact: Matthew Lockwood M.Lockwood@exeter.ac.uk

Date: July 2013
Contents

1. Introduction ........................................................................................................................................3

2. The Governance of Innovation ........................................................................................................6
   2.1 The economic case for innovation policy ......................................................................................6
   2.2 Government failure and the innovation policy debate .................................................................11
   2.3 Innovation policy as a joint discovery process ............................................................................15
   2.4 Conclusions ................................................................................................................................17

3. The Governance of Markets .............................................................................................................18
   3.1 Two views on innovation and market structure ..........................................................................18
   3.2 Innovation and factors of competition ........................................................................................20
   3.3 Innovation in retail energy markets and the debate over competition ......................................23
   3.4 Conclusions ................................................................................................................................26

4. The Governance of Systems ............................................................................................................27
   4.1 Increasing returns, path dependence and lock-in ......................................................................27
   4.2 Low carbon technology races and carbon lock-in .................................................................31
   4.3 Escaping carbon lock-in ............................................................................................................33
   4.4 Conclusions ................................................................................................................................35

5. The Governance of Institutions ....................................................................................................36
   5.1 Institutions and institutional change .........................................................................................36
   5.2 Varieties of capitalism and low carbon innovation ..................................................................38
   5.3 Conclusions ................................................................................................................................41


7. References .........................................................................................................................................44
1. Introduction

The transition to a low carbon economy is one of the major challenges of the 21st century. In the case of the UK, the pace of change to a more sustainable energy system has been quite slow, especially relative to the speed of transformation that will be required over the next two decades (Committee on Climate Change 2012). In some respects, for example in the development of new renewable energy and in energy saving, the UK also lags other countries such as Germany, Denmark and even parts of the US.

To accelerate the pace of change in the energy sector, greater innovation is needed, including not only the development of new technologies, but also new infrastructures, new business models, new services, and new institutions. The question is then how greater innovation in the energy sector may be fostered. In a broad sense the answers to this question must lie with the way that innovation is governed, i.e. the policies and institutions that provide the context for innovative activities by firms, individuals and other actors, and with changes in that governance.

There are a wide range of approaches to understanding this issue, including those from socio-technical and political frameworks (cf. Meadowcroft 2005, Kern 2010, Kuzemko 2013). However, the focus of this paper is on what a number of different strands of economic theory say about how governance arrangements affect innovation, and how governance can and should be changed.

There are a number of reasons for this focus. One is that economic theory has quite a lot to say on the subject; innovation has been an important theme in economics over the last fifty years, and a considerable amount of thinking about why and how it happens, and about related policies has been done. A second reason is that despite many critiques of its dominance (and its failure to anticipate the financial crisis), economics remains the principal language of policy makers. A third reason is to do with the perception of what economic theory says in the energy debate, and in particular the perception that there is a major gap in the approaches to innovation policy taken in economics (or at least the economics used by policy makers) and those taken in technology studies (e.g. Mitchell 2008: 21-61). Anticipating the conclusion of the review, I argue below the converse; that some types of economic theory lead to views on innovation policy that are quite similar to those taken by technology studies.

This last issue raises the question about how such perceptions come about. One factor is the fact that there is more than one view amongst economists, and in particular there are differing

---

1 I am grateful to Tom Foxon, Caroline Kuzemko and Catherine Mitchell for comments.
views about the operation and efficiency of markets relative to governments. In particular, a
distinction can be drawn between a neo-liberal view, in which markets are always superior to
intervention by governments, and a more balanced view in economics, in which both market
failure and government failure are seen as possibilities but not inevitabilities. This distinction
was perhaps muted for much of the past 25 years or so, but has come very much to the fore
again in the current debates in macroeconomics between neo-Keynesians and new classical
economists.\(^2\) There is also an irony that, over the period in which neo-liberal economics has
been dominant in the *policy* realm, economic *theory* has become more diverse, with approaches
such as institutional, evolutionary and behavioural economics moving into the mainstream.
Some of these latter frameworks provide an account of energy system governance and change
that come close to that of the technology studies literature.

The significance of reviewing what economic theory may have to say about innovation,
governance and the energy system transition therefore lies not only in its potential for
understanding the different performance of the UK relative to other countries, but also in the
possibility that it can provide a critique of the neo-liberal paradigm “from within”. However, this is
not to assume that economics will or should become the single framework for managing
transition to a sustainable energy system, not least because a politically-informed perspective is
still needed to understand why paradigms exist and how they are changed (Kuzemko 2013),
and energy systems have particular technological characteristics that make them different from
other parts of the economy. Nevertheless, looking at the range of what economics has to say
about the governance of innovation for energy transitions is useful.

In this paper I review three areas of economic thinking that apply to innovation. It draws on and
builds on earlier work on economics and low carbon innovation, especially Tim Foxon’s 2003
review for the Carbon Trust (Foxon 2003). Much of the debate about low-carbon energy
innovation focuses quite narrowly on *innovation policy*. While this is important, some of the more
interesting insights about innovation in economics come from consideration of the wider context
for innovation. In particular, there are important literatures on *market structure and innovation*,
and on *increasing returns, lock-in and escaping lock-in*. It should also be noted that the aim of
the review is *not* to provide a comprehensive account of the literature; rather, the focus is on the
*basic principles* underlying particular positions.

The review is organised as follows. In section 2, I start by reviewing the basic case for policy to
support innovation, with a focus on externalities and intellectual property rights. This literature

\(^2\) For a version of these debates see Paul Krugman’s essay “How did economists get it so wrong?”
works with a “black box” concept of innovation, and so its policy implications are also quite
genral. The more complex and realistic approach to innovation taken in some of the
economics of innovation literature is then discussed, in the context of debates about market
failures vs. government failures in the area of deployment support for renewable energy. This
discussion leads to the view that the process of policy formulation in itself may be as important
as the design of specific policies, and this point is linked to recent contributions to the theory of
industrial policy. It goes on to ask how far changes in the governance of this process may be
constrained by institutional systems, based on the implications of the institutional economics
literature.

Section 3 looks at theories of the relationships between competition, market power and
innovation, drawing on neo-Schumpeterian debates, and assessing their implications for the
governance of energy markets. The literature suggests that the relationship between market
structure and innovation depends very much on specific factors of competition, such as product
substitutability, intellectual property protection and ease of market entry. The last of these
issues is probably the most important from the point of view of governing markets to encourage
more innovation in energy markets. These ideas are then explored in the case of retail energy
markets in the UK.
The scope is widened further in section 4, which examines the implications of evolutionary
economics for understanding energy systems and system change. Evolutionary accounts of
technology choice emphasise the role of increasing returns, including network externalities, and
the importance of path-dependence, that lead to lock-in at several levels. This goes on to
examine the debate on the role of policy in avoiding or undoing lock-on, with discussion of how
they might apply in cases such as electricity grids and wholesale market reform.

The literature on escaping lock-in implies that the challenge of overcoming coordination
problems lies at the heart of the issue. A key question is then how this challenge may be met.
Some economists argue that the nature of institutions plays a crucial role in how far such
problems are met, and section 5 examines the institutional economic literature. One type of
institutional theory, the ‘varieties of capitalism’ approach, has been used as the basis for
suggesting that some types of market economy may be able to overcome high-carbon lock-in
more easily than others.

The paper concludes by considering what the review of the economic literature implies for a
research agenda on sustainable energy transition.
2. The Governance of Innovation

2.1 The economic case for innovation policy

Much of the low-carbon innovation debate focuses quite narrowly on the innovation process and the case for low carbon technology policy in addition to carbon pricing (e.g. Jaffe et al 2005, Gross et al 2012). The economic case for such policy rests on market failures in innovation that imply that private firms will provide less innovation than is socially optimal. Arrow (1962a) provided one of the earliest arguments for market failure in invention. Invention can be seen as a risky process that may or may not result in new information (or knowledge). Because of the nature of knowledge, it is impossible for all the returns to invention to be captured by the person or company that put in the effort, even where there are mechanisms for intellectual property rights. For this reason, Arrow argues the social returns to invention are much larger than the private returns, and “we expect a free enterprise economy to underinvest in invention and research (as compared with an ideal)” (Arrow 1962a: 615). Mazzucato (2011) puts forward a stronger version of this argument. She emphasises that the returns to basic research are characterised not by risk, in the sense of that a known probability distribution applies to these returns, but by uncertainty, in the sense that there is no basis for forming a probability of certain outcomes.

A large literature on knowledge spillovers has developed, including some counter-arguments (e.g. Cohen and Levinthal 1989, Griffith et al 2004, Girma 2005, Cockburn and Henderson 1998, Loury 1979, Dasgupta and Stiglitz 1980a, 1980b). Nevertheless, much empirical evidence suggests that social rates of return to R&D are substantially higher than private rates of return (Griliches 1992, Bloom et al 2005), and many economists argue the desirability of policy.

---

3 Many economists work with a highly stylised account of the development of new technologies or practices, following Schumpeter (1942) in distinguishing invention, innovation (the first commercial introduction), and diffusion.

4 One problem is that protection against simple copying through various forms of intellectual property rights (patents, copyright etc.) can only ever be partial: “no amount of legal protection can make a thoroughly appropriable commodity of something as intangible as information. The very use of the information in any productive way is bound to reveal it, at least in part. Mobility of personnel among firms provides a way of spreading information. Legally imposed property rights can provide only a partial barrier, since there are obviously enormous difficulties in defining in any sharp way an item of information and differentiating it from other similar sounding items.” (Arrow 1962a: 615). A second problem is that knowledge from invention produces a much wider set of benefits that would not have otherwise existed, that are not based on copying the invention but on building on its existence. Thus the development of the internal combustion engine benefited from the invention of refining petrol from crude oil.

5 i.e. the economic returns to society as a whole, as opposed to the private returns to the inventor

6 This distinction was first formulated by Knight (1921)
to support innovation (e.g. Segerstrom 2000, Aghion and Howitt 1998, Geroski 1995b). This applies not only in general, but also explicitly to policy for technological change to help solve environmental problems, as a complement to classical policy to address environmental externalities, such as carbon pricing (Margolis and Kammen 1999, Foxon 2003, Jaffe et al 2005, Köhler et al 2007, Alic et al 2003, Aghion et al 2012). This view can also be found in official policy documents, especially from the business department (e.g. DTI 2006, CEMEP 2007).

The concept of knowledge externalities, or spillovers, and its potential effects on innovation is particularly relevant for the energy sector because it is predominantly an industry dominated by infrastructure and customer services, in which legal barriers to the spread of knowledge are far less prevalent than in manufacturing or in knowledge-intensive services (i.e. software). Whereas the use of patenting as an institution to partially offset the market failure is widespread in the latter sectors, it is much harder to patent or copyright service innovations such as time of use tariffs or energy service packages.

The literature on knowledge spillovers and R&D subsidies is important, but also limited, in that it focuses narrowly on the development of new technologies at the early R&D phase. A number of other market failures have been identified in later stages of the innovation process. One issue is that as inventions move from the development through the demonstration and niche markets phase, financial risk rises sharply just at the point where policy support in the form of R&D tax credits is no longer available (the so-called “valley of death”). A particular problem in crossing the valley of death is that:

“…unless an inventor is already wealthy, or firms already profitable, some innovations will fail to be provided purely because the cost of external capital is too high, even when they would pass the private-returns hurdle if funds were available at a ‘normal’ interest rate.” (Hall 2002: 36)

Certainly, the use of R&D subsidies is widespread, being one of the largest and fastest growing forms of industrial aid in developed countries. In the EU around one-third of R&D across all sectors is publicly subsidised, which are exempted from state aid rules (Eurostat 2009 cited in Zúñiga-Vicente et al 2012). However, while the case for measures such as R&D subsidies is often accepted in principle, since the 1960s a research programme has developed examining the effectiveness and costs of such policies in practice (see also below section 2.2) (David et al 2000, Zúñiga-Vicente et al 2012: 25, Lentle and Mairesse (2009). However, there is clearly evidence that public support to R&D can have a positive impact on R&D spend and innovation (Hall and Van Reenen 2000, Klette et al 2000), suggesting that different factors ranging from institutional context to the design of subsidies to the characteristics of particular sectors may have an influence on effectiveness.

