ARTICLE FOR OXREP ISSUE: GOVERNMENT AND BUSINESS: A NEW INDUSTRIAL POLICY?

AGGLOMERATION, CLUSTERS AND INDUSTRIAL POLICY
MAX NATHAN¹,² AND HENRY OVERMAN¹

¹ SERC AND LSE
² NATIONAL INSTITUTE OF ECONOMIC AND SOCIAL RESEARCH

ABSTRACT [147 words]

This paper considers the appropriate spatial scale for industrial policy. Should policy focus on particular places, targeting clusters of firms that are spatially concentrated? Or should it, instead, be ‘space neutral’ refusing to discriminate between different areas unless absolutely necessary? We provide an overview of the literature and identify two waves of literature that argue strongly in favour of a cluster approach. We argue that this approach rests on shaky theoretical and empirical foundations. In contrast, we suggest that more attention should be paid to the appropriate spatial scale for horizontal interventions. What can policy do to make cities work better, in ways that help firms to grow? That is, what is the appropriate role for ‘agglomeration’ rather than ‘cluster’ policy? Finally, we consider the possibility that some horizontal industrial policy objectives may be better served by specifically targeting particular places or from decentralised design or delivery.

ACKNOWLEDGEMENTS

This paper draws on NIESR research on industrial policy (funded by LLAKES) and on a number of research projects at the Spatial Economics Research Centre (funded by the ESRC, BIS, DCLG and the Welsh Assembly Government). We would like to acknowledge the support of the funders, and to thank Geoff Mason at NIESR for helpful discussions. The opinions in this paper are those of the authors, not the funders. All errors and omissions are our own.
I. INTRODUCTION

Recent years have seen a re-emergence of interest in industrial policy. As detailed elsewhere in this issue [link to other papers] policy makers are hunting for new approaches to industrial policy, and questions are once again being asked about its appropriate scope and scale. Should the stance be ‘vertical’ and target particular firms and sectors? Or should it be ‘horizontal’ and focus on the general conditions that affect industry? Paralleling this debate, and every bit as controversial, are arguments about the appropriate spatial scale for intervention. Should policy focus on particular places, possibly targeting (inter-linked) clusters of firms that are spatially concentrated? Or should it instead be ‘space neutral’, refusing to discriminate between different areas unless absolutely necessary?

This paper is specifically concerned with the appropriate spatial scale for policy. We are certainly not the first to consider such questions. However, renewed discussion would appear timely because major new initiatives, including Regional Innovation Clusters in the US, the European Commission’s adoption of ‘smart specialisation’ regional policies, and (closer to home) the UK’s Tech City initiative, all have strong cluster components.

A number of factors – both global and country specific – underpin the renewed interest in industrial policy. As argued elsewhere in this issue, the broad contextual factors include: globalisation (including the rise of China which is seen as having a strong industrial policy); climate change, and a need for new infrastructure, goods and services which account for environmental externalities; the current recession, and many national governments’ search for growth mechanisms; and a re-evaluation of evidence on the nature and performance of past industrial policies (in particular, Rodrik (2004) and following), which has accelerated since the recession (see for example Aghion (2012; 2012).

Many of these general factors are at play in the UK too, but local factors also play a role. For instance, the UK government has historically taken an interest in certain sectors (e.g. aerospace and defence, automotive). This historical focus has received renewed attention with the current Coalition’s aims to ‘rebalance’ the economy in both spatial and sectoral terms. To help achieve this Ministers have signalled a return to ‘industrial strategy’, with a strong sectoral component (Cable 2012; Department for Business Innovation and Skills 2012). In turn, this approach has links to the ‘New Industry, New Jobs’ agenda developed by Lord Mandelson in the previous Labour administration (Department for Business Enterprise and Regulatory Reform 2009). Independent voices such as the LSE Growth Commission have also suggested that Government take a more activist role in driving growth, through more effective investments in human capital, strategic infrastructure and the national innovation system (Aghion, Besley et al. 2013).

As with the broader debate, the increased interest in cluster and area-based industrial policies also reflects a mix of global and country specific factors. In particular, in many countries there is renewed interest in the economic role of cities and the importance of agglomeration economies. Partly, this has been driven by the very rapid urbanisation seen in China, India and other developing countries. Partly it reflects the re-emergence of the city as a central driver of economic performance in service orientated developed countries which had, until recently, experienced rapid urban decline as
structural change affected many former industrial cities (see, inter alia Boddy and Parkinson (2004), Buck et al (2004), Overman and Rice (2008) and Glaeser (2011)). At the same time, just as with the broader debate, country specific factors also play a role. In the UK these have been reflected in a raft of policy initiatives – City Deals, Local Enterprise Partnerships, the Heseltine Review – that have embodied the coalition’s commitment to greater decentralisation of the highly centralised British State (Department of Communities and Local Government 2011; Cabinet Office 2012; Lord Heseltine 2012).

In this paper we argue that there are two important paradoxes about the return of cluster policy. First, there is certainly good empirical evidence that agglomeration and co-location matter a great deal in understanding the economic performance of firms and cities. But in moving from this general observation to specific policy recommendations, advocates of cluster policies draw on conceptual frameworks which are remarkably fuzzy – both in terms of providing a tool for analysis and developing specific policy interventions. This would not matter so much if we had good evidence on which policy interventions were effective, which brings us to the second paradox – cluster policies appear to be generally ineffective and have been robustly debunked in the academic literature, yet remain very popular with policy makers.

As we discuss below, much of this second paradox is relatively easy to explain – clusters are a powerful brand, even if cluster thinking often leads to poor policy making (Martin and Sunley 2003). These problems have been widely discussed in the literature, but this has not stopped the emergence of a ‘second wave’ of cluster thinking – a new set of conceptual tools that underpin the recent renewed interest in cluster policy. As we will make clear, we view this second wave of cluster thinking as providing analysis that is conceptually distinct, but no more robust, than that underpinning the first wave of cluster policy, and thus no more likely to succeed. Does this mean that there should be no spatial dimension to industrial policy? We think not, but argue that local strategies should target individual firms and workers, rather than cluster level outcomes and are likely to be limited to a fairly specific set of sectors. We suggest that more attention should be paid to the appropriate spatial scale for horizontal interventions. What can policy do to make cities work better, in ways that help firms to grow? That is, what is the appropriate role for what one might call ‘agglomeration’ rather than ‘cluster’ policy? Finally, we also consider the possibility that some horizontal industrial policy objectives may be better served by specifically targeting particular places or from decentralised design or delivery. As discussed below, however, the literature on which aspects of policy may benefit most from such an approach is surprisingly underdeveloped.