As Jaffe et al (2005: 166) put it: “The realization that the process of technological change is itself characterised by market failures complicates policy analysis, and increases the likelihood that a portfolio of policies, rather than that policy directed at emissions reduction alone, will offer a more complete response to environmental problems.”
Central to this problem is the fact that asymmetries in information are inherent in financial markets (Hall 2002: 38, Jaffe et al 2005). Because banks and other investors have more difficulty in distinguishing good projects from bad ones in long-term innovative areas than in conventional short-term areas, they are more likely to charge a premium, or simply not make credit available at all, for innovation (Leland and Pyle 1977, Stiglitz and Weiss 1981). Amongst other things, this analysis implies that large firms, which have large amounts of internal finance and do not have to seek external finance, will find it easier to innovate than small firms. This argument, originally made by Schumpeter, is discussed further in section 4 below, along with implications for radical vs. incremental innovation. It also implies that, despite the existence of a venture capital sub-sector, there is a potential role for policy in providing finance, and indeed public finance for innovation is part of the policy mix in the US, UK and in Europe (Hall 2002: 46-47).

There are other potential market failures relating to knowledge spillovers in the innovation process at later stages of adoption. Arrow (1962b) argues that new information or knowledge is produced through “learning-by-doing”, i.e. from the repeated experience of applying a technology through investment or use. Learning-by-doing applies to experience gained in production. A related idea of learning-by-using applies to experience gained in the adoption and use of a new technology (Foxon 2003: 11), where again positive externalities are created “in the form of the generation of information about the existence, characteristics and success of the new technology” (Jaffe et al 2005: 167). Learning-by-using is likely to be particularly relevant for the energy industry, especially in electricity generation and networks, where the use of capital goods makes up a large part of activity.

The effects of learning-by-doing or learning-by-using are to lower the costs of a new technology over time, often represented in learning or experience curves (e.g. Junginger et al 2010), although these can also reflect simple economies of scale. These effects are sometimes characterised as “incremental” innovation in the production or use of existing technologies, as

---

9 If all actors in capital markets had equal access to full information about new technologies and new markets, then promising innovations would always be able to obtain finance at market rates. However, this is never the case, and in real economies information about costs and risks is always incomplete (e.g. Stiglitz 1993). In an economic exchange, information is also typically asymmetric: one actor in an exchange or contract almost always knows more about the costs, risks and qualities involved, and since acquiring information has costs, there are limits to how far the less knowledgeable party will go to equal up that asymmetry (Akerlof 1970). Applied to finance for bringing an innovation to market, this implies that the innovator (borrower) will know more about risks and costs than the lender.

10 As before, the company that acquires information or knowledge through learning-by-doing cannot exclude others from that knowledge for long, so they too will benefit; “the presence of learning means that an act of investment benefits future investors, but this benefit is not paid for by the market. Hence, it is to be expected that the aggregate amount of investment under the competitive model of the last section will fall short of the socially optimum level” (Arrow 1962b: 168).
opposed to the “radical” innovation associated with the invention of new technologies.\textsuperscript{11} The policy implications of learning-by-doing and learning-by-using externalities are that the private sector will underprovide both types of activity, and that there is a role, in principle, for government in incentivising both processes, for example through creating initial markets for new technologies outside commercial markets until costs are brought down to a level where they are competitive. Markets can be created either directly by government, for example through procurement, or encouraged indirectly through support mechanisms (e.g. Neuhoff and Sellers 2006).\textsuperscript{12} Mazzucato (2011) points to the key role played by the US government in the creation of markets in key areas such as health, agriculture and IT through ‘mission’ research programmes, that included not only basic research but also the development of technologies through the creation of markets.

One issue in the low-carbon technology policy debate is then the boundary between technology development and “mass deployment” of technology. Critics such as Helm (2010) and Moore (2011) appear to argue that current and planned deployment, for example of off-shore wind in the UK, will not bring costs down, implying either that many renewable technologies are already mature, or that some form of technology development other than deployment would be more effective and efficient. By contrast, the technology literature based on historical learning patterns suggests that, while the balance between R&D and learning-by-doing will vary between technologies, the importance of the latter is in general well-established (e.g. Sagar and van der Zwaan 2006). For example, Crown Estate (2012) argues that a 40% reduction in the levellised cost of electricity from offshore wind can be achieved by 2020, with learning-by-doing in supply chains and installation laying a significant role. However, critics would in turn dispute what they perceive as an optimism bias in such studies.

The discussion so far applies to the economics of innovation \textit{in general}. Additional economic arguments have been made specifically for policy support for innovation in low-carbon technologies for the purposes of mitigating climate change. Such innovation is clearly different from market-driven innovation, both in that it is aimed at producing a public good, and that it is particularly dependent on policy (Pearson and Foxon 2012). One important argument is that, because there is large uncertainty both about the speed and potential costs of climate change and the costs of mitigating climate change (i.e. the long term costs of low carbon technologies to substitute for high carbon ones), there is a (high) value in creating new mitigation technology options, in bringing them forward and in reducing the uncertainty about the performance of

\textsuperscript{11} Sometimes also known respectively as Usherian and Schumpeterian innovation after the economists A. P. Usher and Josef Schumpeter (Ruttan 1959)

\textsuperscript{12} Note that this means mechanisms above and beyond carbon pricing, which is supposed to be correcting the environmental externality rather than innovation externalities (e.g. Stern 2007: 408, Acemoglu et al 2012).
technology options through investment (Papathanasiou and Anderson 2003, Köhler et al 2007: 158), as opposed to “irresolvable ‘paper arguments’ over…merits and costs” (Papathanasiou and Anderson ibid: 80). This argument for the value of technology options under uncertainty is analogous to that for the value of investment options to companies – so-called “real options” (Dixit and Pindyck 1994) – or the hedging value of financial options. However, since the value of policy options accrues to society as a whole, the private sector will not create them on its own, and there is once more a role for policy.\footnote{An important issue here is that the value of mitigation technology options is actually of global benefit, whereas innovation policy is largely undertaken by national governments. In the absence of international collaboration of innovation policy, there is a view that governments should focus their efforts on creating options in areas where they have comparative advantages in the “supply” of technology (e.g. certain engineering skills, low labour costs) or in a particular resource (such as offshore wind or solar irradiation). However, this view can also be criticised for taking a static view of comparative advantage, whereas a more dynamic view would be that comparative advantage can be developed over time.}

A second argument, made by Kalkuhl et al (2011) and Bennear and Stavins (2007) is that if carbon pricing is initially absent or weak (as it is in almost all economies), then support to the development of low carbon technologies as a “second best” policy will be preferable to no policy at all. However, they also argue that permanent subsidies to low carbon technologies are a poor substitute for carbon pricing in the long run.\footnote{It should also be noted that policies to support the development of low carbon electricity technologies through deployment would be expected to interact with carbon pricing through cap-and-trade schemes (e.g. Sorrell and Sijm 2003, Fischer and Preonas 2010). The presence of a cap in the power sector means that developing and using low carbon generation capacity will reduce the demand for permits overall and therefore the price of permits and the equilibrium price of high carbon electricity. At the same time, increasing the share of low carbon electricity especially can reduce the profitability of high carbon power producers, reducing their output and the demand for permits. The effect of a lower carbon price is to shift generation within high carbon sources from gas to coal or oil (Böhringer and Rosendahl 2009).}

Innovation in low carbon technologies, goods and services is also different from innovation in other arenas because of its greater exposure to policy risk. With a few exceptions, low carbon technologies do not offer the prospect of a new user experience. In many cases, low carbon sources of electricity and heat offer the same energy services to end users as do high carbon sources (Fouquet 2009). To this extent, and as long as low carbon sources of energy remain higher cost than high carbon sources, low carbon innovation is more reliant on policy frameworks (including carbon pricing) than conventional innovation, where markets are created by the offer of a new service or experience. All innovators face the risk that their technology may not work, or that they may not be able to find a market, but low carbon innovators also face the risk that the policy which creates their market may be retrenched. Ultimately, this is a political risk.
Even in the area of energy saving, markets appear quite weak. As is well known, many energy savings measures have a negative cost in themselves, but are not taken because of additional transactions costs, market failures in finance (since they often involve upfront), split incentives and lack of “materiality” or salience (e.g. Sorrell 2009, Stern 2007: 427-32). Energy saving can also be a “credence good”, in the sense that consumers have far less information about the effectiveness of measures than providers do, and may be sceptical that they can in fact obtain the results that they are promised (Sorrell et al 2004). There may also be a range of other factors, increasingly explored in behavioural economics, which help explain why the demand for energy saving is weak (Pollitt and Shaorshadze 2011: 13-14). This again means that the market for potential innovation in energy services is unusually dependent on policy interventions (and therefore unusually exposed to policy and political risk) to help strengthen it.

2.2 Government failure and the innovation policy debate

The discussion in the previous section implied a role for policy at various stages of the innovation process (Martin and Scott 2000), including R&D subsidies, finance for innovation, the creation of initial deployment markets to increase learning effects and support to develop low carbon technological options. This case for low carbon technology in principle is widely accepted by policy makers. However, there is considerable debate about exactly which forms that support should take in practice.

Two schools of economic theory provide two different approaches to this issue. One, the Meadean tradition of the economics of public policy: “explores the question of the design of policy in the context of economies which are imperfect in some way – information problems, constraints on taxation, fixed prices – and where an objective can be specified in terms of a social welfare function…” (Stern 2010: 257-58). That is, an economic analysis of policy where the policy maker has the public good at the centre of their objectives. On this view, the key issue in innovation policy is the balance between the benefits and costs of interventions to support innovation, given limitations such as information asymmetries between government and firms. For example, if knowledge spillovers in a particular technology area are relatively small, and at the same time, it is difficult and costly for the government to acquire enough information to ensure that R&D subsidies, say, can be targeted to ensure that they lead to additional R&D, then it simply may not be worthwhile pursuing the policy.

15 The early literature on innovation policy made a distinction between “technology push” policies, focusing particularly on the “supply” of R&D and “demand pull” policies, which saw the strengthening of initial markets as key (for a review see Grubb). However, technology and innovation studies now tend to see both as important and policies for both as part of a wider innovation systems, often with specific national characteristics (Nelson 1993).
Full cost-benefit analysis of innovation policy is in practice difficult, because the full value of the successful development of a new technology cannot be known. In the case of policy aiming at developing technologies for mitigating climate change, the problems are compounded by several levels of uncertainty about climate change and the possibility of catastrophic outcomes (Weitzman 2009, Pindyck 2011). Given this, the cost-effectiveness of measures aimed at reaching particular outcomes (a certain level of investment, a given level of technology cost reduction) becomes the key issue. For example, the Renewables Obligation in the UK has been criticised as a policy for helping induce learning-by-doing because it is less effective in reducing risk than a feed-in tariff (e.g. Mitchell et al 2006). A detailed analysis of innovation policy design is beyond the scope of this paper, but it will again partly be an empirical question, varying from technology to technology and sector to sector, with the most cost-effective interventions differing between them (Martin and Scott 2000).

A second approach to the question of technology policy in practice comes from the Buchanan tradition of public choice, which focuses on “what happens if self-seeking individuals or coalitions try to manage or manipulate the formation of policy for their own benefit” (Stern 2010: 258). The assumption here is that the interests of such groups diverges from the wider public interest. The Buchanan tradition does not assume that government is necessarily benevolent, and indeed at the centre of its analysis is the idea that a combination of rent-seeking by companies and self-interested behaviour by politicians and bureaucrats leads to deeper kinds of government failure. While there is no universally agreed definition, a widely used one is that government failure occurs when intervention leads to an inferior outcome (in terms of efficiency or equity) than laissez faire, even if the latter itself is not socially optimal (Le Grand 1991, Krueger 1990). In other words, government failure happens when attempts to intervene make things worse. Government failure may arise from policy being distorted through the legislature (politicians seeking votes), through the administrative (bureaucrats seeking to maximise their spending) and through regulation (rent-seeking and capture). The extreme neo-liberal view is that, while there may be market failures, government failure is always greater. While the Meadean approach suggests policy reform, the neo-liberal approach suggests a drastic reduction in the size of the state and the withdrawal of policy from large areas of the economy. Applied to innovation, the neo-liberal view might accept that there are market failures in the innovation policy, but would assert that intervention in the form of innovation policy will make things worse, being both ineffective and costly.