The rest of this paper is organised as follows. Section II considers the broad literature available to help structure thinking about industrial policy, with a particular focus on the spatial scope of that policy. Our aim is certainly not to be exhaustive, focusing instead on providing the background against which to assess the additionality of the conceptual frameworks that are used to underpin cluster thinking. Sections III-IV turn to the first of these conceptual frameworks, which we refer to as wave one, and which is widely associated with the work of Michael Porter. This section also runs through some well-established criticisms of this approach. We then turn our attention, in sections V-VI, to wave two of the clusters literature. This literature has received far less attention, but we argue that it is as problematic as wave one in terms of its usefulness as a guide to policy. In section VII, we set out some alternative ways in which industrial policy might become more spatial.
II. FRAMEWORK

Traditionally, debates about the scope of industrial policy have focused on the question of whether it should be sectoral or horizontal. Krugman and Obstfeld (2003) adopt the sectoral definition: ‘Industrial policy is an attempt by a government to encourage resources to move into particular sectors that the government views as important to future economic growth.’ This approach is associated with the targeting of specific firms and sectors – frequently referred to as ‘picking winners’ (often used ironically, given a history of failure to do any such thing) (Owen 2012). In contrast horizontal approaches emphasise ‘neutrality’ and focus on policy interventions – e.g. R&D tax credits or training policies – that could in principle benefit many firms or industries (Aiginger 2007).

Horizontal definitions of industrial policy can become very wide, to incorporate (for example) education and skills policy, attitudes to red tape and regulation and just about anything else that might impact on economic performance. For the purposes of talking about the spatial scope of industrial policy, we restrict the focus to keep things manageable. To do this, we focus on interventions that are directly intended to increase productivity or employment (or both). In many developed countries, innovation policy is seen as key to increasing productivity growth, while business support and enterprise policy are two of the most common tools aimed at increasing both productivity and employment.

For many economists, research on endogenous economic growth has provided a central conceptual framework to help structure thinking about policies intended to stimulate innovation, productivity and employment. More heterodox approaches have drawn on theories of innovation systems and Schumpeterian models of entrepreneurship. The implications of these theories for industrial policy have been widely discussed in the literature (Jaffe 1986; Navarro 2003; Rodrik 2004; Harrison and Rodríguez-Clare 2009; Wade 2010; Weiss 2011; Aghion 2012; Aghion, Dewatripont et al. 2012). For our purposes, we are specifically interested in theories that help structure thinking on the spatial scope of industrial policy. Just as general theories of innovation, productivity and employment growth provide the most useful way of thinking about industrial policy, we believe that general theories that try to explain the spatial structure of the economy provide the best starting point for discussions about the spatial scope of that policy.

Urban economics provides one central model for thinking about the spatial structure of the economy (Combes, Duranton et al. 2005; Overman and Leunig 2008). In urban economics, spatial concentration generates ‘agglomeration economies’ that help firms become more productive. These advantages include a critical mass of workers and infrastructure, and dense networks of suppliers and collaborators. Cities also help new ideas to form and flow, so that firms and workers can learn from each other. Offsetting these benefits are rising costs (e.g. increased land prices or congestion) as economic activity concentrates. At least initially, agglomeration benefits must outweight costs, otherwise all economic activity would be geographically dispersed (an outcome that we do not see in practice). However, at some point as a particular city grows, costs will begin to outweigh benefits and workers and firms will be better off choosing to locate elsewhere. In this model, cities of different sizes emerge when agglomeration benefits differ across economic activities. Differences in
congestion costs may also play a role, although much of the focus there has been on policy induced differences arising from land use policies.

Given the central role of agglomeration economies a large literature has focused on trying to understand these economies. One crucial question is whether these benefits arise from the concentration of similar or diverse activities. If the benefits arise from the concentration of similar activities, the literature often refers to these as localisation economies. Such localisation economies lead to the clustering of similar activities and to cities that tend to specialise. In contrast, the literature often uses the term urbanisation economies to refer to the advantages arising from the concentration of diverse activities. Urbanisation economies tend to lead to diverse cities. Basic data on the structure of cities clearly suggests that both sets of forces may be at work depending on the activity. For example, global cities such as London support a range of different specialisms, from finance, law and business support through to a large arts and creative economy. More rigorous theoretical modelling and econometric evidence that seeks to directly identify the magnitude and scope of agglomeration economies, confirms these basic intuitions (see Duranton and Puga (2004) for an overview of theoretical approaches, and empirical work by Melo et al (2009) and Combes et al (2008) for the UK, Rosenthal and Strange (2004) for the US, Mion and Naticchioni (2009) for Italy and De La Roca and Puga (2011) for Spain among many others).

The discussion so far has been essentially static, concerning predictions about the relative size and structural composition of different cities at any point in time. While these issues continue to receive considerable attention in both the theoretical and empirical literature, attention has turned more recently to questions of dynamics – what explains the growth of cities and the changing spatial structure of the economy over time?

In the urban economics tradition, formal modelling of the dynamics of city growth is in its infancy (see Desmet and Rossi-Hansberg (2009) and Henderson and Venables (2009) for recent contributions) and so the conceptual framework for understanding dynamic issues by necessity draws on a variety of sources. New Economic Geography models (NEG) provide some insights, particularly on the importance of changing transport costs in understanding the changing spatial structure of the economy. The core NEG insight is that a combination of firm level increasing returns, monopolistic competition and (iceberg) transportation costs for goods generate a home market effect whereby, all else equal, firms prefer to locate in larger markets. That is, NEG models provide consistent micro-foundations for the kind of agglomeration economies that had hitherto been simply assumed in most urban economics modelling. As in urban economics, these agglomeration economies, notably upstream-downstream linkages, lead to the clustering of related economic activity (Krugman and Venables 1995). Conversely, congestion, pollution and competition encourage dispersion.. Overall, the balance of ‘centrifugal’ and ‘centripetal’ forces determines the location of economic activity (Fujita, Krugman et al. 1999). Changing transport costs change the balance of these forces and hence the overall spatial structure of the economy.

Recent additions emphasise that the clustering process is characterised by feedback loops, so that existing agglomerations often have first mover advantage (Krugman and Obstfeld 2003); however, technological change and sectoral differences also tend to produce ‘production jumps’ from higher to lower cost regions (Venables 2006). These jumps occur within sectors as well as between them:
Venables gives the example of a financial services firm with offshored call centres, IT services outsourced to local partners, an international network of retail branches and a London-based headquarters. These complex production chains require careful co-ordination, and can imply high search, transaction and management costs (McCann 2008; Saxenian and Sabel 2008). Most recently, the ‘globalisation of innovation’ has seen the international re-organisation of increasingly high-value, ‘knowledge-intensive’ activities (Mudambi 2008).

One issue central to these literatures concerns the existence and nature of knowledge spillovers: that is, the extent to which spatial concentration facilitates learning (among workers) and innovation (for firms), and how this influences the spatial structure of the economy. Using patent citations and R&D data respectively, Jaffe et al (1993) and Audretsch and Feldman (1996; 1996) provided some of the first formal evidence that such knowledge spillovers could attenuate very rapidly with distance. Other researchers (Ellison and Glaeser 1997; Duranton and Puga 2005; Peri 2005; Crescenzi, Rodriguez-Pose et al. 2007; Lobo and Strumsky 2008; Breschi and Lissoni 2009; Ellison, Glaeser et al. 2010) have considered the scope of such spillovers and specifically whether diverse or specialised environments lead to faster growth (sometimes framed as comparison between Jacobs and Marshall-Arrow-Romer externalities). Of course, it is possible that production may benefit from specialisation, while innovation benefits from diversity (or vice-versa). This raises the possibility that firms may wish to move across cities as they evolve – something that Duranton and Puga (2001) formalise in their theory of nursery cities, provide evidence for their theory using data from France.

A second important dynamic issue concerns the role of competition in shaping urban market structure. Cities’ big upstream and downstream markets are a source of agglomeration economies; however, large markets also attract a large number of firms. Competition between firms should lead the least productive companies to exit the market (Melitz 2003; Melitz and Ottaviano 2008). Empirical research suggests that while both agglomeration and selection operate to raise productivity over time, the former is stronger (Combes, Duranton et al. 2012). Entry and selection may also influence innovation, in ways that matter for policymakers. New firm entry may generate straightforward competition, or lead to knowledge spillovers between entrants and incumbents (Markusen and Venables 1999); empirical work suggests that spillovers are strongest in supply chain linkages, and where incumbents have ‘absorptive capacity’ (Javorcik 2004; Meyer and Sinani 2009). As with Melitz models, the selection process itself may be innovation-enhancing, if stronger incumbents innovate to escape competitors (Aghion, Blundell et al. 2009).

A third set of dynamic issues concerns the role of history and ‘first nature’ geography, chance and path-dependence in explaining urban clustering and the growth of sectoral milieux. These issues are less developed in the urban economics literature than in the evolutionary branches of economic geography that we discuss in Section V. But it is worth noting that urban economics and NEG frameworks are fully capable of incorporating these concepts. For instance, Zucker and colleagues (1998) explore the emergence of the US biotech industry in the 1970s and 1980s. They link the historical presence of star scientists, individual universities and federal research funding to localisation economies suggestive of strong knowledge spillovers. More recently Glaeser et al (2012)

1 Thompson and Fox-Kean (2005) question Jaffe and colleagues’ methodology, and show that geographic spillovers (as represented in patent citations) are less extensive than previously thought. More recent work typically incorporates Thompson and Fox-Kean’s improved tests.
look at the long-term impact of heavy industry on current day patterns of entrepreneurship and employment growth. They find strong support for Chinitz-type models, in which areas’ entrepreneurial cultures (or lack of) persist over several decades.

What lessons can we draw from all of this with respect to the appropriate spatial scope of industrial policy? As with theories of endogenous growth, general models that explain the spatial structure of the economy provide a rich theoretical framework and much empirical evidence on which to base discussion around the spatial scope of industrial policy. One basic insight is that spatial disparities and other institutional features tend to persist over time, which raises questions about governments’ ability to ‘rebalance’ national or local economies. Another is that spatial economies are complex, with important differences between sectors, functions and areas, which themselves evolve over time and are subject to unpredictable shocks. From Jane Jacobs (1961) to the evolutionary theorists we discuss below, urbanists have compared cities and clusters to organic systems. For industrial policy designers, this complexity implies that policy interventions at the ‘system’ level may have unexpected consequences.

We return to these points throughout the rest of the paper. For now, however, we turn our attention to the question of what, if anything, cluster theory adds to this framework and body of empirical evidence. We use ‘cluster theory’ to refer to research that specifically seeks to explain the formation and evolution of geographically concentrated sets of linked firms in the same or closely related sectors. Using this definition it is possible to identify two waves of cluster thinking. The first, which we dub ‘wave one’, is most closely associated with Michael Porter. This wave, which concentrates on mapping and static modelling of existing clusters, has received considerable attention in the academic and policy literature. The second wave is associated mainly with ‘evolutionary’ approaches to economic geography, and has a more theoretical flavour. This second wave has, to date, received far less consideration in the economic literature. Critically evaluating these two waves of cluster thinking is crucial, because it is these literatures, rather than the more economic approaches discussed in this section, that underpin the development of most spatially-based industrial policies.