In the context of policy for the development of low carbon technologies in the UK, Helm (2010: 185) argues that “ignoring government failure has led to high-cost, low-effect policies. Government failure, as a concept, is almost entirely absent from the mainstream climate change policy literature.” His view is that policymakers have allowed policy to be distorted by lobbying
and electoral competition for marginal voters. He argues that renewables policy in the UK (i.e. the Renewables Obligation) has taken the form it has because of lobbying by the wind industry, since wind technologies would receive a large share of the available effective subsidy (Helm 2010: 191).\(^{16}\) This is not necessarily the only political economy interpretation of the RO. Mitchell (2008) argues that policies in the electricity sector are shaped amongst other things by the lobbying of the major vertically integrated generator-supplier incumbent firms, who own a mix of thermal plant and renewables including wind. However, the overall point is that in the context the public choice view is that there is an unavoidable “technology pork barrel” (Helm 2010: 194).\(^{17}\)

At a basic level, the argument that government interventions in innovation or industrial policy will always fail is not sustainable, since there are plenty of examples in both fields where government support has led to successful development of technologies (i.e. in terms of bringing costs down) and the development of industries and even specific companies that have subsequently competed successfully without support (e.g. Chang 2009). Watson (2008: 9) provides a number of examples from the energy field, including solar PV and building energy efficiency. This does not mean, of course, that innovation support always succeeds, and again there are examples of failure in energy, such as fast-breeder reactors. At the same time, in the case of what appear to be successes, it is impossible to prove whether the development of a particular technology would have been slower and less cost-effective if the intervention had not been made, i.e. under laissez-faire. However, in many cases (for example solar PV) the counter-factual is not particularly plausible.

Common to both Meadean and public choice approaches (and amongst some policy makers) is the view that support to technological innovation should in some sense be general and “technology neutral”,\(^{18}\) and that governments should avoid trying to “pick winners” through technology-specific support (Watson 2008: 9). The IEA, for example, frequently takes this view with respect to energy technology policy (e.g. IEA 2007). Arguments about the dangers of picking winners, also appear in the debate about industrial policy (Lin and Chang 2009, Rodrik 2007: 99-152, Pack and Soggi 2006). The picking winners debate can be confusing, since it is not always made clear whether the “winners” refers to technologies (or industries) or particular companies. In the UK context, the poor performance in attempts at support to particular

---

\(^{16}\) In fact most ROCs up to around 2007 were generated by electricity from landfill gas (Woodman and Mitchell 2011: 3916), which also probably received the highest level of rent per ROC.

\(^{17}\) Such a view would give equal recognition to “pork barrels” captured by incumbent technologies and fuel lobbies.

\(^{18}\) Even if a general technology class (such as “renewable electricity”) is specified in a policy, it should still, on this view, be neutral between specific technologies within this class (wind, marine, biomass, solar PV etc.)
companies in the 1970s has helped muddle the debate. Here the focus is on whether policy should be targeted at specific technologies or technology areas.

In the public economics tradition, opposition to the idea of government targeting support to particular technologies is based mainly on the idea that there is always an asymmetry in information between governments and companies about the costs and potential of different technologies. It is assumed that civil servants know a lot less than the private sector about technologies, and so general support mechanisms that allow self-selecting take-up by private sector actors are to be preferred. On a government failure view, the greater the information asymmetry, the more likely it is that companies can take advantage of officials and extract rent, in the form of support being more than it needs to be to develop the technology. Information asymmetry is worse the more specific support is to particular technologies. Thus policies that define support in terms of outcome characteristics (i.e. low carbon) are argued to be more cost effective.

There are a number of counter-arguments to the view that picking winners should be avoided. First, whether in industrial policy (Chang 2009: 13-15) or in technology policy (Watson 2008) it is not possible in practice to be industry- or technology-neutral. Support mechanisms that are not explicitly technology-specific (such as the Renewables Obligation) in practice favour technologies that have the lowest costs at the time the mechanism is introduced, and are nearest to market (Mitchell 2008: 40-42, Watson 2008: 10, Unger and Ahlgren 2005). Even wider market-based instruments, such as tax concessions for R&D, will have to select what counts as R&D and what does not, and may have other implicit targeting aspects. It is also the case that the wider an instrument is, the more costly it is to monitor performance and assess the impact and additionality of policy (Chang 2010). This is important because it is more difficult to withdraw support if success or failure if not clearly defined (Rodrik 2007: 114-15).

Second, the appropriateness of targeting depends on the objective of policy. If the objective is simply to reduce the costs of the nearest-to-market technology, this may be appropriate. However, if the objective of policy is to develop as broad a range of technology options as possible, then more technology-specific policies will be preferable.

---

19 The understanding of what a "market-based" support mechanism sometimes appears to vary in the policy debate. For Helm (2010), the RO is not a market based instrument because it requires a target or quota, and an example of a market based mechanisms would be tax or subsidy. For Mitchell (2008) the RO is a market based instrument because it creates a market.
Third, there is evidence that technology-specific policy support is sometimes more cost-effective and less open to capture than generalised support. Some of the cases of success in supporting technology development mentioned above also involve technology-specific programmes, and so also demonstrate that it is possible for the state to pick winners (in the sense of technologies whose costs come down significantly in a reasonably short period of time, and which may be competitive – for example on-shore wind). Moreover, comparative evidence from Europe on non-technology specific quota mechanisms and technology specific feed-in tariffs for wind energy suggests that the former have higher costs, are less effective and produce high levels of windfall profits (e.g. Ragwitz et al 2007, Lipp 2007, IEA 2008: 105, Mulder 2008, Soderholm and Klassen 2007, Butler and Neuhoff 2008). Pollitt (2010) argues that the poor performance of the UK version of the quota mechanism (i.e. the Renewables Obligation) is due principally to space and planning constraints rather than to the design of the mechanism itself. Nevertheless, the evidence does provide considerable support for the view that technology-specific feed-in tariffs have been successful. Interestingly, this suggests that in countries with feed-in tariffs, governments have been able to overcome information asymmetries, and find ways to avoid capture (for example by digression in tariffs to match costs over time). It also suggests that while the design of particular innovation policies matters, since the design of policy can also be distorted by government failures, the process by which innovation policy is designed and implemented also matters.

2.3 Innovation policy as a joint discovery process

The theme that the process by which governments intervene in economic activities characterised by market failures is central to the success or failure of that intervention lies at the heart of a different approach to industrial policy recently developed by Rodrik (2005, 2007: 99-119), based in part on considerable empirical evidence on where that policy has been successful and unsuccessful. While there are important differences between industrial policy and innovation policy, they share the key characteristic that both are aimed at addressing knowledge externalities and coordination failures.

The standard approach in economics, as described in sections 1.1 and 1.2 above, is to enumerate various market failures, develop policy interventions to address them, and then consider the desirability of different policy options from the point of view of informational requirements or political economy considerations. Rodrik, however, starts from a view that the “location and magnitude of [the] market failures is highly uncertain” (Rodrik 2007: 100). It is worth quoting at length from Rodrik about what follows from this approach:

“A central argument…is that the task of industrial policy is as much about eliciting information from the private sector on significant externalities and their remedies as it is...
about implementing appropriate policies. The right model for industrial policy is not that of an autonomous government applying Pigovian taxes and subsidies, but of strategic collaboration between the private sector and the government with the aim of uncovering the most significant obstacles...and what interventions are most likely to remove them. Correspondingly, the analysis of industrial policy needs to focus not on the policy outcomes...but on getting the policy process right. We need to worry about how we design a setting in which private and public actors come together to solve problems in the productive sphere, each side learning about the opportunities and constraints faced by the other...Hence the right way of thinking about industrial policy is as a discovery process – one where firms and the government learn about underlying costs and opportunities and engage in strategic coordination. (Rodrik 2007: 100-101).

Rodrik goes on to argue that such a perspective makes the problem of picking winners irrelevant:

“Yes, the government has imperfect information, but...so does the private sector. It is the information externalities generated by ignorance in the private sector that creates a useful public role – even when the public sector has worse information that the private sector. Similarly, the idea that government needs to keep private firms at arm’s length to minimize corruption and rent-seeking gets turned on its head. Yes the government needs to maintain its autonomy from private interests. But it can elicit useful information from the private sector only when it is engaged in an on-going relationship with it...” (Rodrik 2007: 101).

Rodrik’s approach therefore suggests that governments need to be close enough to the private sector to be effective in coordinating activity to discover costs and overcome barriers, but not so close that they are captured. The key to policy is then in getting institutions for policy making that get this balance right, rather than focusing on specific details (Rodrik 2007 p 111).20

Rodrik goes on to propose three critical elements for such institutions. The first is political leadership at the top, to increase the prioritisation and visibility of policy, to provide coordination and it sets up a clearly identifiable high level politician as accountable for the success of policy. The second element is some form of public-private deliberation and coordination bodies within which information exchange and learning can take place. Rodrik specifically argues that “to avoid the biases of incumbents and insiders, these should go beyond the typical ‘peak’

20 This approach shares some similarities with the idea of adaptive (as opposed to optimising) policy deriving from evolutionary theories of innovation (e.g. Metcalf 1995), and with the “transitions management” literature (e.g. Voß et al 2009).
organizations that include only well-organized groups and business organizations.” Finally, a successful institutional system includes mechanisms of transparency and accountability, to maintain public support of the process.

This emphasis on getting institutions right for the innovation policy process raises the question of how feasible it is to change institutions within any particular country. Rodrik himself argues that policy does not “travel well” between different countries because of differences in wider institutional settings (Rodrik 2008), and Pollitt (2010: 39) uses this point to argue that “general” market-based support mechanisms remain better suited to the UK, as it is more committed to liberalised energy markets than Germany and other continental European countries, in which deliberation plays a greater role. This view that particular policy approaches cannot be easily transplanted between different institutional systems is also found in the “varieties of capitalism” (VoC) literature, discussed in section 4 below (see e.g. Soskice 1997).

However, this analysis is also disputed, even from within the VoC literature. For example, Culpepper (2001: 275) argues that “political initiatives that aspire to create coordination in policy areas where it has previously not existed...can succeed.” Despite the dominance of arms-length market-based economic relationships between amongst companies and between government and companies in the UK in theory, there are examples which come closer to the deliberation model such as the Carbon Trust “Accelerator” programmes in which the Trust, private sector firms and universities worked closely on specific technologies such as marine renewables, with considerable success. At the same time, institutional systems at the higher level are not immutable, especially in periods of crisis. Change in such systems is complex and bound up with shifts in fundamental ideas, or policy paradigms (for a discussion see Kuzemko 2013). Shifts in the deep-seated ideological and institutional foundations of policy go beyond the scope of this paper. However, it is worth noting that it is increasingly recognised that the governance of energy in the UK is already shifting away from a simple liberalised market framework, and that there is an active debate about whether there should be further shifts (e.g. Keay et al 2012).

2.4 Conclusions

Economic theory provides several reasons for policy to support the innovation of new technologies, services and business models, based on a number of well-established market failure arguments. There are additional specific rationales for support to the development of low-carbon technologies given the urgency of climate change and uncertainty about technology costs. In the absence of carbon pricing these arguments are even stronger. Economic arguments do not, of course, provide the only rationale for low carbon technology support; there are also, for example, political arguments (Lauber and Jacobsson 2006).
Concerns about government failure, centring on asymmetric information and the risk of capture, have tended to lead economists to argue for “technology neutral” market-based support mechanisms. However, for both theoretical and empirical reasons, the case against technology specific support policies is not as strong as it appears, and the more important underlying policy question is how governments can overcome asymmetric information problems and the risk of capture to provide effective support without excessive rent. The answer to this question may vary with wider institutional settings, but it lies at the heart of understanding both differences in comparative experience and in how policy in the UK might be made more effective.

3. The Governance of Markets

The discussion in section 2 above focused on what economic theory says about the direct governance of innovation processes through policies aimed at tackling market failures. However, innovation as a process does not happen in an institutional vacuum, and economic theory also posits that institutional context has an effect on the ability and incentive for actors (essentially firms) to undertake innovation. The most immediate institutional context shaping innovation is market structure. In this section I examine debates on the influence of market structure on innovation and explore their implications for the UK retail energy market.

3.1 Two views on innovation and market structure

Historically there have been two contrasting views within economics on the relationship between market structure and innovation. One is that competitive markets are needed to deliver innovation. This view may be seen as related to Hayek’s characterisation of competition as a discovery mechanism (Hayek 2002). Competition leads not only to the discovery of the least cost way of producing goods and services in an economy, but also to the discovery of better goods and services. These both involve innovation – doing things in a new way (i.e. process innovation) or producing new things. The view that competitive markets would deliver innovation was certainly embedded in the project of privatising and liberalising the retail energy market in the UK. Some of the main architects of this projects, including Stephen Littlechild and Michael Beesley, were heavily influenced by Hayek (Helm 2004: 57-60, Littlechild 1981).

An alternative view was proposed by Hayek’s contemporary Joseph Schumpeter (1934, 1942). Schumpeter’s argument had two parts. The first is that large firms may be more able to innovate, because they are more likely to have the resources and access to finance internally to

21 As Vives (2008: 420) notes: “Schumpeter himself oscillated between thinking that monopoly rents or competitive pressure (in particular the entry threat of rival innovators) were the drivers of innovation although usually only the monopoly driver is emphasized in the interpretation of his work.”
generate innovation than small firms (see also above Section 1.1).\textsuperscript{22} The second is that highly competitive markets could work against innovation, by reducing the returns to innovative activity. Innovation can produce rewards for companies, but it is also costly and risky, involving R&D, demonstration and pre-commercial deployment costs before commercial returns can be secured. The balance of risk and reward determines how much innovation companies will undertake. If a firm develops a new product, process or service that no competitor offers, then it is in a monopoly position and can earn monopoly rents, charging a price above the marginal cost of production. This is how the innovative firm earns a return on its invention. How long such a monopoly rent remains in place depends on the appropriability of the innovation (see above section 1.1). Intellectual property rights (IPR) mechanisms such as patents are intended to increase appropriability and raise the revenue a company gets from marketing inventions, because it allows, in theory, companies to prevent copying or to raises revenue through licencing. In practice, IPR is never perfect (Gilbert 2006: 163), innovations can be adapted, and companies that undertook the innovation first incur far higher costs than those adapting it. Competitors can then offer their product at a lower price, and take a part of the market. The original firm may be able to recapture that market share, but only by reducing its price. Either way, its revenue falls relative to what it could have earned if it could perfectly protect its invention. Thus very high levels of competition could be expected to discourage innovation because monopoly rents are competed away very quickly.

Arrow (1962) argued, in a Hayekian vein, that an incumbent monopoly firm in a market has less incentive to innovate than a new firm entering that market. If a company with a cost-reducing innovation entered a competitive market it would win market share from other firms. By contrast, if a monopoly company came up with a cost-reducing innovation, it would effectively be taking away market share from itself. Its incentive to innovate is reduced by the amount of profit that it currently earns and would lose by innovating (the “replacement effect”).\textsuperscript{23} This is especially relevant for radical innovations that would completely destroy the value of sunk costs and existing profits. New entrants play an important role in innovation precisely because they do not have vested interests in the pre-innovation way of producing.

\textsuperscript{22} In a recent review Jamasb and Pollitt (2008: 999) note evidence for a positive relationship between firm size and R&D in the electricity industry.

\textsuperscript{23} See also Reinganum 1983. In the context of the UK energy retail market, this point can be seen in Littlechild’s (2012) argument that restricting companies to four tariffs will disincentivise innovation in new products, because companies will not be able to offer a new tariff without withdrawing an existing one. Without such restrictions it is not clear that the Big Six do actually have sufficient product differentiation to be considered as monopolistic competitors, but it may apply to some of the niche “green” suppliers such as Ecotricity and Good Energy. Their incentive to develop new products then depends less on concerns that other firms (or new entrants) will also offer them, and more on whether the new products make their existing products obsolete.
Following Schumpeter and Arrow’s contrasting hypotheses, a large literature has developed on the relationship between market structure and innovation (see Cohen and Levin 1989, Gilbert 2006 for reviews). However, empirical studies have not shown that one of these hypotheses can be definitively regarded as correct (Scherer 1967, Cohen et al 1987, Cohen and Levin 1989, Nickell 1996, Blundell et al 1999, Gayle 2004).

Some studies show an inverted-U shaped relationship, where innovation at first increases with the concentration of markets and then decreases, so that innovation is low in very competitive and in monopolistic markets but higher in between (Levin et al 1985, Aghion et al 2005, Askenazy et al 2007). Aghion et al (2005) develop a model to explain this pattern, which draws on both Schumpeterian and Arrovian arguments. In concentrated markets with a few players, companies tend to be neck-and-neck, and there is an incentive for leaders to innovate to get ahead of rivals (an “escape competition” effect). At this end of the scale, the more competitive the market the more innovation there is. In markets with many firms, there will be more technological laggards, who have an incentive to innovate to catch up. However, when markets are extremely competitive (i.e. have many firms), gains are quickly eroded, so at this end the less competition there is the more innovation there is. Overall, this provides an explanation for the inverted-U shape.

3.2 Innovation and factors of competition

As the debate on market structure and innovation has proven inconclusive at a general level, research has increasingly focused on the need to disaggregate the relationship between competition and innovation by specifying them more carefully:

“As a general statement, the incentive to innovate is the difference in profit that a firm can earn if it invests in R&D compared to what it would earn if it did not invest. These incentives depend upon many factors including: the characteristics of the invention, the strength of intellectual property protection, the extent of competition before and after innovation, barriers to entry in production and R&D, and the dynamics of R&D. Economic theory does not offer a prediction about the effects of competition on innovation that is robust to all of these different market and technological conditions. Instead, there are many predictions and one reason why empirical studies have not generated clear conclusions about the relationship between competition and innovation is a failure of many of these studies to account for different market and technological conditions.” Gilbert (2006: 162)

In other words, there can be many models of market structure and innovation, and it is important to know which one is appropriate for each context (Gilbert 2006: 165). This point is particularly
relevant for thinking about how innovative the energy markets are likely to be, because as an
industry energy retailing has different characteristics from those often included in the general
empirical studies of market structure and innovation, which tend to be based on manufacturing
industries.\textsuperscript{24}

One problem has been how to define and measure ‘competition’ (Boone 2008). As Dasgupta
and Stiglitz (1980) point out, defining competition with reference to market concentration is not
useful since a concentrated market (i.e. a market with a few large firms) is not necessarily an
uncompetitive market. Other frequently used measures of competition, such as price-cost
margins, are also not necessarily a good guide (Boone 2008).\textsuperscript{25} Instead, more fundamental
“factors of competition” need to be considered, especially: the degree to which products in a
market can be substituted for one another \textit{(product substitutability)}; the costs of entry into the
market; and the \textit{degree of intellectual property protection} available to innovating firms (Boone
product substitutability is high, where entry costs are low, and where intellectual property
protection is poor. Another issue is the nature of the innovation in question, and in particular
whether it involves a new \textit{process} that has the effect of reducing the costs of producing an
existing good or service, or an entirely new \textit{product}, which creates a new market (and may
destroy existing ones).

Gilbert (2006: 204) argues that the underlying principles in the relationship between the factors
and nature of competition on the one hand, and innovation on the other, are fairly
straightforward: “The incentives to invest in R&D increase with the profits that a firm can earn or
protect by innovating and decrease with the profits that a firm can earn if it does not innovate.”
The question is what theory says about how these principles operate in relation to different
factors of competition and types of innovation, and what the overall effects might be. This
question is complicated by the fact that there are multiple theoretical interactions between the
factors, and that effects are also somewhat different for product vs. process innovation.

For \textit{product} innovation, low \textit{product substitutability} allows firms to gain a degree of monopoly
rent, and the incentive for innovation is driven by expected profits \textit{net} of any replacement effect
(see above section 3.1) (Gilbert 2006: 204) and subject to a degree of \textit{intellectual property
protection} (Vives 2008, Gilbert 2006: 167). By contrast, high product substitutability prevents
product differentiation and so implies a low level of product innovation.

\textsuperscript{24} Although Tingvall and Karpaty (2011) find a similar inverted U shape relationship for service industries
\textsuperscript{25} Price-cost margins (PCM) are used by Ofgem in its monitoring of energy retail markets. A high cost PCM is seen
as a sign of insufficient competition. However, Vives (2008) argues that high PCMs may also result from a selection
effect, where low efficiency firms (who have low PCMs) exit from a market because they cannot compete.
The effects of market entry costs on product innovation depend on the degree of product substitutability and on the strength of the replacement effect. Arrow's analysis implies that monopolistic markets are harmful for innovation where there is no threat of entry. However, where an incumbent monopoly firm faces the threat of entry by other firms, the situation may be different. Innovation can then become a defensive strategy (along with other strategies such as limit pricing and buying out potential competitors). Various neo-Schumpeterian hypotheses have been put forward linking a degree of monopoly power to product innovation where entry is possible. Etro (2004) argues that if market entry is free, Arrow's replacement effect disappears and monopolists have a greater incentive to innovate than new entrants. Gilbert and Newbery (1982) similarly propose that monopolists have an incentive to develop new products if by so doing they can prevent entry by potential competitors which would lower their profits by more than the costs of innovation. However, this effect depends on the ability to protect intellectual property (i.e. pre-emptive patenting). Vives (2008) also argues that increasing the number of firms in a market by removing restrictions or lowering entry costs will lead to greater product innovation, but that the greater competition will also reduce the incentive for cost-reducing process innovation, especially where intellectual property protection is weak (Dasgupta and Stiglitz 1980).

For process innovations that reduce the costs of production, the returns to innovation are higher the larger is the total market, but also the fewer are the number of firms in the market, and also the larger the price elasticity of demand for the variety. Situations where there are no entry restrictions and costs are low will militate against process innovation, especially where there are low levels of intellectual property protection:

“In the case of nonexclusive intellectual property rights the presence of rival firms that can independently invent and adopt the new process technology reduces the value of discovery to each potential inventor. As the number of firms that compete in research and development increases it is likely, though not necessary: that each firm’s share of the total output using the new technology would fall and so would its corresponding benefit from invention. In this case, competition in R&D and in the market for the new technology reduces the value of innovation. With nonexclusive rights to a process technology and profit-maximizing inventors, competition can be bad for R&D.” Gilbert (2006: 164)

---

26 There can also be degrees of market entry. For example, the literature on competition with switching costs implies that entering a market and attracting a small number of customers may be easy, but subsequently expanding that base quickly becomes harder because as competition shifts to the “stickier” part of the market (Farrell and Klemperer 2007). This would appear to apply to the UK retail energy market, where there are new entrants but they have tended to remain relatively small.
This view contrasts with Arrow’s (1962) view (see section 3.1 above), but his analysis assumes that intellectual property protection is strong.

Overall then, low levels of IP protection are hypothesised to have a corrosive effect on all innovation, especially when combined with low market entry costs. Low market entry costs with higher levels of IP protection can stimulate incumbents to innovate as a defensive strategy. High levels of product substitutability work against product innovation, but where there is good IP protection, can stimulate process innovation. The combined effects of these factors of competition can thus be expected to vary according to particular conditions in different markets. The implications for a transition to sustainable energy systems therefore depend on the characteristics of particular energy markets. The next section considers the example of the UK retail household energy market.

3.3 Innovation in retail energy markets and the debate over competition

There are different views on how competitive the UK retail energy market actually is. Some express concern about cartel-like behaviour, low levels of switching by consumers and market segmentation (Waddams-Price 2005, NAO 2008, Defeuilley (2009: 380-81), while others take the view that markets have actually been reasonably competitive (Littlechild 2012). Ofgem seems to have evolved from being sanguine to being concerned about declining levels of switching and increasing complexity of tariffs (Ofgem 2007, 2011).