III. CLUSTERS, WAVE 1: PORTER

At the heart of wave 1 of cluster thinking lies Porter’s famous ‘Diamond’, which purports to map out the underlying sources of competitiveness. It was originally applied to nations (Porter 1990). The later focus on clusters arose because of the suggestion that diamond forces were stronger when activity was geographically concentrated (Porter 2000; Porter 2003). According to Porter, clusters might arise for historical or geographic reasons, but their key role in competitiveness means that ‘it is clusters that drive economic development’ (Porter 1996) [p57] by enhancing the ‘microeconomic business environment’, which in turn helps increase productivity and innovation, and stimulates firm entry and entrepreneurship (Porter 2000).

The Diamond model applied to clusters appears to be complex, with many different elements all feeding positively back on one another. As Martin and Sunley (2003) note, the existence of these positive feedbacks is one thing that helps make the framework so appealing to policymakers.
Reinforcing these positive feedbacks is itself positive, justifying an active cluster policy and a range of ‘innovative’ cluster support measures. Porter proposes that national economic development policies should be re-orientated around clusters, arguing that traditional sectorally-focused industrial policies risk being captured by vested interests. The new strategy should be based on ‘identifying and upgrading’ clusters, which entails ‘a recognition that a cluster is present’, followed by ‘removing obstacles, relaxing constraints and eliminating inefficiencies that impede productivity and innovation’ (Porter 2000 [p26]).

Some of the specific interventions proposed may be sensible (developing an evidence base for the cluster, promoting university-industry links, improving local infrastructure), given what we know from the wider evidence (see Harrison and Rodriguez-Clare 2009). Other interventions are considerably more speculative (cluster-orientated free trade zones and industrial parks, focusing export promotion and government departments around clusters). Porter also places a great deal of emphasis on local and area-based actions, even for high-value tradable sectors: ‘in low-productivity economies, the focus is heavily on foreign markets. Advancement requires the development of more local markets’ (ibid, p.21)

Despite its widespread appeal amongst policy makers (and some academics), most mainstream economists and a number of leading economic geographers, are sceptical, if not hostile, to this approach. The first, well-documented problem is one of definition (Gordon and McCann 2000; Martin and Sunley 2003; Duranton 2011). Just what is meant by a cluster? The literature provides a large number of rather vague answers to this question and Porter himself is no clearer: ‘drawing cluster boundaries is often a matter of degree and involves a creative process’ (Porter 2000) [p17].

Even if this objection could be met by tightening the definition, a second, more significant problem remains. As noted by Duranton (2011) “the [main] problem with the cluster policy literature is one of a lack of well-articulated theory: what is the ‘problem’ that cluster initiatives are trying to fix?” This problem lies at the heart of economists’ objections to this approach.

A common answer to the question about the role of cluster policy is that it aims to improve local “competitiveness” or productivity (although it is not obvious that these are equivalent concepts). The broader problem with this answer is that it does not clearly set out the source of any possible inefficiencies (or inequities) and thus cannot explain how to correct for them. Porter’s Diamond model may appear complex, but this complexity is actually rather superficial as all of the different elements feedback positively to other elements. Of course, in reality this will not be the case because of negative feedback – of the kinds suggested by urban economics and NEG frameworks. As discussed in Section II, reducing barriers to entry in a sector might lead some incumbents to innovate and others to exit (Aghion, Blundell et al. 2009). The overall outcome may be to increase innovation but the effects on employment are ambiguous: this matters for policymakers who are concerned about employment and jobs growth, as well as the fortunes of ‘native’ firms. It would be possible to identify many other examples where carefully

---

2 Porter (2000) briefly worries about ‘lock-in’ to specific technological pathways, and lists various cluster constraints, but these are all removable through policy actions.
specified economic models and available empirical evidence point to a negative, rather than a positive, feedback between different elements of the Diamond.

The third problem with the Diamond model is that, despite its apparent complexity, it pays no attention to some fundamental drivers. For example, what is being assumed about labour or firm mobility? If firms are mobile, but workers are not, how can one be sure that encouraging larger clusters in a particular place is a good idea? Similarly, what is being assumed about the functioning of the land market? It is quite possible that any surplus generated by increasing the size of the cluster just translates into higher rents for owners of land. Models of urban economics show that the answers to such questions are fundamentally important in understanding the functioning of the spatial economy and in assessing the role, if any, for policy. Yet the Diamond model is silent on these issues. This is particularly important in the UK context, where planning for housing and commercial land use is one of the key policy levers available at the sub-national level.

Finally, even if there is positive feedback between the different elements of the Diamond model, this does not actually provide a justification for policy intervention. Such a justification needs to be based on carefully identifying reasons why the market ignores these positive feedbacks and produces an inefficient outcome. That is, we need to look for market failures and construct policy to address them accordingly. Note, however, that the existence of several un-priced externalities makes it very difficult to know what a cluster-level policy should be seeking to do in practice. Cluster advocates essentially only think about one of the three types of externalities at work in clusters (agglomeration externalities) while ignoring the other two (cost externalities and coordination failures). We would argue strongly that this is not a good basis for policy.

There are two other related problems for those seeking to make policy on the basis of cluster analysis. First, the emphasis on ‘local’ actions – however defined – encourage easy-to-implement physical or area-based initiatives (such as zoning a neighbourhood as a ‘tech quarter’, or constructing a flagship building). But this risks missing the more pressing challenges that firms actually face. Second, Porter is almost wholly silent on how to deliver the policies he recommends, even down to which level of government should be involved in which functions. Given the conceptual and measurement problems facing clusters, this may produce competing initiatives, co-ordination problems or both.

These conceptual issues might not matter if cluster advocates could point to rigorous empirical evidence that suggested that feedbacks were unambiguously positive and so supported the cluster upgrading approach. Unfortunately, cluster concepts are sufficiently vague and hard to measure that the mechanisms have been subject to little, if any robust empirical modelling (Martin and Sunley 2003). Even this might not matter if we had (reduced form) evidence that carefully designed cluster policy had a positive ‘treatment effect’ on either the size of clusters or their competitiveness. Unfortunately, an early meta-survey of more than 750 clusters (van der Linde 2003) finds little evidence of any positive effects of cluster policy. Instead, the evidence suggests that government interventions are one of “the least important determinants in competitive clusters, while they play a much more important role in uncompetitive clusters” (p.144). Cluster creation policies perform even worse: only one competitive cluster has been established as a result of a specific government policy to attract it.
This case study evidence is reinforced by more recent econometric evaluations. Empirical studies typically find no evidence for any beneficial impact of Porter-type cluster policies (McDonald, Huang et al. 2007; Martin, Mayer et al. 2011; Yu and Jackson 2011; Huber 2012; Kline and Moretti 2012). In the limited number of cases where positive impacts are found, benefits may be small relative to costs. For example, Falck et al (2010) find positive effects of Bavarian cluster policy on firms' innovative activity, but note that private R&D declined, suggesting additionality may be limited. They also note the huge expense of the state’s "High Tech Offensive" programme: €1.35bn. In short, even if we wanted to, simplistic implementation of cluster policies appears to do little to create or increase the competitiveness of clusters.

IV. CLUSTER POLICIES: A TECH CITY EXAMPLE

Many of the limitations of ‘cluster thinking’ are evident in the UK government’s Tech City initiative, which is focused on the digital economy cluster around Shoreditch, East London. Since the early 1990s, the area has been a home for new media and creative industries, some of which relocated from the West End and Soho. Over the past decade and a half, a large number of ICT and digital content firms have appeared, centred on Old St roundabout, or ‘Silicon Roundabout’. Conservative estimates for 2010 suggest at least 1,500 firms in the three wards, Clerkenwell, Hoxton and Haggerston, which surround the roundabout (Nathan, Vandore et al. 2012).

Many of these businesses have close functional links to each other, and to the wider creative economy. For instance, there is considerable overlap between tools developed in coding and software activity and their application by ‘creative digital’ firms in advertising and marketing who work across online and physical platforms (Foord 2013). Qualitative research suggests strong localisation economies – driven by shared amenities (such as the edgy environment, trendy bars, cafes and restaurants), linkages between producers (generally production side collaborations) and knowledge spillovers (helped by proliferating ‘meetups’ and professional networks, as well as serendipity) (Nathan, Vandore et al. 2012). In short, it is a cluster, as most people would understand it.

The Tech City initiative is aimed at accelerating the growth of this cluster, and is seen as symbolic of the Government’s new approach to industrial strategy. Launching the programme, Prime Minister David Cameron announced that:

“Our ambition is to bring together the creativity and energy of Shoreditch and the incredible possibilities of the Olympic Park to help make East London one of the world’s great technology centres.” (Cameron 2010).

Eighteen months later, the Chancellor stated that:

“When people ask: give me an example of the Government’s industrial strategy I say this: we want nothing less than to make the UK the technology centre of Europe … Tech City [is] at the heart of this ambition.” (Osborne 2012).
A number of the ‘cluster thinking’ problems we identified above are evident here. First, the definition of the cluster is elastic at best – stretching from the existing concentration of firms in Shoreditch to an area five miles further east, where the data indicates very few digital economy firms are currently based. Worse, research with Shoreditch-based companies indicates little enthusiasm for relocating to the Olympic Park, and some cynicism about political motives for connecting the two (Nathan, Vandore et al. 2012). It may be the case, that the iCity site within the Olympic Park is well suited as a campus for large organisations - BT, data centre firm Infinity and Loughborough University have all announced offices there - but as yet there are no meaningful connections with the existing Shoreditch core.\(^3\)

Second, the overall strategy for the cluster shows a number of tensions and confusions. It involves a mix of horizontal interventions aimed at SMEs and entrepreneurs (such as the availability of an Entrepreneur Visa\(^4\) and a range of tax incentives aimed at encouraging seed and venture finance), investment programmes aimed largely at multinationals, and area-based initiatives – including the remediation and promotion of Olympic Park sites. The structure of, and governance arrangements for, these interventions are not clearly defined, with formal responsibilities scattered across central government, local government and local agencies.

As discussed, there is no obvious economic logic connecting the Olympic Park to Shoreditch. Policymakers clearly hope to repeat the success of Canary Wharf in encouraging financial services firms to relocate out of the City, but location patterns for the digital economy sector in London suggest a number of alternative locations exist, offering both cheaper sites for small companies and more expensive, high-specification spaces for larger businesses. More worryingly, if the intention of the policy is to promote the growth of local firms in the existing cluster, the focus on inward investment is misplaced. While some forms of FDI are linked to gains for ‘indigenous’ firms (via production complementarities, spillovers, or competition-induced innovation), encouraging entry may be just as likely to displace local businesses. In the face of such criticisms, the Tech City Investment Organisation has started to target potential investors who provide auxiliary services for technology firms (such as specialist finance) as well as developing a stronger emphasis on export promotion for locally-based businesses (Tech City Investment Organisation 2012).

Third, the strategy is also likely to have unintended distributional effects. Raising the area’s profile in this way has helped firms in the cluster sell themselves to customers and investors, in ways that may help them to grow. It will also encourage new firms to enter, which as we have seen has ambiguous effects on incumbents. There are also clear signs that the area’s landlords and developers are starting to respond to these developments by raising property costs, potentially displacing a number of smaller, younger firms (Nathan, Vandore et al. Forthcoming). While clusters are mobile in the long term, as we argue in the next section, displacement may involve substantial short term co-

---

\(^3\) The co-working space provider TechHub is also planning to open a space on the site, which should introduce some smaller firms. However, this still does not create the functional linkages of a working cluster, and given the geography there is no meaningful co-location with the Shoreditch core.

\(^4\) The UK has no restrictions on migration from European Economic Area countries. For non-EEA migrants a Points-Based System was introduced in 2008, with quotas for each active tier introduced in 2010. The Entrepreneur Visa is designed for highly skilled (Tier 1) migrants who have a business idea and can prove they have at least £50,000 of backing finance. 645 Entrepreneur Visas were granted between 2008 and 2011.
ordination costs. To repeat – cluster thinking fails to fully think through issues of costs and mobility and these distributonal impacts of the Tech City policy demonstrate the consequences of this.

V. CLUSTERS, WAVE 2: DYNAMIC MODELS

Dissatisfaction with the simple, static cluster models associated with Porter has led many researchers to develop richer models that focus more on the dynamics of clusters. It is this ‘second wave’ of cluster thinking that is underpinning recent developments in cluster policy. Its roots are in developments in regional innovation systems thinking, and in evolutionary economic geography.

Originally defined by Freeman (1987) as ‘the network of institutions in the public and private sectors whose activities and interactions initiate, import, modify and diffuse new technologies’, innovation systems are now viewed broadly as including social institutions, education and communications infrastructures and the norms and rules that regulate economic and social interaction (Lundvall, Joseph et al. 2009). ‘Regional innovation systems’ models (RIS) apply these frameworks to specific regions and clusters (Saxenian 1994; Cooke 2002; Asheim and Gertler 2005; Asheim, Smith et al. 2011). The basic argument is that given the existence of agglomeration economies, area-level institutions and networks substantially influence firms’ capabilities. Developing this insight, Storper (1997) suggests that regional outcomes are governed by three ‘spaces’ – territory, organizations and technologies and the local ‘untraded interdependencies’ that regulate agents’ behaviour.

Given the emphasis on institutions and interactions, RIS studies play close attention to describing the key agents in a cluster and their relationships, which are seen as governing the evolution of the cluster. These include universities and public agencies, networks (e.g. public-private partnerships) and social institutions (rules, customs and norms). Beyond this, national-level institutions (such as legal frameworks or spending programmes) and sectoral factors (industry-specific conditions or technological trends/shocks) are argued to set underlying operating conditions.

Evolutionary economic geography (EEG) has distinct intellectual foundations to RIS, as it draws on both the thinking of urbanists like Jane Jacobs (1969; 1984) and evolutionary economics (Nelson and Winter 1982). In particular, EEG frameworks are highly reliant on the use of notions and metaphors drawn largely from the biological sciences (hence the ‘evolutionary’ tag line). While research in this field is not uniquely focused on clusters, there has certainly been a strong theoretical and empirical emphasis on understanding how geographical clusters of related activities emerge and evolve (see Boschma and Lambooy (1999) for an early statement of intent, and Boschma and Frenken (2011) for an overview of recent empirical work).

One central debate in this literature has been the role of cities’ and regions’ current or historic economic structure in influencing future development trajectories. Much of this begins from Jane Jacobs’ observation that urban economic diversity both protects a city against external shocks, and enables knowledge spillovers across sectors as well as within them (hence ‘Jacobs spillovers’). Over time, these knowledge flows allow entrepreneurs, firms and other actors to recombine existing products and services into new ones. However, Jacobs is silent on what mixes of related activities are most likely to generate dynamic spillovers, and what might be most important for regional
resilience. Frenken et al (2007) distinguish Jacobs-style ‘related variety’ within sectors, from pure or ‘unrelated variety’, which they associate with urbanisation economies. In tests on Dutch regions, they find that related variety is most likely to generate subsequent regional employment growth, with unrelated variety acting as a dampener on unemployment. A series of other empirical studies generate similar results, for example Boschma et al (2013), Boschma et al (2012), or Boschma and Iammarino (2009).

Another key debate in the evolutionary literature concerns the nature of cluster development and change. Evolutionary geographers are at one in suggesting that dynamic causation and path-dependency are central to an understanding of the spatial structure of the economy. However, they disagree about the best way to frame this. Linear models of path dependence, as popularised by Paul David (1994) imply that cities and regions are set on particular trajectories by random events, which only external shocks can break. Martin and Sunley (2010) propose a more open-ended approach, which allows for the role of spatial factors in business formation and the ‘support, selection and ... emergence of new trajectories’ [p30]. Applying this to clusters, they argue against dominant ‘life-cycle’ frameworks and in favour of more open ‘adaptive cycle’ frameworks, which see clusters as complex adaptive systems that can evolve along a number of distinct paths (Martin and Sunley 2011).\(^5\)

In contrast to wave 1 cluster theory, these models emphasise tensions between positive and negative feedbacks, and explicitly include consideration of dynamics. They thus provide a richer conceptual and descriptive framework than Porter-type models. For instance, RIS frameworks emphasise the internal benefits of the cluster as a socio-economic system, but also suggest potential disadvantages, such as the ‘lock-in’ that may result from firms’ excessive reliance on local networks and sources of ideas (Boschma 2005). Similarly, EEG theorists suggest a variety of growth and decay trajectories for clusters, with interactions between different parts of the system generating positive, negative or ambiguous outcomes (Martin and Sunley 2011).

However, because much of the literature still treats ‘places’ as the central actors in these systems or evolutionary metaphors, the theory provides little guidance on what determines how these conflicting forces play out. In EEG, the reliance on metaphor and high-level analogy allegedly acts as an additional barrier to clearly understanding micro mechanisms, causal impacts and feedback loops.

This places substantial limits on what empirical work can tell us. There is a frustrating contrast between the very rich descriptive literature that systems and evolutionary models can generate, and the limited tests of how the component parts influence each other. For instance, empirical papers on the role of related variety – or ‘branching’ – on areas’ economic evolution generally set up panels or cross-sections of regions or industry-region cells, then simply test for correlations between area-level outcomes and past area-level conditions. Relatively few theoretical studies (such as Storper (1997; 2011) or McCann (2011), and empirical studies (Neffke, Henning et al. 2011; D’Este, Guy et al.\(^5\) Note: for reasons unknown, Martin and Sunley are major critics of Porterian clusters thinking (wave 1), but major figures in the evolutionary clusters literature (wave 2).
explicitly model micro-economic channels (such as patterns of firm entry and exit, or university-industry collaborations) rather than observing area-level outcomes. As we shall see in the next section, such limitations also mean that these approaches are hard to translate into workable policies.

This stands in stark contrast to recent research in urban economics which emphasises the importance of understanding the micro-economic actors and mechanisms that may underpin similar aggregate behaviour (e.g. the tendency for productivity to increase with city size). We discussed in section II, how these more conventional economic approaches are able to incorporate a number of ‘systems’ and ‘evolutionary’ concepts without losing this micro detail. In the small amount of evolutionary literature that does emphasise firms, workers and households as the underlying actors, the strong dynamic complexity comes at the cost of much simpler models of individual behaviour. This then makes it hard to understand the impact of policy precisely because this depends on the behavioural response of individual firms, workers, etc.

A second, related issue is that by concentrating on places – or place-based systems – rather than firms and individuals, the conceptual and empirical analysis in Wave 2 cluster thinking tends to focus on physically bounded, immobile objects rather than the (relatively) mobile actors who comprise a given cluster. Urban economic models suggest that while there are strong forces that create and sustain clustering, there are also dispersal forces, and that concentrations of firms may relocate across space over time. There is a great deal of empirical evidence for this, notably for creative industries clusters in global cities such as London and New York, where shifting property market pressures help relocate existing hotspots and provide founding conditions for new ones (Hutton 2008; Pratt and Jeffcut 2009). Because they focus on ‘the cluster’ rather than its actors, however, evolutionary and systems models are more or less silent on these crucial dynamics.

VI. CLUSTER POLICIES: SMART SPECIALISATION

Smart Specialisation provides a good example of the impact this second wave of cluster thinking is having on major policy initiatives. The development of the EU’s Smart Specialisation policy is described in detail elsewhere in this issue [REF]. A sectoral development framework, in and of itself, does not necessarily require a cluster component. However, smart specialisation is now being extended into the EU’s regional growth programmes, and is a central theme in the EU’s reformed Cohesion Policy: from 2014 it will be compulsory for any region accepting Structural Funds to have in place a regional smart specialisation strategy (European Commission 2011).

Setting aside the detail, the area-based version of smart specialisation strategy will require ‘integrated, place-based economic transformation agendas’ that involve ‘each region building on its own strengths’ (Foray, Mowery et al. 2012) [p9]. As we discuss below, cluster policies are proposed as a central component of practical strategy development.

As McCann and Ortega-Arguielés (2011) explain, this version of the policy is built on Schumpeterian frameworks, regional innovation systems and evolutionary economic geography. Regions develop through an ‘entrepreneurial process of discovery’. Importantly, these entrepreneurial discovery
processes work differently in different places. So regions need to identify their own areas of competitive advantage. ‘Context matters’ for the evolution of the regional system – especially the position of entrepreneurs, and the performance of public-private, university-industry linkages. Regional leaders need to exploit agglomeration economies, and to deal with diseconomies of agglomeration in overheating regions. At the same time, leaders also need to facilitate branching of the industrial structure into related sectors and technologies.

Discovery processes are inherently uncertain, and this has important implications for policy delivery and management. Local public and private sector actors have a strong shared interest in collaboration, but public leaders need to avoid capture by vested interests (Jaffe 1996; Rodrik 2004). Following Rodrik, smart specialisation design guides emphasise the importance of clear processes, robust evaluation and systems that allow policymakers to drop failing interventions. In light of the arguments in this paper it is surprising, therefore, to find that clusters feature as an important part of the ‘delivery toolkit’. Foray et al (2012) suggest that ‘due to their inherent capability to support co-operation … clusters are powerful instruments to foster industrial competitiveness, innovation and regional growth’ [p67]. Specifically, they argue that clusters can be used both in designing and implementing a regional smart specialisation approach – as a way of identifying and benchmarking regional strengths, and as a platform for bringing public and private actors together.

In its area-based form, smart specialisation is a policy concept that is unusually well-grounded in the literature – particularly when compared to interventions such as Tech City. However, as even its proponents acknowledge, the choice of building blocks makes it difficult to develop actionable policies. McCann and Ortega-Arguilés (ibid) identify three ‘challenges’ in using the ideas to develop concrete interventions. First, systems and evolutionary models provide very rich frameworks, but tend to collapse to individual descriptive case studies. Pulling out generalizable findings is not straightforward. Second, regions are not closed systems – but the sense of regional bounding inherent in the frameworks generates a double risk of under-bounded interventions and inappropriate area-level initiatives. Finally, and most alarming for policymakers, we don’t know what works, or as McCann and Ortega-Arguilés put it, there is little empirical research ‘regarding the relationships between the policy objectives and policy instruments’ [p4].

All of this makes the endorsement of cluster policies all the more puzzling – since this is one set of interventions where there is fairly clear evidence of ineffectiveness. For example, the smart specialisation design guide (Foray, Mowery et al. 2012) urges policymakers to use cluster mapping tools from the European Cluster Observatory6 – even though these only measure co-location, rather than links between firms, and thus do not identify clusters even in the basic Porter sense. Revealingly, the authors of the guide also suggests that regional leaders avoid setting up new cluster initiatives, and streamline existing ones.

Overall, we suggest that wave 2 cluster thinking is not much better than wave 1 as a means of designing industrial / economic development policy. Even when strategy is better-grounded in theory, as is the case here, it feels frustratingly hard to operationalise as a basis for policymaking. The coming years will show how our predictions play out.

6 http://www.clusterobservatory.eu
VII. CONCLUSIONS: CLUSTER POLICY AND AGGLOMERATION POLICY

Economic activity is spatially concentrated. In this sense, much observed clustering is real, which is what makes clusters useful as a descriptive device. But the problems with cluster thinking outlined above, mean that the cluster concept is far less useful as an analytical tool or as a means for making policy. As Duranton (2011) explains, a physical cluster is the outcome of what entrepreneurs, firms and workers do. As we have seen, because the cluster is an emergent property of these interactions it is very difficult to make policy that targets cluster outcomes and manipulate the cluster itself.

What would a more effective approach look like? We think it should consist of two elements. First, some national industrial policies may be more effective if they are more place-sensitive in design and delivery. Second, policymakers should focus more on horizontal policies that aim to make cities work better, in ways that help firms to grow. That is, we need what we might call ‘agglomeration policy’, not cluster policy. We discuss these points in order.

If we cannot manipulate cluster outcomes directly, a better basis for policy is to focus on market failures that affect individual firms and people within the cluster. Such an approach might draw on general evidence about market failures (e.g. in terms of access to finance) but may also benefit from a more in-depth understanding of individual actors inside a cluster. This information and analysis could then be used to develop a suite of policies, which may include encouraging entrepreneurship, subsidising venture or other early stage finance, building workforce skills and management capacity, and helping firms forge international links (Bresnahan and Gambardella 2004). For very young firms, the popularity of co-working spaces and accelerators suggests a further set of potentially useful interventions (see Moretti (2012) for more on this).

Several comments are in order. First, although these policies may benefit from more detailed local understanding, they require much less information than many of the complex cluster policy proposals. Even then in many of these areas effective interventions are likely to prove hard to design and implement. Robust evidence on the effectiveness of specific interventions is still much thinner than we would like, and further work is needed to clearly identify causal effects. All of this suggests the need for experimentation (which may be aided by decentralisation) and rigorous evaluation (which may not be). See Overman (Forthcoming) for further discussion.

Second, it is important to recognise that, as always, in situations with multiple interacting market failures fixing one of these failures need not necessarily improve efficiency. Also, in stark contrast to cluster thinking, there is no guarantee that fixing market failures will automatically increase cluster employment or productivity. Indeed, as explained above, it is possible that un-priced externalities may mean some clusters are too large rather than too small.

Third, in the UK at least, these policy levers are still largely held at national rather than local level. A more decentralised approach to any cluster orientated components of industrial policy thus leaves local decision-makers with rather less to do – advocacy, co-ordination, planning and public realm (although the City Deals process is beginning to change this, with a number of major cities taking
control of elements of adult skills, and at least one – Manchester – calling for devolved inward investment policy).

Fourth, and most importantly, many of these policy actions need not be specific to the sectors or places that policymakers are most interested in (most SMEs would benefit from actions in all of the areas highlighted by Bresnahan and Gambardella). This has two important implications for the shape of policy: first, its effects may ‘spill-over’ to other sectors; second, these effects need not be physically confined to a specific targeted area.

If policy interventions are, at least in principle, not specific to the sector and potentially have spillovers to other firms (both in other sectors and other areas outside of the cluster), this raises the question of whether any sectoral component is still desirable. In turn, this may depend on the extent to which market failures are particularly prevalent in some sectors. For example, in activities at or close to the technological frontier challenges to business formation and growth may be particularly severe because market structure, information asymmetries and externalities all may lead to under-investment. For example, while the costs of entry are now very low, many firms in the ‘digital economy’ face acute problems in raising finance and hiring skilled staff from international markets, which most firms in (say) retail do not generally face (Nathan, Vandore et al. 2012). If policymakers want to encourage the development of the digital economy and other sectors – because they are high-value added and/or exhibit social returns that are higher than private returns – then there may be good reasons for addressing the acute issues in these sectors, over and above horizontal policies that benefit all sectors. To the extent that the resulting interventions are sectorally focused, even the delivery of space-neutral policies may end up following actual cluster geography. For example, national policies to subsidise venture finance for high-tech firms – such as Enterprise Capital Funds – are likely to result in higher levels of VC activity in the largest clusters of tech firms, such as East London.

This leads us to our second conclusion: outside of some specific sectors horizontal policies – targeting aspects of places rather than sectors as a means of encouraging growth – may more effective than a pure cluster-based approach. At the same time, horizontal industrial policy objectives may be better served by specifically targeting particular places – rather than pursuing a place-blind strategy. The basic argument is that cities deliver productivity benefits so that getting urban policy wrong may significantly weaken aggregate economic performance. An ‘agglomeration policy’ approach recognises this and seeks to develop interventions that increase the benefits of urban location (in terms of productivity) while damping down the disadvantages (in terms of costs). For instance, in many countries the evidence would suggest that local policymakers should consider infrastructure investments and improving skills (which raise the ‘productivity curve’), or more market-based housing policies (which lower the cost curve). In highly centralised economies (such as the UK) national policymakers may also need to decentralise economic development powers and resources to major agglomerations, where reallocating responsibilities to city-regional level makes sense (Cheshire and Magrini 2009). In the UK the City Deals process, and the Coalition’s localism agenda, may represent the beginnings of such a shift (Nathan and Steeples Forthcoming).

One recent example of a locally-led, agglomeration-focused approach to urban economic strategy is the Manchester Independent Economic Review (Manchester Independent Economic Review 2009).
Drawing on research by LSE and others, the Review’s Commissioners found that labour productivity in the city was lower than would be expected given the size of the city-regional economy. A large part of the productivity ‘gap’ was due to low skills and poor skill utilisation; transport networks within the city were also inadequate. On the cost curve side, evidence suggested undersupply of housing in the most popular neighbourhoods. The Review made 10 recommendations, including improving early years and school education; productivity-focused transport policy and better intra-city connectivity; a shift towards demand-led housing strategy, and a unified planning, regeneration and neighbourhood renewal framework. Traditional cluster and FDI policies were not recommended, although the city-region has since called for devolved inward investment powers.

To conclude, this paper makes four basic points. First, as governments in the UK and elsewhere rediscover industrial strategy, policy designers need to pay careful attention to the spatial economy and to the spatial footprint of industrial policy interventions. Second, the cluster ‘cookbook’ represents a superficially attractive set of interventions, but as we have argued, makes for generally ineffective interventions. Third, better-targeted horizontal interventions and a greater focus on boosting agglomeration economies are likely to deliver greater results. Fourth and finally, the evidence base on specific spatial interventions and delivery architecture remains under-developed. As Rodrik (2004) makes clear, industrial policy is inherently experimental, and this is as true for specific places as for nations. Policy experimentation is to be encouraged, and in the UK at least, devolution to city-regions could help bring this about.
REFERENCES


Nathan, M., E. Vandore, et al. (Forthcoming). 'The first rule of Tech City is, you don’t talk about Tech City ...': the politics of place-naming in digital East London.