However, the debate on competition has largely been about pricing and costs to consumers. The issue of the impact on innovation is less discussed. There has been innovation in the retail market, but this has focused almost entirely on tariff packages, and it does not involve “a broad redefinition of retail market attributes nor challenge incumbents’ business models by disqualifying their offers both technically and commercially.” (Defeuilley 2009: 382). The early years following the liberalisation of retail energy markets saw a proliferation in the number of different types of contract that did not exist at liberalisation, in terms of duration, bundling of gas and electricity, payment type, origin of the electricity including clean energy products, pricing options, as well as non-price deals involving loyalty schemes etc.  

27 The combined sale of electricity and gas (dual fuel deals) has been the primary innovation, with about a third of domestic customers on such a contract by 2007 (Ofgem, 2007: 4). Fixed price guarantees have also been relatively popular. The first of these was introduced in 2003, and by 2007 all suppliers offered one, covering 6 million product accounts, or 13% of the market (Ofgem 2007: 13). By contrast, green tariffs for electricity have been much less popular (around 300,000 accounts by 2007, or a little over 1% of that market). There were also new channels of retail supply (sales over the Internet) and joint offers (sales of energy associated with telephony or Internet access and more recently other services such as plumbing and electrical maintenance). First introduced in 2005, there were 2.5 million online accounts with the Big Six by 2007, representing approximately 5% of the market (Ofgem 2007: 14-15).
Suppliers have also developed offers of products and services in the area of energy efficiency, such as free home energy surveys, discounted loft and cavity wall insulation, energy efficient boilers and energy saving appliances, along with information on energy efficiency in various forms, and some incentive schemes for energy saving (Ofgem 2007: 16). However, these developments have nothing to do with market structure and incentives to innovate, being driven instead by the increasing obligations on suppliers under different generations of energy efficiency policy. Suppliers have had no incentive to promote these services in the wider market, once they have reached their obligation. Equally, the development of some new value-added services such as demand monitoring and new technologies, such as web-based auditing and energy management software (OFGEM, 2007, NAO, 2008), is not a core innovative activity driven by competition.  

From the perspective of the factors of competition outlined in section 3.2 above, the UK retail energy market is characterised by countervailing tendencies. On the one hand, there is a very high degree of substitutability in the underlying products – i.e. gas and electricity. However, as noted above since the late 1990s energy suppliers have tried to “innovate” by offering new pricing deals and packages, and indeed the plethora of tariffs on the market by 2012 might be seen as an attempt by energy suppliers to create “product” differentiation and extract a degree of monopoly power. But each supplier is at the same time limited in the extent to which they can extract monopoly rents by the low levels of intellectual property protection in the market. This is because the energy market is basically a service market; virtually nothing that energy companies do can be patented, and the innovations of rivals are easily monitored, not least through the switching websites that have grown up around the industry.  

On the other hand, energy retailing in the UK appears to have some barriers to entry. According to Littlechild (2005), these include fixed costs and economies of scale in areas such as billing; regulatory costs, such as licensing, poor data and limited metering services. But of special interest is the possibility that the vertical integration of suppliers upstream into electricity generation appears to have led to market foreclosure. Although public energy companies were unbundled at or after privatisation, there was a wave of vertical integration in the UK market in  

28 British Gas did anticipate the official smart meter programme, starting its own roll-out from 2010, but it is questionable whether this would have happened in the absence of the official programme, which was in the advanced planning stages at the time British Gas initiated its programme.  
29 As Defeuilley (2009: 382) puts it: “These new products, tools and contracts, are easily reproducible from a supplier to another one and may be quickly disseminated among all market participants. They seem unable to give a clear-cut and long-lasting competitive advantage to an innovative new entrant in the retail market. To date, this new entrant is not in position to create, what J. Schumpeter coined a temporary monopoly position, from which he will gain overprofit and exploit his competitive advantage at the expense of the incumbent companies.”  
30 Vertical integration makes sense where the risk on investments by one firm can be hedged by securing a market through taking over another (e.g. Grossman and Hart 1986, Holmstrom and Roberts 1998), and this will apply especially for electricity where technological intensity is high (so investment costs are large) (Acemoglu et al 2010). This view is supported by empirical studies of the electricity industry (Jamasb and Pollitt 2008: 999-1000).
the late 1990s after liberalisation of retail markets (Rocques et al 2005, Helm 2004, Mitchell 2007). While some take the view that vertical integration will not affect competition (and may increase innovation because positive externalities from such innovation can be captured within the firm and because of the firm size effect), others argue that it can reduce competition through market foreclosure, where potential rivals, and especially new entrants in retail may be foreclosed from obtaining supply from wholesale markets (Salinger 1988, Ordover et al 1990). Such markets are then effectively “uncontestable” for new entrants (Baumol et al 1982).

Overall, then, the UK’s liberalised retail energy market tends to have a low degree of product differentiation and low levels of intellectual property rights protection, but quite high costs of entry, partly relating to foreclosure through vertical integration. The theory discussed above in section 3.2 suggests that in such markets there will be a tension between the search for temporary monopoly profits through developing new products on the one hand, and a combination of the replacement effect and the lack of IP protection on the other. The latter in particular means that it is hard to escape competition through innovation, in the sense proposed by Aghion et al (2005) in a market with a relatively small number of head-to-head firms. This may be a reason why energy companies may fluctuate between seeking to get ahead of rivals through innovation (e.g. on pricing packages) and seeking to coordinate to reduce competitive pressure on each other.

At the same time, high costs of entry and certain restrictions on entry are bad for product innovation, and lowering costs and removing restrictions would encourage more new entrants and more products. This is particularly relevant for thinking about inducing innovation for sustainability through policy. The significance of entry conditions lies in the fact that it is an area that governments and regulators can take action on, in contrast to the degree of substitutability and the degree of intellectual property protection, which are more inherently determined by the nature of the product and technology. This is the case not only for such “products” as tariff packages, but also for products aimed more at sustainability such as demand response deals and even home automation systems and services. There may be considerable IP protection through patents for the hardware in such systems, but not for the service that suppliers can offer in installing, commissioning or helping to run them.

The original intention behind liberalisation was to make energy markets contestable (Beesley and Littlechild 1989: 464-468), and this reasoning was behind pressure from the regulator that led to the break-up of British Gas Trading in 1997 (Helm 2004: 246-47). However, when it came to vertical integration the government and regulator took a more relaxed view (Helm 2004: 234-38). Littlechild (2005: 47-48) finds some evidence to support the suspicion that there is foreclosure, with very low liquidity in wholesale electricity markets. Mitchell (2008: 28-29, 50) discusses the unintended consequences of the liberalisation paradigm for regulation and corporate power.

Insofar as there is competition between incumbents, because the potential gains from cost-reducing process innovations would be diluted as the market would become shared amongst many more actors, improving entry might have a negative impact on affordability over time.
The evidence from empirical studies appears to be mixed. In a literature review of the impacts of liberalisation in the electricity industry on R&D, Jamasb and Pollitt (2008) report an overall finding of a drop in R&D spending, although it is important to distinguish the effects of uncertainty created by the liberalisation process itself and the effects of competition once introduced. They also cite the finding of Sanyal and Cohen (2004) that some factors of competition (e.g. the degree to which customers can switch supplier) have a positive effect on R&D spending in the US electricity industry. Makard et al (2004) note that liberalisation and unbundling has led to a shift in the focus of innovation in electricity utilities from technology to customer products (as discussed above) and firm organisation.

Finally, Defeuilley (2009: 384) argues that competition is not the only driver of innovation, since different sectors have different characteristics related to the underlying nature of the technologies found in them. Drawing on Breschi et al 2000, Defeuilley suggests that the nature of technological learning in the electricity sector – which is largely cumulative and driven by equipment suppliers – places a fundamental limit on how innovative the retail stage of the value chain can be. This may indeed be the case as long as there is no innovation in the business model, but if a switch to an energy services model occurred, then the location of technological learning shifts (for example to automated home control systems) and the space for more radical innovation will arguably be larger. However, as noted in section 2.1 above, a key barrier to innovation in areas such as energy services is a lack of demand partly due to hidden costs and behavioural factors. Increasing market access without policies to strengthen demand may well not be particularly effective in increasing innovation in such areas.

3.4 Conclusions

One of the clearest messages from the theoretical literature on market structure and innovation is that removing barriers to market entry and reducing the costs of market entry would be expected to increase the amount of product innovation in retail energy markets. Since the energy industry does involve quite large sunk costs, active policy is needed to maintain ease of entry. However, political economy approaches in economics suggest that such attempts to reduce barriers to entry will be resisted by incumbents, since they tend to prefer to lobby for regulation that creates such barriers, and that restricts the number of firms in an industry, in preference to subsidy, which may be diluted if entry costs are low (Stigler 1971, Dal Bo 2006). This may be seen in the tolerance of vertical integration and foreclosure in retail markets discussed above. It may also be seen in UK wholesale electricity markets, where regulation has historically reinforced costs of entry for small generators and new technologies such as renewables, licensing, technical codes and the design of network user fees and the high cost of participation in the balancing and settlement mechanism required of generators of even quite
small size (Mitchell 2008: 139-50). Other countries, including Germany and Denmark, appear to have regulatory arrangements which allow a much greater range of actors in both the generation and supply of energy. Thus an important research question is how different institutional arrangements (and political forces) in different countries minimise the risk of regulatory capture, and indeed how institutional arrangements in the UK may be changed to reverse capture or minimise the risks.

Finally, it should be noted that the economic literature on competition and lowering entry costs frame the issue in terms of a market in which the actors are all firms, albeit varying in size. However, the German Energiewende experience suggests that opening markets up to different types of actor, including private individuals, community groups and local authorities can be equally, if not more effective in stimulating innovation. Mitchell (2008: 53) argues that the potential of such actors has been underestimated in the UK. There is an additional point here that encouraging a wide range of actors to invest in low carbon energy production creates a new set of vested interests that can underpin the political sustainability of policies to encourage innovation in sustainable energy (Lockwood 2013).

4. The Governance of Systems

The institutional context for sustainable innovation extends well beyond the structure of markets, and is affected by a wide range of factors and interactions (e.g. Smith et al 2010: 437-38) In this section I examine the implications for innovation, particularly in sustainable low-carbon energy technologies of debates on two issues: increasing returns and lock-in, and institutional systems

4.1 Increasing returns, path dependence and lock-in

The concept that competitive markets will lead an economy to a unique, stable equilibrium that optimises welfare is dependent on a number of simplifying assumptions, one of which is the idea of convexity, or decreasing returns. However, in real economies, there are many circumstances in which returns do not decrease, but rather increase with scale. This is especially the case for technologies, products and services that are relatively new. There are a number of reasons why increasing returns may exist.

---

33 On the supply side, decreasing returns mean that as more of a good is produced, costs per unit of production rise. This implies that firms will not want to increase production without limit, which in turn implies that markets will tend to have many firms competing with one another rather than a single monopoly. On the demand side, decreasing returns imply that consuming a bit more of a good when we already have a lot of it is valued less than consuming a bit more when we have none, or very little of it; in other words, that consumers reach a point of satiation for any particular good. Decreasing returns produce limits to what economic agents are willing to supply or want to demand, and so mean that unique prices can be found to balance the two.
One is fixed costs of production, which means that average costs decrease with the level of production, so that there are increasing returns to scale. Thus, for example, printing a newspaper involves heavy investment in a press, but then printing a single newspaper is very cheap. This means that the cost of printing the first newspaper is extremely high, but the average cost of printing the hundred thousandth is very low.

Another reason why costs can be expected to come down as more of a good or service is produced is learning-by-doing and learning-by-using (see above section 1). Such effects mean that producers and users of capital goods learn how to be more efficient in their activities, producing increasing returns (i.e. for a given amount of labour, capital and other inputs at higher levels of production, they can produce more output than at lower levels of production).

Particularly important for the analysis of technology development are network externalities. These effects arise from situations where the benefit from using a particular good, service or technology depends on how many other users there are of compatible items (Katz and Shapiro 1986; see Farrell and Klemperer 2007: 2007-09 for a wider review). For a new good or technology, positive network externalities can set up a feedback effect; the more that people adopt it, the more valuable it becomes to others, so the faster it spreads. Expectations about what people later in the process will do can also play a role: “if players expect others to adopt, they too will adopt” (Farrell and Klemperer 2007: 2025). A similar effect can arise through adaptive expectations. Especially with new goods or services when there is more than one option, potential buyers may seek information on the quality, performance and permanence of different options from those who have already made purchases. Preferences are not given, as in conventional economics, but are discovered through adaptive interaction with others. Under some circumstances this can create feedback effects, so that the more a particular number of people have bought a particular option, the more potential purchasers are drawn to that option (Arthur 1991).

Analysing markets with increasing returns has been difficult for neo-classical economics. Unlike decreasing returns, which produce stable equilibrium prices and competitive markets, increasing returns or non-convexity produce the possibility that there is no price that clears that market, and that where a market equilibrium does exist, it might not maximise welfare (e.g. Mas-Collel

---

34 Many examples involve the provision of a durable good and a complementary good or service – a frequently used example is that of software and computers, where the value of a particular type of software depends on how many other people have computers that are compatible with that software. Other examples include physical networks such as telephones, roads and railways. Network externalities also tend to apply to any technology requiring specific training, because the training is more valuable if the associated technology is more widely adopted.

35 De Vany and Walls (1996) analyse the example of new movies, where word-of-mouth feedback effects create hits and flops from amongst competing new releases.
1987). As Krugman (1997: 59) notes, there are still no general models of economies characterised by increasing returns in neo-classical economics.

However, there have been a number of different attempts to incorporate increasing returns into specific economic issues. One distinctive approach comes from evolutionary economics. Unlike neo-classical economics (where economic actors react only to price signals) or game theory (where actors respond to each other strategically), evolutionary economics allows actors to be influenced by the decisions of other actors in a non-strategic way, and market outcomes emerge from those interactions (e.g. Anderson et al 1988). This approach fits well with the analysis of increasing returns arising from network externalities and adaptive expectations, as illustrated in Arthur (1989)'s analysis of markets for new, competing technologies. He identifies four ways in which outcomes with increasing returns differ from those with decreasing returns:

- **Unpredictability** – with decreasing returns and given preferences for supply and demand, a market or economy will converge to a single, unique equilibrium, and it is possible to predict what the prices and market share of new technologies will be. With increasing returns, when new goods or technologies compete for market share, one will eventually come to dominate the market, but *a priori* it is not possible to predict which one.

- **Inflexibility or lock-in** – increasing returns such as network externalities or information contagion are self-reinforcing; they strengthen the position of the technology in the lead and weaken that of others. Beyond a certain point it is impossible for the laggards to catch up, and no intervention, such as a subsidy or tax, can rescue them (Arthur 1989: 118-19), unlike cases with decreasing returns.

- **Path dependence** – small events early on in the history of the evolution of competing technologies (or cities in economic geography) are not “averaged away” and forgotten, but rather can place a market on a path in which the future is determined by that event.

---

36 These include explaining patterns of growth and development (Murphy et al 1989), trade (Helpman and Krugman 1987) and economic geography (Fukita et al 2001). These examples all come from within conventional neo-classical economics, where increasing returns are mainly understood as arising from fixed costs or possibly learning-by-doing, and are handled by modelling markets as having monopoly-like characteristics.

37 The existence of several possible equilibria is also found in the new economic geography, and in Murphy et al’s (1989) account of economic development (see also Krugman 1991: 651-52).

38 Network externalities are not the only way of modelling lock-in. Kalkuhl et al (2012) offer a model where lock-in is produced by a low-carbon technology having learning effects, meaning that it is initially high cost and cannot realise its potential cost reductions because it is squeezed out early on by a high carbon mid-cost rival. One example might be electricity systems, which as a physical network are particularly prone to lock-in to a single system technology. The late 19th century saw a struggle between alternating current (AC) and direct current (DC) systems, in which, once established as the leader, AC was bound to win completely (Hughes 1983).

39 As the former dynamic is known as an ergodic process, path dependent processes are sometimes referred to as displaying non-ergodicity.
A frequently cited example is that of the QWERTY keyboard, one of several designs competing for the emerging typewriter market at the turn of the last century. According to David (1985), particular events, such as the winning of a speed typing competition by a contestant using the QWERTY design helped propel its popularity early on. Because each competing keyboard design required specific training, network externalities amplified such small, almost random effects, and QWERTY became established as the market leader with almost complete domination.

- **Possibly sub-optimal final outcomes** – dynamics created by increasing returns imply that it is possible for a technology to establish a lead early on and become locked-in to market dominance, even though it might not be the technology that a majority of people would have chosen had they not been constrained by those dynamics.\(^{40}\)

The above account of the implications of increasing returns is not without its critics. In a series of papers, Liebowitz and Margolis have argued that lock-in, path-dependence and sub-optimal outcomes are less useful concepts than they might first appear (e.g. Liebowitz and Margolis 1990, 1994, 1995). Much of their critique is focused on the last characteristic identified above, sometimes called ‘path inefficiency’. They argue that most path-dependent processes are not ‘remediable’, in the sense that given the information people had at the time, mistakes made were unavoidable, and it does not make sense to call the outcome inefficient or suboptimal. Moreover, if nothing could have been done at the time of the process, or by the time lock-in has occurred, then the process in question is of no interest from a policy point of view.

At the same time, they argue that in practice, in most cases, processes such as technology choice are efficient.\(^{41}\) This is partly because early consumers can see which technology is superior, and as they expect that later consumers will also see that superiority and buy in to that technology, they will also. It is also because that even if some technologies have difficulty in establishing a lead early on, if they are demonstrably superior, they will be able to attract finance backed by patents, and will in the long term be able to win market share that reflects their true value. Thus, the ability to capture the returns from long-term investments in a well-functioning market economy “prevents bad choices” (Pierson 2000: 256).

\(^{40}\) Taking the QWERTY example again, David (1985) argues that other keyboard designs, such as the Dvorak Simplified Keyboard (DSK), are actually better, offering superior speed. However, while many typists or companies might have chosen the DSK design in a theoretical one-off choice, once many others had already chosen the QWERTY layout, network externalities outweighed the value of technical superiority. Similar arguments are sometimes made about dominant software designs, such as Microsoft.

\(^{41}\) Picking up on David’s (1985) example of the QWERTY keyboard, they argue that the evidence for its technical inferiority relative to alternatives is weak (Liebowitz and Margolis 1990). However, as Farrell and Klemperer (2007: 2012) point out, despite the fact that network effects associated with keyboard design have weakened with the rise of computers and the decline of typists, even weak effects are enough to maintain the near complete dominance of QWERTY.
At its core, Liebowitz and Margolis’s argument is that expectations (i.e. people’s actions are influenced by what they think others will do in the future) outweigh history (i.e. people’s actions are influenced by what others have already done). In Arthur’s (1989) model, history plays the decisive role. Krugman (1991) provides a general framework for deciding under what conditions which of these forces predominates. He argues that the importance of history, relative to that of expectations, depends on adjustment costs. In the case of the adoption of a new technology, adjustment costs might be interpreted as the costs of acquiring information about that technology. If everyone already knows what they think and their choice is simply influenced by network externalities, then expectations will dominate in the determination of which technology will dominate. If adjustment is slow because people have to acquire information through learning from what others have already done, then history matters. In the case of energy technologies, where investments are often large scale and long-lived, it seems plausible that history will play a dominant role.

4.2 Low carbon technology races and carbon lock-in

The idea of increasing returns and the outcomes of path-dependence and lock-in are of particular relevance to two kinds of governance problem. One is how to manage situations in which two or more technologies are emerging that are not compatible, so that only one can come to dominate (Katz and Shapiro 1985: 424-25). In many examples of technology competition with network effects and adaptive expectations, the dynamics are driven by consumers choosing which option to buy into. In cases of physical networks, such as railroads or electricity, the issue is more about the choice of design of the network, which will be more a commercial or political decision, where technologies compete through lobbying (David and Bunn 1987). Within the context of sustainable energy, this situation may well apply to a contest between nuclear power, which is a stable but inflexible source of power operating at large scale, requiring a largely one-way transmission network, and renewable electricity, which is often intermittent but can operate at multiple scales and would benefit from a smarter network (e.g. Verbruggen 2008). Once one technology family is established, technology-specific complementarities with grid infrastructure are likely to crowd out the other family.

Another possibility is that one initially high-cost technology may have learning-by-doing and learning-by-using externalities, but that it can never come to exploit those externalities because it is crowded out early on by a competitor (see Kalkuhl et al 2012 for such a model applied to low and high carbon technologies). Such an approach requires that Leibowitz and Margolis’s assumptions about self-fulfilling and optimal expectations do not apply because of some form of market failure.
If choice of complementary infrastructure is determined by history and not expectations (see above), then events and timing matter, and the outcome can be far from what would have been best with hindsight. Establishing an early lead can create bandwagon effects and crowd out an incompatible competitor (Farrell and Klemperer (2007: 2035-36), Choi 1994, Choi and Thum 1998) if there are network effects, so process does to some extent resemble a fight. In markets where firms battle to shape consumer expectations about market leadership and therefore choices, the reputation of a firm may play a major role in establishing a bandwagon effect (Katz and Shapiro 1985, Farrell and Klemperer 2007: 2033). In situations where public policy has a role in network choice, then strength of lobbying or reputation of the competing industry would be the appropriate political analogy.

Katz and Shapiro (1986) study the case where there are two inherently incompatible technologies subject to network externalities, and where the choice between the two is affected by both by the history of other consumers’ choice to date and expectations about the future success of the competing products (see above section 4.1). They argue that, precisely because of first-mover advantage in markets with network externalities, firms with sole rights to a given technology will be willing to make it available initially at loss-making prices, because these losses can later be recouped at higher prices which include a monopoly rent. Such firms effectively “sponsor” a technology. Again, the political economy analogy would be the willingness of firms (or industries) to spend considerable resources on influencing a public policy or regulatory decision leading to policies or network design that favoured their product. In both cases, the outcome may be socially sub-optimal, in the sense that the sponsored technology wins the market regardless of whether it is superior in performance or other characteristics. However, in cases where there are two sponsors and a degree of forward-lookingness amongst consumers or public policy makers then the best technology wins.

A second governance problem is how to manage situations in which an existing technology or set of technologies is already locked-in, but where it is desirable to accelerate the process by which they are displaced by new ones, as is the case with a low-carbon transition. The complexities involved may be seen in Unruh (2000)'s detailed application of the concepts of increasing returns to the energy and transport systems that produce the vast majority of man-made carbon emissions.

Unruh argues that there is lock-in to existing systems at multiple levels. At the firm level, the emergence and domination of particular technologies leads the firms producing those technologies to invest heavily in specialised knowledge and complementary assets. Corporate strategy tends to focus on incremental improvement of the existing dominant technology rather than radically new alternatives (Unruh 2000: 181-82). At the level of systems, network
externalities play a major role in relationships between technologies, infrastructures and inter-dependent industries. Thus the rise of the automobile in the first part of the 20th century saw the accompanying rise of supply industries in petrol, rubber, glass etc., as well as the development of the road network, service stations and ultimately a series of developments complementary to the existence of the car, such as motels, supermarkets and ultimately suburbs (ibid: 822). Network externalities were increased where coordination was facilitated by technical standards and conventions (see also Farrell and Klemperer (2007: 2022-23). At the social level, energy and transport technologies create new constituencies who have an interest in the maintenance of those technologies. These can include professional associations, unions of workers in associated industries and groups of users and enthusiasts. Crucially, they also typically include institutions of training, which replicate and refine the body of knowledge underlying the dominant technology (Unruh: 823). Finally, at the level of public institutions, dominant technologies become institutionally locked-in via regulation and policy, frequently receiving subsidy, or effective protection from new potential competing technologies through legal or technical barriers (ibid pp 824-25).

Unruh labels the interacting totality of these different levels as ‘techno-institutional complexes’ (TICs), positing that increasing returns and lock-in operate at and between each level. He then argues that because electricity, heat and transport in particular rely heavily on fossil fuels in most economies, these constitute TICs in which there is ‘carbon lock-in’: “he carbon-based TICs discussed here are possibly the largest techno-institutional systems in history and therefore have no real precedent” (Unruh 2000: 828). In relation to successful innovation in new low-carbon energy and transport technologies, the point is that policy makers need to ensure not only that innovation support policies are in place, and that markets are sufficiently open to new entry, but also that whole the whole techno-institutional complex based on mutual benefits across dense and extensive networks of firms, other social actors and government, is transformed.

4.3 Escaping carbon lock-in

Lock-in, or ‘inertia’ is a powerful concept, and now widely used in the literature on low carbon transition. But lock-in is never absolute – there are plenty of examples where ‘disruptive’ technologies have broken through and displaced dominant technologies despite the fact that the latter benefitted from large network effects (David 2000: 11, Bresnahan and Greenstein 1999). In the history of energy, new fuels have broken through in a series of waves (wood to coal to oil

42 "Network effects, by binding together different users’ choices, might generate a stronger and more worrying form of inertia, locking society into an inefficient product (or behavior) because it is hard to coordinate a switch to something better but incompatible – especially where network effects coexist with individual switching costs." (Farrell and Klemperer 2007: 2028).
and electricity and more recently gas) despite the existence of established networks and supporting institutions for the existing fuel (Fouquet 2009). The key question is therefore what factors and conditions can lead to an escape from lock-in (e.g. Foxon 2011).

At a very basic level, a key condition is that the disruptive technology must in some way be better than the dominant one. It might in fact need to be a lot better - Intel's CEO Andy Grove has been quoted as saying that an improvement that is incompatible with existing networked systems must be “ten times better” to break through. In the case of sustainable energy, the superiority of the product in many cases lies mainly in its inherent characteristics (i.e. it is low carbon) rather than in the quality of its service. This makes breaking lock-in especially dependent on policy (see discussion below).

But even a much better (or policy-supported) technology faces problems in the face of lock-in, because it requires consumers to switch away from a pervasive, known existing network, or to persuade firms or governments to change existing networks. The former applies to electric vehicles, where a new “refuelling” network needs to be constructed alongside the existing petrol/diesel network. The latter applies to electricity, and is even more difficult because it means confronting vested interests directly rather than working around them. In some network industries, incumbents try to lock out potential new entrants by setting up exclusive club networks (Shapiro 1999). In the case of electricity, where basic economics and regulation has created a single monopoly network, incumbents may instead seek to exclude new technologies by shaping the technical and operating rules of the network to their advantage.

For an individual consumer, joining a new network involves both loss of network externalities, and may also involve switching costs, such as getting hold of information on the new option. Thus even if a new technology is considerably better than an existing one, and everyone would be better off if they used it, there may be a penalty for an individual switching (Farrell and Klemperer 2007: 2025-2032). Overcoming this may mean pricing new options at below cost (or at least with a discount) initially. For sustainable energy this implies that policy may have to be designed not only to produce price parity with conventional energy technologies or supplies, but actually offer investors and consumers greater profits and lower prices. The success of such a strategy may be seen in the example of the solar PV feed-in tariff in the UK, which on introduction in 2010 offered sufficiently high returns to overcome any barriers and saw rapid take-up.

Thus, in practice, many of the network effects identified by Unruh (2000) are not absolute, can be overcome and in many cases have been overcome. While energy networks and corporate institutions remain largely based on fossil fuel technologies almost everywhere and have yet to
be reconfigured to facilitate switchover to a sustainable energy system, there are no absolute technical barriers to connecting renewable energy technologies to existing networks; training in new technologies is being introduced, new professional associations are forming, support policies have been adopted, and so on. In some countries, notably Germany and some other European countries, this process is well under way.

4.4 Conclusions

The possibility of increasing returns introduces a new set of relevant issues for the governance of innovation. Applied to low-carbon innovation, with its specific characteristics, increasing returns point to the essentially political nature of the governance problem.

Ultimately, the problem of escaping lock-in is one of coordination. The more people can be coordinated to switch to an emerging technology, the lower the costs for each of them. Breakthrough by a new and disruptive technology is also more likely if innovators are untied and coordinated themselves; splintering amongst innovators tends to preserve the status quo (Kretschmer 2008). But coordination is not easy, and suggests a key role for a coordinating agent. Liebowitz and Margolis (1994, 1995) argue that entrepreneurs can effectively play such a role. However, as noted above, the case of carbon lock-in is different from that of conventional technological disruption, for the reason that that in many cases, low-carbon technologies do not really provide a new service, but rather that same service (electricity, heat, mobility etc,) with different production characteristics (i.e. they are low carbon) (Fouquet 2009). This difference marks out aspects of the sustainable energy transition from previous energy transitions – i.e. the latter offered tangible private benefits as well as wider social ones. It also means that while entrepreneurship may play a major role in delivering transition, breaking lock-in initially requires public policy.

However, as Unruh (2000) notes, public institutions themselves are locked in to the high-carbon techno-industrial complex, which implies that escaping lock-in must at root be a political project, involving a redefining of policy goals and a challenge to vested interests. It is for this reason that Unruh argues in a later paper that argues that such policy will only come about as the result of external political pressure by civil society (Unruh 2002). Such pressure will have to be substantial to be effective, since increasing returns apply to political dynamics even more strongly than to economies, and political lock-in is particularly powerful Pierson (2000). Where high-carbon interests have influence over or access to state resources and authority, while excluding rivals, they can be difficult to dislodge. Success is likely only where challenges to

43 For example, in the UK, the renewable energy lobby is quite splintered; there are a number of associations, including the Renewable Energy Association, which is further sub-divided into technology-specific committees, RenewableUK, the Solar Trade Association, and the Micro-Power Council. These associations have not always agreed on key policy questions, such as the desirability of feed-in tariffs.
lock-in can themselves harness increasing political returns – i.e. where coalitions press for policy changes that create new constituencies and interest groups that favour change (Lockwood 2013).

5. The Governance of Institutions

If coordination is the key to escaping lock-in, an interesting question is how the feasibility and ease of coordination varies with the nature of institutions,\textsuperscript{44} which vary widely even within modern market economies. In this section, I consider what institutional economics says about this issue, and the implications for the governance of innovation.

5.1 Institutions and institutional change

There are a variety of attempts at explaining why certain institutions and their economic consequences come about, persist and change (for reviews see Chavance 2009, Ménard and Shirley 2008, Alston 2008 and Brousseau and Glachant 2008). One literature focuses on why such institutions as property rights, contracts and firms exist in market economies, with an explanation centring on transaction costs (Coase 1937, Williamson 1985, North 1990) or the imperative to solve coordination or collective action problems (Ostrom 1990). North’s (1990) account also incorporates increasing returns, and leads to an understanding of stasis and change quite similar to that of Arthur (1989) and Unruh (2000); institutions represent equilibrium outcomes that are hard to change because they are locked in. They tend to remain quite stable for long periods. Tensions can build up as relative prices, terms of trade or technologies change, and then there can be discontinuous abrupt movement to a new set of institutions.

Given Krugman’s (1991) analysis noted above, it is not surprising that the transactions cost approach to institutions takes history seriously and embraces the notion of path dependence (e.g. North 1990: 92-104). A contrasting school of thought, applying evolutionary game theory, gives expectations the key role (Aoki 2001, Grief 1997, Dixit 2004). Rather than simply following rules, this view sees institutions as stable arrangements of consistent, self-enforcing expectations that evolve out of repeated interactions. Institutional change here is then not about changing rules but about expectations (Aoki 2001, Brousseau et al 2011). When shared beliefs no longer work as a guide to behaviour of others then the institution breaks down. Destabilisation of beliefs can be caused by external factors, including technological change, changes in prices, new information, etc. although a change in external environment that doesn’t change expectations will not lead to institutional change (Kingston and Caballero 2008).

\textsuperscript{44} Institutions can be understood both as constituted by both formal rules or laws and informal expectations.
However, destabilisation can also occur from within if institutions fail, and expectations of some parties are no longer met. This school of thought comes close in some respects to the analysis of sets of ideas and institutions as ‘policy paradigms’ in approaches from political science (Hall 1993, Campbell 1998).

A third view of institutions, mainly applied to understanding the dynamics of political patronage and economic underdevelopment in poor countries, takes a social conflict view. Institutions are not chosen by the whole of society as ways of solving transactions and collective action problems; rather they are chosen by groups that control political power at a given time (Acemoglu et al 2004, North 1981). The outcomes will not be institutions that maximise the size of the economic pie, but the slice of the pie taken by powerful groups. This implies that the efficiency of institutions cannot be separated from their distributional effects, and that political and economic institutions and outcomes are inextricably linked (Acemoglu et al 2004). As with the other explanations, this school provides a better explanation of persistence than change. Institutions are constructed by those with political and economic power to reinforce and maintain that power, so again there is a tendency to persistence. Change can happen because of price or technology shocks, or because of feedback effects that work between economic and political power. Groups without political power may find over time that economic changes make it easier for them to solve their collective action problem, gain de facto power and alter institutions.

These accounts of institutions and institutional change tend to treat technological innovation as an exogenous black box, as a cause of institutional change rather than a process determined by institutions, although Unruh (2000) and others have applied North’s ideas about the stability of dominant institutions to explain why low carbon transition is so hard. If institutions in energy and transport have evolved to solve existing problems of collective action and transactions problems, then they are not likely to be very suitable for solving new problems, like climate change.

Equally, it is possible to see how the social conflict view of institutions might be applied to an analysis of dominant high carbon interests and emergent low carbon interests, even if these might coexist within a single firm or government department. One key hypothesis in the social conflict approach to institutions is that where actors have de facto economic or political power, they will attempt to convert it into de jure political power, mainly because the latter is a more secure form of power (Acemoglu et al 2004: 63-64). Private companies can never gain de jure political power directly, but they can attempt to influence those with such power to create

---

45 Obverse to the socio-technical transitions literature, where economic and political institutions tend to be treated as a black box

46 Military force being an example of the latter
durable institutional arrangements to secure their economic power. In the energy sector, such arrangements, as discussed in section 3 above, would include regulations effectively increasing the cost of market entry and technical codes that favour the core activities of incumbent firms. In some cases, the committees responsible for technical codes in the energy industry are effectively run by incumbents, so they do come close to wielding political power. The practice of secondment from firms into government is another example of this influence. Such influence is not so much about the interests of specific firms, so much as the interests of the industry. In the social conflict view, however, the interests of the industry do not equate with the wider public interest.

5.2 Varieties of capitalism and low carbon innovation

At a wider level, one particularly influential application of institutional economics has been Hall and Soskice’s (2000) ‘varieties of capitalism’ (VoC) approach. They provide an analysis of interlocking economic institutions which produce distinctive national “political economies”, with implications for the nature of innovation associated with those political economies. Unlike Unruh’s framework, which focuses on the institutions governing energy and energy innovations, the VoC approach is about the deep structure of institutions across the whole economy.

Hall and Soskice are concerned with modern capitalist economies, and they place the firm at the centre of their analysis (2000: 5). Building on the transactions costs approach to institutions, they argue that firms face coordination problems in a number of different areas, five of which are particularly important (ibid: 6-7). First, in industrial relations, they must coordinate bargaining over wages and working conditions with their own labour force, unions and other employers. Second, in vocational training and education, they face the problem of how to secure a workforce with suitable skills. Third, they must organise corporate governance and accountability in such a way as to ensure access to finance. Fourth, they must coordinate relationships with suppliers and clients in ways that work effectively. Finally, they must organise their own employees in ways that incentivise them appropriately and in ways that give them enough information to function well.

At the core of Hall and Soskice’s thesis is the idea that there are different systems by which firms solve these problems, and that distinctive systems can be found in different countries. In particular, they distinguish two systems: liberal markets economies (LMEs) and coordinated market economies (CMEs). In LMEs, firms coordinate activities “primarily via hierarchies and competitive market arrangements (Hall and Soskice 2000: 8). Market relationships are

For example, the governance of the electricity codes that determine how the system is run is managed largely the firms in the industry (e.g. http://www.nationalgrid.com/uk/Electricity/Codes/gridcode/reviewpanelinfo/2012/)
“characterized by the arm’s-length exchange of goods or services in a context of competition and formal contracting.” The US and the UK are proposed as archetypal LMEs. By contrast, in CMEs, “firms depend more heavily on non-market relationships to coordinate their endeavours with other actors and to construct their core competencies”. This set of institutional solutions entails “more extensive relational or incomplete contracting, network monitoring based on the exchange of private information inside networks and more reliance on collaborative…relationships”. Germany stands as the paradigmatic CME.48

There are a number of implications of the VoC approach. One is that economies are institutional systems, where particular policies and practices work because they are elements of a wider whole. This means that taking particular elements out of the system and applying them in another system will not work well (Soskice 1997).49 Pollitt (2010: 39) makes this argument for renewable energy policy.

A second corollary of the VoC approach is that the comparative advantage of economies depends on their institutional structure (Hall and Soskice 2000: 37). Firms in one system do some things better and some things worse, than their counterparts in the other system. In particular, because they can shift from one line of production to another with fewer adjustment costs, in fluid labour markets, Hall and Soskice argue that firms in LMEs will be better at radical, Schumpeterian innovation, involving the development of new products. By contrast, firms in CMEs, which are tied into longer term relationships with their workers and other firms, will tend to invest more in Usherian, incremental innovation (ibid: 38-44). In support of their argument, Hall and Soskice point to evidence that the US dominates patents in areas involving radical innovation, such as biotechnology, ICTs and pharmaceuticals, where as Germany dominates patenting in industries such as civil engineering, machine tools and consumer durables, where incremental improvement of existing products is more common and important.

Thus in Germany, ‘patient’ finance is available as banks and firms have close long-standing relationships and share private information, whereas in the US and the UK firms’ access to finance is heavily dependent on valuation in equity markets and other publicly available information. Mergers and takeovers are more common and acceptable. In Germany, the management of firms depends partly on supervisory boards which involve other stakeholders including employees, while in the UK and US industrial relations rely on liberalised labour markets and hire-and-fire; trade unions are present but are weaker. German industrial relations deliver wage equality at each skill level, handled through consensus building, but in the US and UK wages are set through market competition. In Germany, vocational training is organised through industry-wide employer associations and unions, whereas in the UK and US fluid labour markets mean that firms are reluctant to train, so training takes place mostly in formal education system. Finally, in Germany, there are pervasive relations for facilitating technology transfer between firms, who tend to be organised in complementary ways through product differentiation and niche production rather than direct competition, contrasting with the US and UK where relationships between firms are mostly competitive and formally contracted.

48 This idea that policy does not travel well is also found in Rodrik’s (2007) approach to industrial policy. Institutional context matters.
Despite a number of criticisms of Hall and Soskice’s original model, the idea of institutional elements, with implications for how innovation works differently in different contexts, may have some application to comparative analyses of innovation for sustainable energy. To date, only a handful of studies exist. In a study of the evolution of electric vehicles, Mikler and Harrison (2012) argue that CMEs will be better than LMEs at innovation aimed at sustainability because firms and government are able to reach consensus on environmental goals through deliberation, while LMEs have to rely on arms-length regulation that industry will seek to game and erode. They apply this framework directly to a comparison of Germany and the US, arguing that the car industry in the former will have the edge over the latter. The view that aspects of Germany’s CME institutional system may also better facilitate innovation for sustainable energy may also find support in the experience of the Energiewende, including the role of local financial institutions and close deliberation and coordination between manufacturing industry, unions, civil society and the government (especially the Ministry of Environment) (Lauber and Jacobsson 2006).

However, it is not clear that wider evidence supports such a sharp difference. For example, Dechezlepetre et al (2011) find that both Germany and the US rank highly (3rd and 2nd respectively) in patents for mitigation technologies across a wide range of types of technologies. Over the period 1980 to 2007, Germany did have a small lead, with an average of almost 20% of global inventions across different mitigation technology classes, but the US had over 15%. There was also considerable overlap in the technology classes in which patents fell between Germany, the US and the UK.

On the other hand, patent data may provide some support for Hall and Soskice’s hypothesis that CMEs tend to specialise in innovation in established technology areas where innovation is likely to be more incremental in nature, while the US tends to specialise in more recent fields where innovation is more radical. Dechezlepetre and Martin (2010), examining patents for a wider range of ‘clean’ technology classes, find that Germany has a clear lead over the USA in the share of global patents in fuel injection for cleaner ICEs, geothermal energy, wind power, solar thermal, heating, cement, insulation and methane destruction. Conversely the USA leads Germany in biomass, marine energy, solar CSP, nuclear, lighting, batteries and fuel cells. They produce roughly similar proportions of global patents in electric and hybrid vehicles, hydro

---

50 The VoC framework has been much extended, debated and critiqued. Hall and Soskice’s predictions on patterns of innovation have been challenged (Taylor 2004, Akkermans et al 2007). The framework has been criticised for not having a role for governments. It has been argued that since Hall and Soskice put forward their arguments, Germany itself has changed, coming to resemble an LME more, especially in its labour markets. Crouch (2005) points out that the fit between LMEs, CMEs and particular economies is not absolute. Even the US, the paradigmatic LME, has sectors and industries that have more CME-like qualities, such as the defence industry. Crouch argues that Hall and Soskice’s institutional system should be seen as ideal types rather than particular economies, which are better understood as having elements of both systems, with one perhaps more dominant that the other, or applying to particular sectors.
power, waste-to-energy, and solar PV. Although the pattern is far from definitive, it appears that Germany is tending to specialise in innovation in more established areas (e.g. fuel injection, cement, heating and insulation), while the USA leads in several newer technology classes (e.g. marine, solar CSP and fuel cells).

5.3 Conclusions

The governance of innovation occurs largely through specific institutions, most of which are economic in nature. Different schools of institutional economics provide accounts of the rationale and persistence of institutional arrangements, and of institutional change. These accounts are based primarily on interests, or expectations about interests, and do not easily incorporate the influence of changes in ideas (such as climate science) within them.

However, one application of the transactions costs approach - varieties of capitalism – may provide a framework for understanding why new goals, such as sustainability, may be easier to adopt under certain institutional arrangements than under others. This difference may work in two ways. One is that actors such as firms and government may reach stronger agreement on goals, more quickly, in institutional systems based on principles of deliberation than in those depending on arms-length, rules-based relationships. A second is that institutional systems based on deliberation and long-term institutions of trust are better at managing the pervasive asymmetric information problems inherent in the transformation of a whole sector, like energy.


This paper has reviewed the implications of a number of areas of economic theory for the governance of innovation for sustainable energy. The review is far from complete, and it neglects issues at many other levels, especially at the level of politics and discourse. However, it does suggest a broad research agenda for understanding why innovation towards sustainable energy may proceed at different rates in under different governance arrangements, and also how those arrangements might change. It also shows the considerable range and scope of economic analysis that may be relevant for understanding innovation in sustainable energy. Finally, it shows the gap between the sometimes simplistic approach to economic analysis in the policy process and the subtlety of ideas available in economic theory more widely.

Neo-classical economics provides a clear theoretical rationale for policies to support innovation, based on a number of well-established market failure arguments. There are additional specific rationales for support to the development of low-carbon technologies given the urgency of
climate change and uncertainty about technology costs. At the same time, economics is also concerned with the costs of policy, and concerns about government failure, centring on asymmetric information and the risk of capture, have tended to lead many economists to argue for ‘technology neutral’ market-based support mechanisms as the most cost-effective. However, the case against technology specific support policies is not as strong as it appears, and the more general underlying question is how the policy process can overcome asymmetric information problems and the risk of capture to provide effective support without excessive rent. The answer to this question may vary with wider institutional settings, but it lies at the heart of understanding both differences in comparative experience and in how policy in the UK might be made more effective.

The focus on innovation policy is very common, but other areas of economic theory are also relevant. A clear message from the theoretical literature on market structure and innovation is that removing barriers to market entry and reducing the costs of market entry would be expected to increase the amount of product innovation in retail energy markets and technological innovation in the generation of electricity and the production of energy. Since the energy industry does involve quite large sunk costs, active policy is needed to maintain ease of entry. However, political economy approaches in economics suggest that such attempts to reduce barriers to entry will be resisted by incumbents. An important research question is how different institutional arrangements (and political forces) in different countries minimise the risk of regulatory capture, and indeed how institutional arrangements in the UK may be changed to reverse capture or minimise the risks.

The possibility of increasing returns and the perspective of evolutionary economics introduces a new set of relevant issues for the governance of innovation, and especially the problem of ‘lock-in’ within firms, in industry, in public institutions, amongst end users and more widely in social institutions. Applied to low-carbon innovation, with its specific characteristics, increasing returns point to the essentially political nature of the governance problem. Ultimately, the problem of escaping lock-in is one of coordination. The fact that sustainable energy technologies are often distinguished mainly by inherent characteristics than by new services implies that while entrepreneurship may play a major role in delivering transition, that coordination requires public policy, at least initially. However, since public institutions themselves are subject to lock-in, escaping it must at root be a political project. A key question is then how far different arrangements in economic and political institutions facilitate or block such a project. One application of institutional economics - varieties of capitalism – may provide a framework for understanding why new goals, such as sustainability, may be easier to adopt under certain institutional arrangements than under others. Actors such as firms and government may reach stronger agreement on goals, more quickly, in institutional systems based on principles of
deliberation than in those depending on arms-length, rules-based relationships. At the same time, institutional systems based on deliberation and long-term institutions of trust may be better at managing the pervasive asymmetric information problems inherent in the transformation of a whole sector, like energy than those based purely on market mechanisms that can be gamed.

All of these three areas of research would require detailed comparative case studies, examining evidence on how institutional arrangements affect the use of information, incentives for actors involved in governance arrangements, and coordination. Such case studies would have to use a mix of methods, incorporating not only collection and comparison of data on formal institutions but also on informal relationships, and the analysis of these in relation to specific outcomes (such as the rate of investment in renewable energy).

The agenda reviewed here is based on the view that: “...the energy sector is subject to multiple externalities like carbon emissions, local air pollution, innovation and learning spillovers, imperfect competition, network effects or energy security concerns” (Kalkuhl et al 2011: 2). This agenda contrasts with the much narrower agenda associated with what is effectively a neo-liberal view of economics, both by adherents and critics of that view. A neo-liberal view abstracts from the complexities of the energy sector, or takes the view that attempts to address them would end up doing more damage than the original problems pose. The main element of the neo-liberal agenda is a carbon price. Interestingly, many neo-liberals apparently overlook the fact that carbon pricing itself has suffered from multiple policy failures and political challenges (Laing et al 2013), and that government failure in carbon pricing may ironically be a bigger problem than in low carbon innovation policy. Moreover, where carbon pricing is weak or non-existent, as it still is in most cases, then this agenda is even more relevant.

A final point is that the agenda generated by economic theory and sketched out above has some points of similarities with the approach taken in the socio-technical transitions and multi-level perspectives literature. Not surprisingly, this is especially the case for the evolutionary economics literature, since the socio-technical systems view is partly built on evolutionary economics (Safarzyńska et al 2012). The discussion of innovation policy is relevant for the creation and management of “niches”, and the discussions of market structure and innovation, and of lock-in, are relevant both for understanding “regimes” and how niches come to challenge regimes. The value of economic theory in this context is that it provides a finer grain theoretical analysis, and hypotheses about the effects of specific kinds of institutional or policy change.
7. References


Knight, F. (1921) Uncertainty and Profit New York: Augustus M. Kelley


Lin, J. and Chang, H-J. 2009 (forthcoming). Should Industrial Policy in Developing Countries Conform to Comparative Advantage or Defy It? – A Debate between Justin Lin and Ha-Joon Chang, Development Policy Review


Pindyck, R. (2011) "Fat tails, thin tails and climate change policy" *Review of Environmental Economics and Policy* 5, 2: 258-274


Ruttan, V. (1959) “Usher and Schumpeter on invention, innovation, and technological change” *Quarterly Journal of Economics* 73, 4: 596-606


Tingvall, P. and Karpaty, P. (2011) “Service sector competition, innovation and R&D” Economics of Innovation and New Technology 20, 1:


