[The Visible Hand/Priming the Pump:] Sub-National Economic Complexity and Applications to UK Industrial Policy

Joshua Bailey*

Yale University

Working draft: April 10, 2024

Abstract

The UK's has some of the most significant spatial economic disparities found anywhere in the developed world. Against this backdrop, and the increasingly challenging global geopolitical context, states, including the UK, are looking to a more active industrial and growth policy to address these longstanding issues. Despite the recent increase in interest, policy has mostly focused on broadly targetted subsidies and tax-breaks, despite the very different structure of local economies. This paper applies economic complexity methods to develop a detailed picture of the UK's sub-national economic structure, utilising novel sub-national data on employment and exports to develop a multidimensional measure of sub-national economic complexity. I then show that economic complexity is highly predictive of key sub-national economic outcomes, even after controlling for traditional capital and labour inputs. Applying the economic complexity estimates, I finally develop a framework to quantitatively inform a sectorally-led industrial policy. This framework generates specific sectoral recommendations, based on a network of capability 'relatedness', that policymakers can apply when developing industrial policy interventions.

JEL classification: L52, R1, O25, O40, O52

Acknowledgements:

*Yale University, Jackson School of Global Affairs, `e-mail:` joshua.bailey@yale.edu

1 Introduction

To follow

2 The UK's Spatial Economic Divide and Economic Complexity: Literature Review

2.1 The UK's spatial economic divide

2.1.1 The State of Play: Britain's Significant Economic Disparities

The UK's significant regional economic challenge is shown starkly in Figure 1. It shows GDP per capita across the UK's regions from the turn of the century. The UK began the century with a large divergence in GDP per capita, that has persisted and even grown in the more than 20 years that have followed.

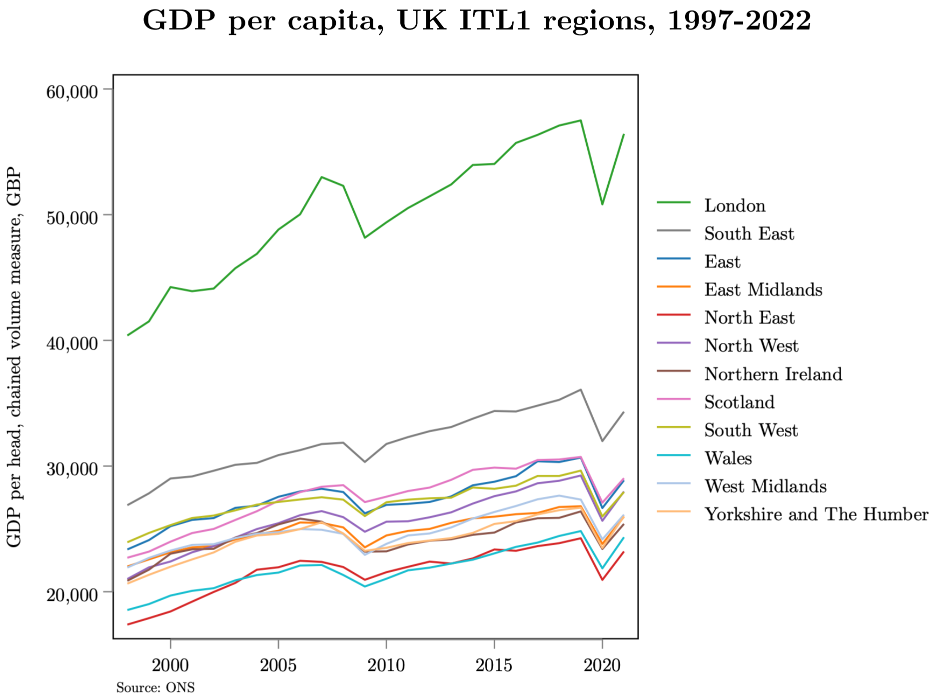


Figure 1: Differences in output across the UK are wide and getting wider.

While many countries have spatial disparities, it is reasonable to expect that fiscal policy, through redistribution, will limit the size of these gaps. But in the UK, the system of redistribution isn't achieving this either. Figure 2 shows Gross Disposable Household Income (GDHI) per head for the UK's regions over the same period as Figure 1. Indexed against the UK average, even after the effect of taxes and welfare, significant income disparities between London and the South East, and the rest of the UK persist. These gaps, especially between London and the rest, have grown substantially since 1997. Even if housing costs are taken into account, which are higher in London and the South East, these gaps

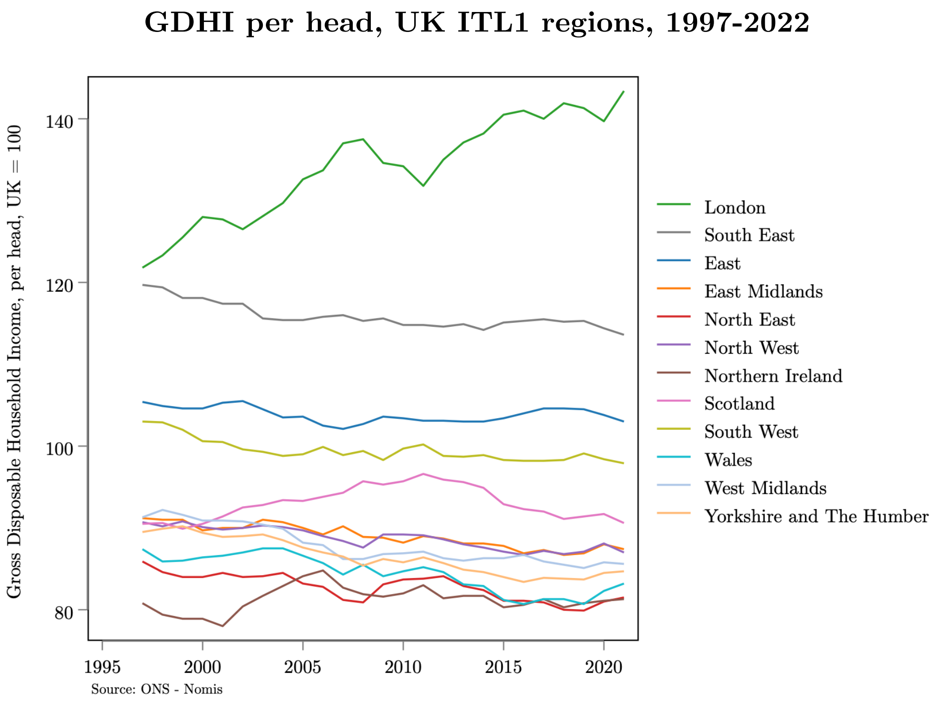


Figure 2: Even after redistribution, regional inequality is getting higher in the UK.

continue to persist. The effect of housing makes the gap between London and the rest somewhat smaller — the difference in median disposable income between London and the North East before and after housing costs in 2019 narrows from 25% to 11% — but the gap between the South East and the rest of the UK is persistently high (HM Government 2022).

Underlying the significant differences in economic outcomes is the productivity divide. Figure 3 draws on European regional data service, ARDECO, data covering regional GVA across the EU.[1] In per worker terms, Figure 3 shows the UK's significant and longstanding divide in regional productivity. The South East, and especially London, have consistently outpaced the rest of the UK. A similar pattern is observed using other measures of productivity (Martin et al. 2021). The data available at the European level goes back to the 1980s, but other studies have shown that the UK's regional productivity disparities have been fairly consistent since at the start of the 20th century with some narrowing in the post-war period, following by an accelerated divergence since the 1980s (Geary and Stark 2018).

The UK's significant spatial productivity disparities are also large by inter-

1. And the UK as a non-EU country since 2021.

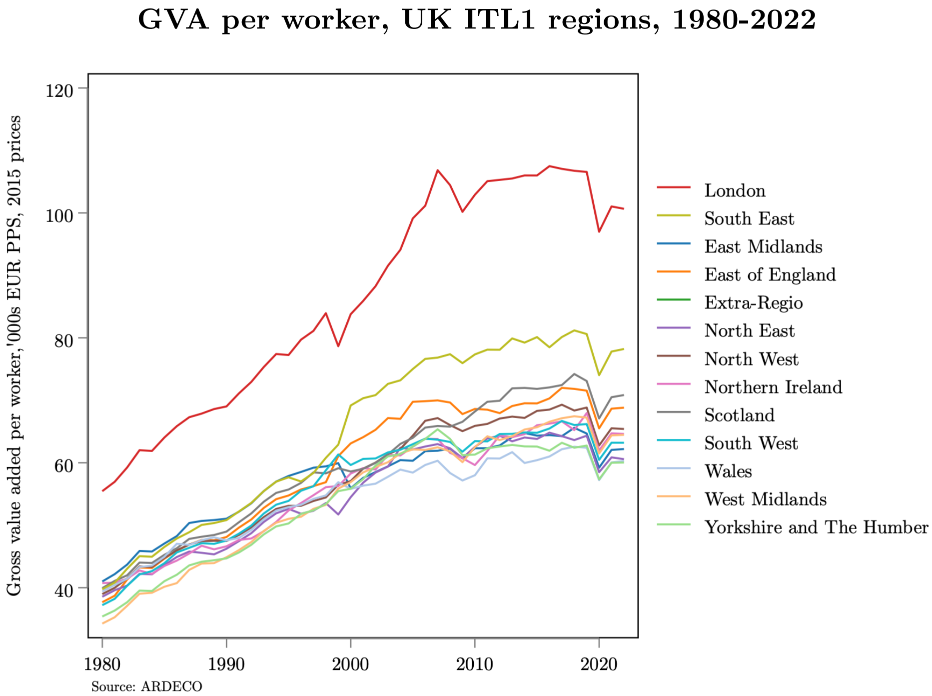


Figure 3: The UK's longstanding spatial economic divide.

national standards. Figure 4 uses the ARDECO data to show the distribution of GVA per worker at the ITL2 level in 2019. Not only is the UK's median regional productivity (shown by the markers) lower than many of its peers, but the range and variation is also the largest in Europe. As the shape of the violin plots shows, the UK is also unusual in the clustering of regions at lower levels of productivity, with one superstar performer in London.

The significant inequality in regional productivity has also widened over time. The change in the regional range of GVA per worker – as measured between the highest and lowest performing region – is shown in Figure 5. The plot shows that not only does the UK have the largest regional range in 2019 (the vertical axis) but that the gap has widened significantly versus its peers since 1980 as well. Only France approaches the UK in terms of the overall worsening of its regional productivity inequality problem.

Productivity is now the central regional divide in Britain. This wasn't always the case. In the early period of deindustrialisation (discussed further below), employment (or rather, unemployment) was the dominant cleavage (Blanchard and Summers 1986; Martin 1988). But while structural regional employment outcomes have generally improved, productivity gaps have widened, as Figure 5

Range of GVA per worker, highest and lowest ITL2 regions, 2019

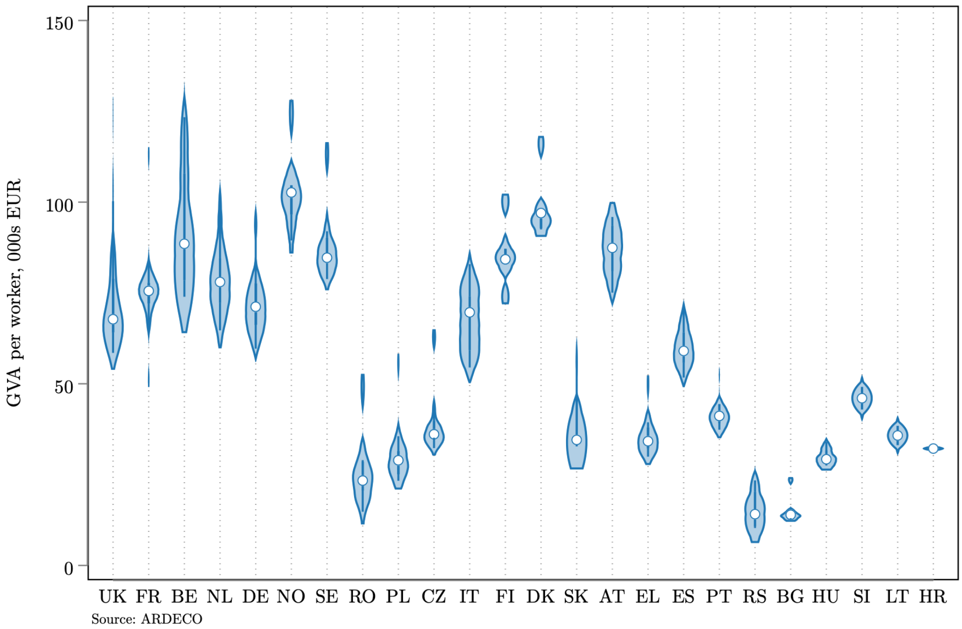


Figure 4: The UK's spatial productivity divide is large by international standards.

Change in range of GVA per worker, highest and lowest ITL2 regions, 1980 - 2019

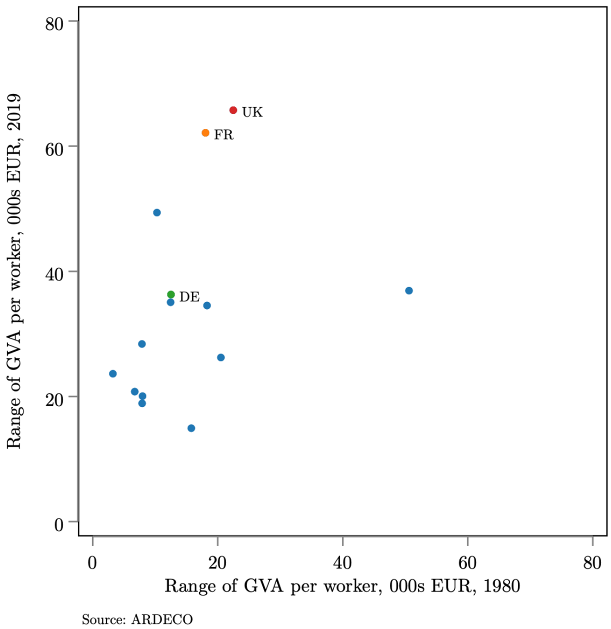


Figure 5: The UK's spatial productivity divide has increased, both absolutely and relative to peer economiues, in recent decades.

shows.

2.1.2 How Did We Get Here? Reviewing Potential Explanations

Explanations for the UK's regional divides cover a range of issues. Several of the most plausible and widely accepted are reviewed in the following section.

Deindustrialisation

The UK suffered a more rapid and severe deindustrialisation than many of its European peers. Figure 6 shows employment shares in manufacturing and mining in 1980 and the change in those shares over the period 1980-2018 for NUTS1 regions in Western European countries. The plot shows that the UK's regions had a particularly high share of manufacturing and mining employment in 1980 compared to its peers (shown by the number of places towards the right end of the horizontal axis). Second, of those places, the UK's regions suffered a more precipitous decline in manufacturing and mining employment shares in the four decades that followed (shown by the places further down the vertical axis). Taken together, regions like Yorkshire and the Humber, and the West and East Midlands were particularly hard hit. The West Midlands, for instance, had a manufacturing and mining employment share of 40 percent in 1980. That fell by 30 percentage points in the decades that followed. The only other European regions that saw a ten-year period of deindustrialisation as fast as the Midlands or Yorkshire experienced in the 1980s were former Communist countries in the years that followed the transition away from state-led economies (Stansbury, Turner, and Balls 2023). This has had permanent effects on overall productivity (Rice and Venables 2021). The resulting weakness in skilled employment has been persistent, with many places never recovering from the impact of deindustrialisation (Beatty and Fothergill 2020). In particular, these regions saw a decline in middle-skilled occupations. Growth of these professions, like retail, care, and hospitality, has lagged the rest of the country (Martin and Becker 2023). The promised jobs of the new economy never arrived for many of the places who felt the brunt of the 1980s.

Agglomeration (or the lack thereof)

Moving to the present, the underperformance of the UK's cities also explains differences in regional outcomes. Beyond the capital, the UK's 'second cities'

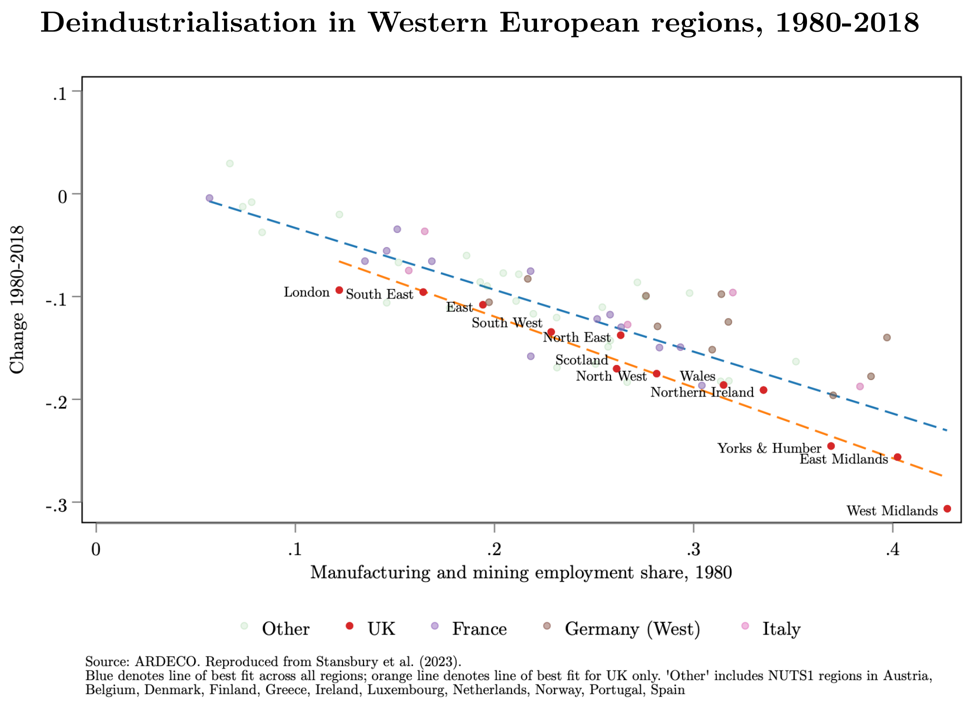


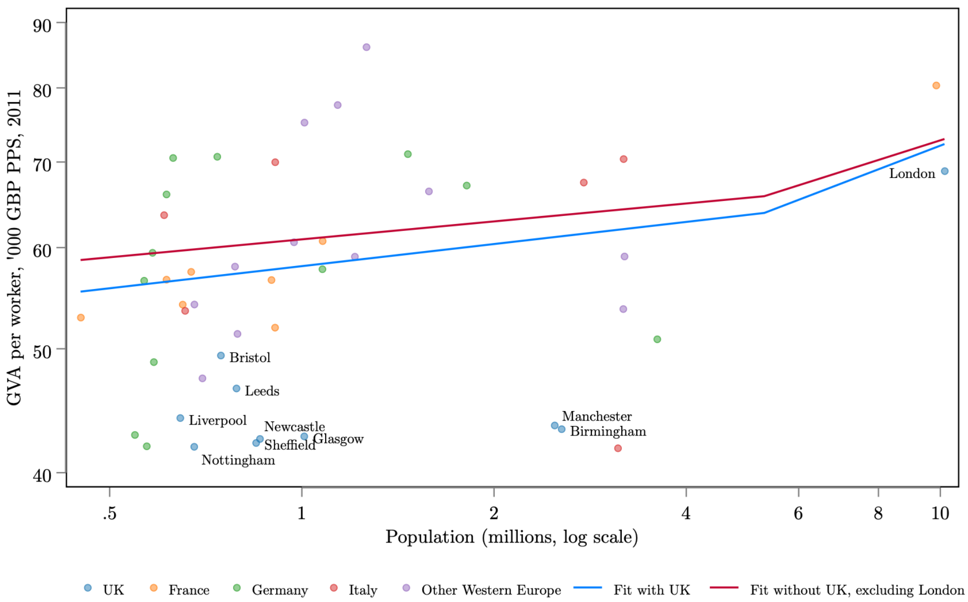
Figure 6: The UK suffered more severe deindustrialisation than its European peers.

tend not to show the same returns to agglomeration[2] as other wealthy countries. Figure 7 shows GVA per worker in cities across the OECD against population. What we can see is that UK cities hug the vertical axis, showing negligible returns to increasing scale. Indeed, the underperformance of the UK's second cities significantly drags down the trend within the whole group, as shown by the divergence between the red and blue fitted trends. This would form at least a partial explanation for the UK's growing regional productivity gap. As cities grow in size, they aren't becoming more productive and therefore fall further behind London and the South East.

This underperformance begs the question of what is driving the breakdown between productivity and city size. One plausible explanation is the state of the UK's transport infrastructure. Research tells us that good transport infrastructure is needed for the agglomeration effect to operate by increasing the effective size of a place. For instance, it allows the matching process between workers and firms to operate efficiently, free from the friction of needing to live near your workplace (Pérez Pérez, Vial Lecaros, and Zárate 2022; Overman and Puga 2008). Once more people are connected, they can share resources more

2. Agglomeration as a general economic phenomenon is discussed further in Section 2.2.1

Productivity and city populations, Western Europe



Source: OECD. Reproduced from Stansbury et al. (2023), originally from Rodrigues and Breach (2021).

Figure 7: The UK's cities don't exhibit the same returns to agglomeration as their European counterparts.

efficiently and avoid duplication of shared assets (Kline and Moretti 2014; Giuliano, Kang, and Yuan 2019). In turn, firms want to locate in places with large populations to be nearer to their customers, effectively increasing their market size (Ellison, Glaeser, and Kerr 2010; Glaeser and Kerr 2009). As places become effectively larger, productivity (Glaeser 2010; Saito and Gopinath 2009) and produce quality tend to improve (Saito and Toshiyuki 2016).

If a place does not have good transport infrastructure, its effective size will be limited. But by international standards, the UK has poor public and private transport infrastructure. One way to estimate this is by the area that transport infrastructure facilitates access to in a place. Figure 8 shows data from (Conwell, Eckert, and Mobarak 2023) estimating the area accessible from the city centre of US and European cities by road and public transport at rush hour. This gives a useful combined sense of the overall quality of transport infrastructure in cities. When looking at the US and Europe, stereotypes are borne out. US cities are generally well-connected by road, but have poorer public transport. European cities have the inverse. The UK is unfortunately clustered in the bottom left of the plot, suffering from poor road and public transport.

Poorer transport infrastructure limits the effective size of UK cities. Ad-

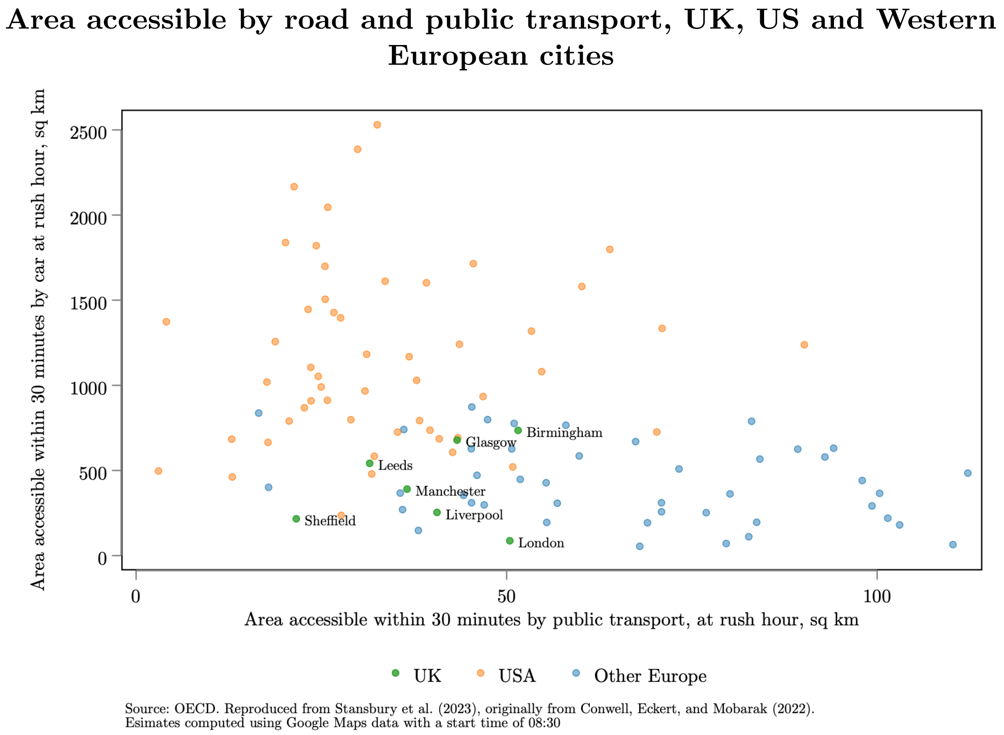


Figure 8: The UK's cities have poorer roads than the US and public transport than the rest of Europe.

justing for the effective size of UK cities, i.e. the number of people who can realistically travel into the centre for work, the expected agglomeration pattern reasserts itself (Breach and Rodrigues 2021). Birmingham, for instance, goes from having an OECD population of around 1.9 million, to 0.9 million, once its effective size is computed. This makes sense. Birmingham has no metro system. Neither does Manchester or Leeds, the two next largest English cities. By comparison, France's second, third and fourth cities have 8 metro lines between them (four in Lyon, two in Marseille and Lille). Manchester and Lyon have similar tram systems, with around 100 stations each but Marseille and Lille have 3 and 2 tram lines respectively, to Birmingham's 1 and Leed's 0 (Forth 2019). The UK's second cities are just not that large because its transport infrastructure fragments what should be much larger cities.

Labour mobility and housing costs

In the face of the productivity gap, theory tells us that workers should move to from lower productivity regions to higher ones (Blanchard and Katz 1992). Despite this, the UK exhibits the opposite pattern, as people tend to leave the capital for other parts of the UK (HM Government 2022). As with the productivity gaps, only France follows a similar pattern among the UK's peers (Stansbury, Turner, and Balls 2023).

The most plausible explanation for this is the UK's prohibitive housing costs. As discussed above, the gap in disposable income after housing costs between London and the rest of the UK is much smaller than the difference before housing costs (HM Government 2022). (Agrawal and Phillips 2020) find that median household income in London is 14% higher than the UK average before housing costs, but only 1% higher after. Workers are therefore being driven away from more productive jobs into less productive ones as wages fail to keep up with London house prices. High-skilled workers in the regions are then trapped in low productivity roles, that cannot improve because the normal forces that drive greater productivity in cities have stalled and cannot be traded for better roles in London because house and rental prices are too high.

2.1.3 Addressing the challenge: Correcting market failures and the return of industrial policy

The literature tells us why it is the case that much of this required investment doesn't happen. Places suffer from coordination failures, where no individual stands to profit from, say, a new housing development on their street when they already own a house or a new metro line when they already drive to work. This is because the benefit of such an investment to an individual depends on the behaviour of others. If only your street allows a new housing development, that won't allow enough new workers into the city to, say, improve child-care services or start exciting new businesses, but it will disrupt your commute and increase competition for places at the local school. The same is true if only your neighbourhood opts for the new metro stop. Places can get stuck in bad equilibria when it would otherwise make everyone better off to invest (Rodrik 2009). Here there is a clear role for policy to correct these failures and overcome local veto-playing. A classic case of this in the UK is the planning system. Since the Town and Country Planning Act 1947, the UK planning system has been designed around a series of local veto-points, allowing small minorities to object to new projects they don't like (Breach 2020). In effect, it institutionalised the coordination problem as a feature of policy.

Even if the UK were to address the infrastructure and housing challenges it faces, it is unlikely that this alone would close regional productivity gaps. Research tells us that 'knowledge' is a crucial component in productivity growth but that absent intervention, firms will tend to underinvest in it. This is because knowledge is in many ways a public good. It is partially non-excludable – in the sense that despite the existence of intellectual property laws, knowledge tends to spill over from people and firms – and non-rivalrous in its use (once you have knowledge, it is by definition not consumable) (Jones and Williams 2000). Another way of thinking about this are the 'cost-discovery' externalities of developing new knowledge (Hausmann and Rodrik 2003). Here, the social value of discovery is not realised in equilibrium because firms and entrepreneurs face an uncertain cost and demand dynamic when investing in new knowledge, while also providing information to potential competitors by investing first.

Relevant to the discussion above, correcting these failures typically requires activity-specific interventions. Take the infrastructure deficit from Section 2.1.2. Once the general need for transport infrastructure investment is agreed upon, governments face a choice as to which projects to support. This inevitably means choosing one sector over another. (Juhász, Lane, and Rodrik 2023) use the example of upgrading a port or the road transport network. Many would call this a 'horizontal' form of industrial policy, not favouring a particular sector, but this is rarely the case. The port will benefit maritime industries and exporters who favour sea freight, whereas the road will benefit local firms who sell their goods to places nearby and construction firms who are good at building roads. The latter is probably of more use to the city's services businesses, enabling more people to get in and out of the city. In this sense, places are "doomed to choose" certain interventions over others (Hausmann, Hwang, and Rodrik 2007). The role of policy, therefore, is to identify which sectors require more support. That is the exercise conducted in Section 4.

There is increasing consensus that in the face of these market failures, governments have not done enough to correct them, especially since the 1980s and the return to a more *laissez-faire* form of economic policy (Juhász, Lane, and Rodrik 2023). There has therefore been a resurgence in state-led industrial policy. Across Western countries and particularly in the US, this was predominantly a reaction to the perceived sense of increasing economic competition with China (Bown 2023). After years of a more hands-off approach to shaping industrial growth, states are now investing trillions in supporting chosen industries and firms. The oft-cited examples are, of course, the US Inflation Recovery Act (IRA), the CHIPS and Science Act (CHIPS) and China's suite of industrial policies, typified by Made in China 2025. The Congressional Budget Office (CBO) and Joint Committee on Taxation estimate that the IRA alone authorises $891 billion in total spending (Committee for a Responsible Federal Budget 2022). The EU has also embarked on a significant round of industrial policy through its Next Generation EU (NGEU) post-pandemic recovery fund. NGEU is capitalised to €750bn. Italy alone stands to receive €191bn in total to be spent within the 2021–2023 period.

The toolkit of industrial policy has also evolved. In (Juhász, Lane, and Rodrik 2023), the authors review the current state of industrial policy, highlighting that taxes and subsidies are now used alongside new tools such as business services, like marketing, management & tech assistance, customized training, infrastructure, and seed capital/loans for specific technologies. More broadly, the sectors industrial policy targets have evolved from traditional manufacturing sectors to a broader range of knowledge services. The government does not necessarily assume it can identify market failures *ex ante* and instead seeks to rely on knowledge dispersed throughout the economy. The state won't necessarily take control or direct a sector, but instead shape its incentives, partnering with it and providing direct support and capacity where needed. (Juhász, Lane, and Rodrik 2023), building on (Rodrik and Sabel 2020), have developed a full typology comparing what they call the New Industrial Policy (NIP) to traditional industrial policy. The typology is recreated and modified below in Table 1.

Quantifying the shift in policy is challenging, not least choosing what to classify as 'industrial policy' over some other form of economic and social investment. One such effort is led by researchers at the IMF and Global Trade Alert called the New Industrial Policy Observatory (NIPO) (Evenett 2024). The NIPO uses granular data from Global Trade Alert to categorise government spending that is "targeted government intervention aimed at developing or supporting specific

Table 1: Comparison of Traditional vs. New Industrial Policy from (Juhász, Lane, and Rodrik 2023)

Aspect	**Traditional Industrial Policy**	**New Industrial Policy**
Market Failures Targeted	R&D, innovation, learning externalities; coordination failures in investment	Traditional plus; Good-job externalities, direction of innovation, missing public inputs
Sectors	Manufacturing, tradable sectors	Services & manufacturing
Firms	Large, globally competitive	All sizes, including SMEs
Government Assumptions	Can identify market failures ex ante; Insulated from capture	Knowledge about market failures is dispersed; Government faces uncertainty; State capacity is endogenous
Incentives	Tax, credit subsidies	Business services; Marketing, management & tech assistance; Customized training, infrastructure, seed capital/loans
Incentive Application	Fixed schedule; Negotiated for large firms	Customized to firm needs
Relationship with Recipients	Pre-specified	Voluntary participation
Conditionality	Hard; rigid criteria	Soft; provisional, open-ended

domestic firms, industries, or economic activities to achieve national economic or noneconomic (e.g. security, environmental) objectives." Being a new database, there is not a significant historical time series but Figure 9 shows the data for 2023, when there were over $2tr of industrial policy interventions, about 2% of global output in 2023. The UK, incidentally, was responsible for the single largest intervention in the database for 2023, the over $600m bailout of Tata Steel's steel plant at Port Talbot in Wales (Sweney 2023).

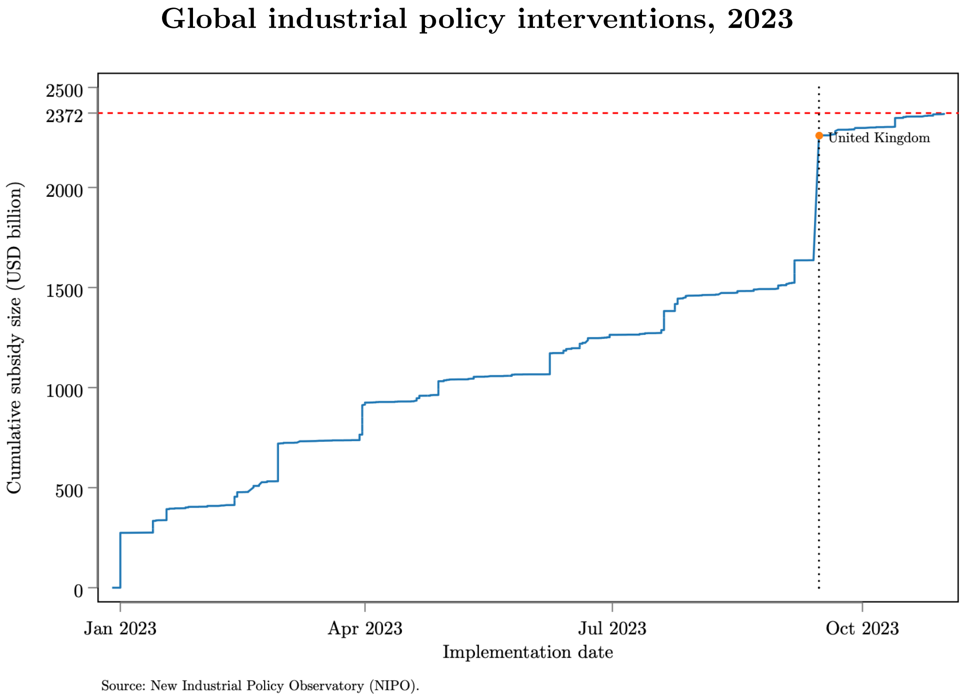


Figure 9: Over $2tr of industrial policy was conducted in 2023.

Despite headline-grabbing interventions like that of Port Talbot and the general resurgence in the popularity of industrial policy, UK industrial policy has historically been a disorganised affair. Policy has tended to be short-term, highly centralised and top-down (Coyle and Muhtar 2021). Most of the period since the 1970s has lacked a coordinated industrial strategy of any sort. There have been periods when this is different, but these episodes have proven to be short-lived. For instance, in 2017, then Prime Minister, Theresa May introduced a formal industrial strategy, (HM Government 2017). Alongside specific sector plans, it established an Industrial Strategy Council to oversee and provide expert advice on the strategy. An empowered department for Business, Energy and Industrial Strategy was charged with leading the strategy. But less than four years later, the strategy and the council were abolished by Prime Minister, Boris Johnson. His Chancellor, Rishi Sunak, replaced it with a more hands-off, Treasury-led document, the Plan for Growth, (HM Government 2021). Despite these challenges, the political focus on a renewed British industrial strategy only increases. The shadow Chancellor and Labour Party politician, Rachel Reeves, has talked up her concept of industrial policy, 'Securonomics' (Reeves 2023). In it, Reeves explicitly borrows from the more interventionist industrial policy pursued by the Biden administration.

It is clear, therefore, that industrial policy is both justified by the theory, can play an important part in closing the UK's regional divides and is here to stay as a political matter. The economic and policy question therefore is what methods can we use to shape a quantitively informed industrial policy? This is the question I turn to next, introducing the ideas of economic complexity and how they can help answer this question.

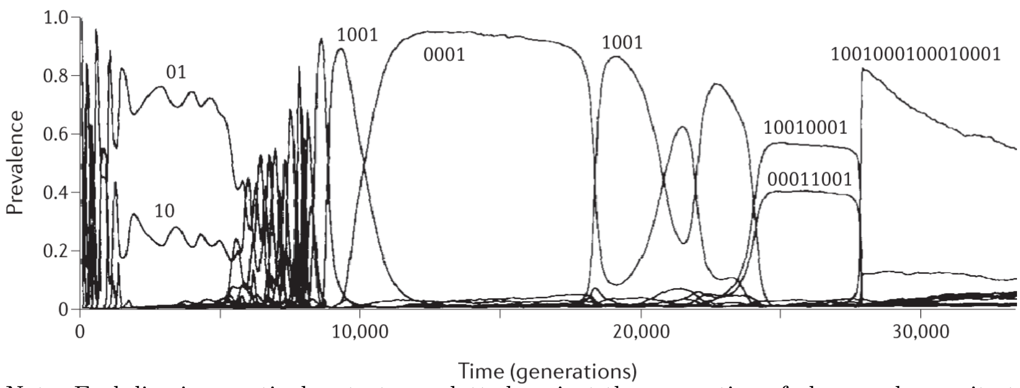
2.2 Economic complexity

2.2.1 The economic complexity approach to understanding economic development

The importance of knowledge in generating higher levels of productivity and the role of an interventionist policy to support its creation begs the question of how to think about 'knowledge' in a systematic way. As I argued in Section 2.1, in the UK's case this also needs to be grounded in a sub-national context, given the shape of Britain's economic geography. Complexity economics offers one such toolkit to support this analysis. In this section, I briefly review the history of complexity economics and then connect it to the economic geography and agglomeration literature referenced in the earlier discussion of the UK.

Complexity economics as a field was originally incubated in the 1980s at the Santa Fe Institute (SFI) (Arthur 2021). One of its early proponents, Kristian Lindgren, applied systems thinking from his training as an engineering physicist, to develop a computerised tournament where randomly chosen pairs of players would compete in the prisoner's dilemma game. Instead of a fixed strategy, which would imply some sort of rational expectations formation, players learn from their interactions with other agents. Agents had a finite memory and evolved their strategies over time (Lindgren 1997). Lindgren found that the system very rarely found a stable equilibrium. Figure 10 shows the evolution of player strategies over time, with simple strategies dominating at the start of the game but more sophisticated and varied approaches emerging as the game continues. The game exhibits significant fluctuations, the outcome varied each time and was extremely sensitive to initial conditions.

Evolution of strategies in a simulated tournament of the prisoner's dilemma.



Note: Each line is a particular strategy, plotted against the proportion of players who use it at a begin stage of the game. The labels indicate the memory depth of strategies i.e. how many previous moves in the game they take into account

Figure 10: In a complex system, strategies evolve in response to changing conditions.

The model underscores what is different about economic complexity. Complexity theorists content that to understand a complex system, one must relax some of the classical assumptions, including the construction of agents, how they are meant to behave, the existence of equilibrium conditions, modes of interaction within a network and exogenous structural parameters (Arthur 2021). Herbert Simon offered a helpful summation of this way of thinking when he defined a complex system (Simon 1962):

> "Roughly by a complex system I mean one made up of a large number of parts that interact in a nonsimple way. In such systems, the whole is more than the sum of the parts, not in an ultimate metaphysical sense, but in the important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole. In the face of complexity, an in-principle reductionist may be at the same time a pragmatic holist."

In the field of empirical economics, economic complexity emphasises granular data that builds up to a picture of the aggregate (Hidalgo 2021). Importantly, as another early complexity theorist and mathematical economist J. Barkley Rosser Jr. put it, researchers should "only [assume] local relationships among individual actors, and [allow] aggregate behaviors or structures [to] emerge out of self-organization rather than simply being imposed or assumed" (Rosser 1999). Since Rosser's remarks in the 1990s, our computational capacity has vastly improved, and with it our ability to apply economic complexity approaches to a range of empirical questions. Economic complexity methods have now been applied in most sub-fields, including macroeconomics (Hommes 2021; LeBaron and Tesfatsion 2008), finance (Battiston et al. 2016) and labour economics (Axtell, Guerrero, and López 2019).

In growth and development economics, complexity approaches have been applied by researchers to "use of network science and machine learning techniques to explain, predict, and advice changes in economic structures. The focus on economic structure is motivated by work showing that these structures explain and predict important macroeconomic outcomes, from economic growth to the intensity of greenhouse gas emissions and income inequality" (Hidalgo 2023). Here, the central, and in many ways original, question of economics is why some nations are richer than others (Smith 1776). Adam Smith's original insight was that the division of labour drove development. As labour specialised, it became more productive, allowing the product of the system to be larger than if the same workers tried to accomplish all the tasks of production. Given the existence of differences in GDP per capita across countries, it must therefore be the case that there are differences in capabilities, like infrastructure, natural resources, institutions, and specific labour skills, needed to produce certain goods and services in a particular place.

Traditional approaches have often emphasised the proportion of different factors that contribute to production (Ohlin and Heckscher 1991; Solow 1956, 1957). In this account, countries become richer as they specialise in capital-intensive goods requiring a broader range of skills and knowledge. Theory has further developed the importance of underlying differences in technology as a way to explain these differences in capabilities (Romer 1994).

Within these models, there is an implicit notion of a hierarchy of capital, knowledge and technology. Surely, this is correct at one level. Everything required to fabricate advanced semiconductors – from specialist process engineers, to extreme ultraviolet lithography machines – seems space-age in its sophistication compared to, say, farming a smallholding (even if one sets aside that most semiconductor engineers are probably not very good farmers). But at another level, this linear notion of more or less sophisticated becomes complicated. We may characterise many activities as being of similar complexity but nevertheless different in some other dimension, such that not all sophisticated economies are capable of doing all equally sophisticated activities. Many would say that transporting oil through pipelines is of comparable difficulty to transporting chemicals safely, and as yet not all advanced economies do both. We often try to abstract away from these complications, collapsing the economic sophistication of a country to a single productivity statistic.

Drawing on the complexity tradition of trying to observe a system as it operates with a view to understanding its emergent properties, economic complexity reintroduces this multidimensional notion of sophistication by mapping the activities of an economy as a network (Hidalgo et al. 2007; Hidalgo and Hausmann 2009). Specifically, it focuses on what researchers have termed the "capability approach" to understand economic development (Hausmann and Hidalgo 2014). Economic complexity begins with the notion that economic units will tend to evolve existing economic activities into related activities, that utilise and build on their existing capabilities. Instead of trying to impose a structure on how the various capabilities, like labour, land, capital, knowledge, are combined to produce certain activities, economic complexity takes an agnostic approach, looking at the outcomes of production, with activities that tend to require similar capabilities, clustering with other activities that require similar capabilities. This gives rise to a key claim made by economic complexity researchers; that while it may be challenging to directly observe the capabilities need to engage in more sophisticated forms of economic production, these capabilities can be inferred from their relative scarcity in an economy (Hidalgo and Hausmann 2009). If an economy can produce high voltage electrical transmission equipment, it must possess the myriad capabilities required to do so. If fewer places engage in a certain economic activity, it is likely to be a more sophisticated activity. Through this, it allows us to infer the presence of certain, connected capabilities, needed to produce complex economic outcomes.

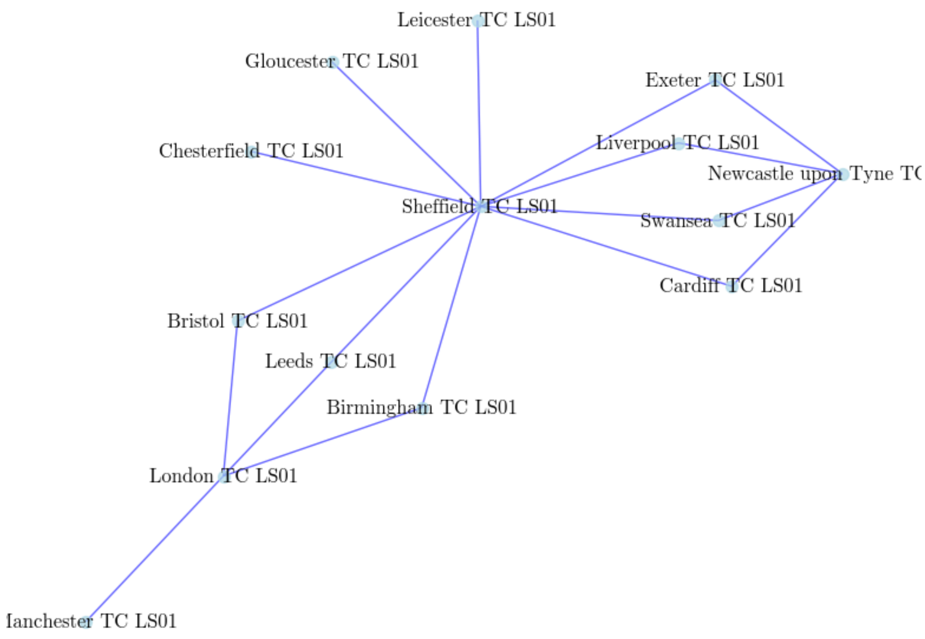
To understand the innovation of economic complexity further, let us consider the standard approach taken in trade analysis and compare how economic complexity builds on it to give a more sophisticated picture of the production structure. Economic complexity measures begins with Balassa's original notion of revealed comparative advantage (RCA) (Balassa 1965) (detailed formally in Section 3.2). A related idea is that so-called Krugman Index (Krugman 1992). It is a regarded as a standard index to measure relative specialisation. The index compares the industrial structures of two geographic areas. It is calculated as follows:

$$\mathrm{K_r} = \sum_{\mathrm{F1}}^{\mathrm{k}} \left| \frac{\mathrm{E_{ir}}}{\mathrm{E_r}} - \frac{E_{\mathrm{i}}}{E} \right| \tag{1}$$

Where $\mathrm{Ei}_{,r}$ is the share of industry i in area r.[3] The Krugman Index compares the industrial structures of two geographical areas. If $K_r = 0$, area A has the same employment structure across industries as the reference area. If $K_r = 2$, the employment structure is in entirely different industries to each other. Figure 11 shows the Krugman Index for selected UK urban areas in 2016 (London, Newcastle, Sheffield, and Oxford). Plotted as a network and using a threshold of $K_r = 1$, it gives an overarching sense of industrial similarity. Sheffield's sectoral composition is similar to several places, whereas Manchester's is only similar to London's, and Oxford is not shown because it has a comparatively unique sectoral mix.

3. The reference area can vary. It was one other country in (Krugman 1992) but can be all other countries (Palan 2010). The same goes for other geographic units, like cities or regions.

Krugman Index, selected UK urban areas, 2016



Note: Plot uses a Krugman Index threshold of 1. Only London, Newcastle, Sheffield, and Oxford have been shown.

Figure 11: The Krugman Index gives a useful overall view of industrial similarity.

The Krugman Index facilitates the easy comparison of two places' sectoral structure, but is less useful in trying to understand the next layer down, *how* places differ. (Hidalgo et al. 2007) begins with this challenge, asking (rhetorically) whether we should care if, in David Ricardo's example, it mattered that Britain specialised in cloth and Portugal in wine? To address their question, they construct a bipartite network (which they call the 'product space') that relates industries to one another. Through the product space, they can show that countries transition into products that are 'proximate' to one another, defined by the industrial mix places tend to have when they develop a particular new industry.

The emergent properties of these networks are extremely powerful. In the original analysis conducted in (Hidalgo and Hausmann 2009), the product-space allow the authors to differentiate the income levels of countries through the information about the ubiquity of the products they export and the diversification of the corresponding countries related to them in industrial composition. This allows them to tell Singapore and Pakistan apart purely through their export patterns. Singapore is connected to diversified countries that export similarly rare products, whereas Pakistan is connected to predominantly poorly diversified countries that export common products. Importantly, these networks take account of scarcity of production. Pakistan exports numerous products and is therefore well diversified, but the scarcity weighting tells us that this is likely due to factors that are not indicative of higher levels of economic complexity, and are more likely to be the result of other factors, like having a large population.

2.2.2 Application of economic complexity methods to sub-national units

While the use of economic complexity measures at the national level that utilise global trade data have become an increasingly popular way to conceptualise national economic performance and the underlying capabilities that support these outcomes, their application at the sub-national level is a relatively newer phenomena, despite this appearing to be a natural extension of existing methods. A body of literature has developed, including applications to China (Gao and Zhou 2018), Australia (Reynolds et al. 2018) and the United States (Fritz and Manduca 2021). (Gao and Zhou 2018) calculated sub-national economic complexity measures for Chinese provinces using data on publicly traded firms. (Chávez, Mosqueda, and Gómez-Zaldívar 2017) was one of the earlier papers to use sub-national employment data across broad sectors to compute complexity measures for Mexico's states. Relevant to this paper, (Fritz and Manduca 2021) constructed economic complexity measures using employment data at the US Metropolitan Statistical Area (MSA). The paper finds that employment-based measures of complexity perform similarly to their national export-based measures. This research has raised important questions about the applicability of economic complexity analysis at the sub-national level (addressed further in Section 3).

Sub-national economic units are worthy of study with regard to economic development. There is a longstanding interest in trying to understand the structure of production geographically, variously referred to as the 'industrial economic base' (Andrews 1953) (Heilburn 1981), 'growth poles' (Perroux 1955) and, latterly, 'clusters' of industries (M. E. Porter 1998). As discussed above in Section 2.1, the UK's extreme spatial disparities support further investigation at the sub-national level.

The now well-developed literature on economic geography also emphasises the importance of agglomeration as a driving force in understanding how places grow (Glaeser et al. 1992; Glaeser 2010). Industries develop at the level of functional economic units, as people, capital, and knowledge come together to form production networks. Indeed, the apparent breakdown between increasing returns to urban scale and productivity gains is an important explanation for the UK's current spatial disparities (Nathan and Overman 2013) (as discussed in Section 2.1.2).

The rising importance of agglomeration within the study of economic geography is an important justification for the use of economic complexity methods. The agglomeration literature emphasises understanding economic production as a network process, where different factors of production interact in heterogeneous ways to form production processes (Glaeser 2010). It also adopts a similar frame to the capabilities approach through the explanatory power given to cheaper transport, knowledge spillovers, common labour pools and specialised intermediate inputs (P. Krugman 1991; Overman and Puga 2008; Ellison, Glaeser, and Kerr 2010). (Balland et al. 2020) investigated the link between agglomeration and complexity explicitly. They found that complexity explains a significant proportion of the variance in the urban concentration of jobs, industries, scientific fields and technologies.

Having grounded the discussion of the UK's economic geography in a set of tools that can be used to understand it and develop policy recommendations to help improve it, I now move to more formally define the computation of economic complexity statistics and the analytical framework pursued in the remainder of the paper.

3 Methods

As the previous section set out, economic complexity is a toolkit that can be applied to places with a view to better understanding their relative economics strengths. The interest of this paper is the sub-national economies of the UK. This section therefore begins where the previous section left off, discussing the applicability of economic complexity approaches to sub-national economic units (Section 3.1. I then move to formally state the economic complexity methods employed in the rest of the paper (Section 3.2), and the data used to compute them (Section 3.3).

3.1 Applying economic complexity at the sub-national level

Traditionally, economic complexity measures are computed with national export data (Hidalgo and Hausmann 2009; Hausmann and Hidalgo 2014). The use of sub-national employment data has emerged as a useful alternative when looking at the sub-national level. It is useful to first examine how employment and export data compare. One critique that might be made of using sub-national employment data to compute complexity measures is that it captures too many domestic distortions to be useful. Whereas national export data offers a macroscopic view of a country's competitive advantage on a global scale, sub-national employment data might reflect a disparate array of local economic activities that do not necessarily equate to global competitiveness. Such a critique could further contend that the inclusion of broader sectoral activities in employment data – going beyond export-oriented industries – may dilute the perceived economic complexity, as these sectors do not directly contribute to the global market positioning of a nation.

One response to this is that global trade data is also a product of myriad distortions. Trade barriers, both formal and informal, distort where export specialisation occurs. More constructively, the integration of sub-national employment data offers a more detailed and contextually rich exploration of economic diversification and capacity. A second response is that employment data captures all industries, including, crucially, services. Typically, economic complexity measures are computed using goods export data (*Observatory of Economic*

Complexity (OEC) 2024). Service complexity is inferred from goods complexity. Using employment data removes the need for this assumption. Moreover, abstracting away from this distinction makes the analysis more powerful, not least because the services/goods distinction is often an arbitrary distinction in the modern economy. Take Rolls-Royce for example. They produce the Rolls-Royce Trent XWB, a market-leading turbofan jet engine used to power the Airbus A350. Rolls-Royce manufactures the WXB at their plant in Derby, England. They export the finished engines to Airbus to be bolted onto new Airbus A350s at their assembly facility in Toulouse. That shows up in the goods section of the UK's current account and would contribute to an export-based economic complexity calculation. But Rolls-Royce now makes more money on after market services, like maintenance and in-flight monitoring, than it does on the original equipment (Rolls-Royce 2024). This would be lost in an analysis that focused on goods exports alone. This is particularly problematic when an increasing share of industrial in advanced economies, including and especially the UK, comes from high value-added services (De Lyon et al. 2022).

This paper also supplements employment-based measures of economic complexity, by developing a sub-national export-based measure, utilising new sub-national trade data available in the UK (discussed further in Section 3.3). Importantly, this sub-national export data covers both goods and services exports.

3.2 Theoretical framework

3.2.1 Economic complexity measures

The economic complexity literature has developed a flexible framework that allows complexity and its related metrics to be estimates at different geographies and using different data as inputs. The general idea, as discussed in Section 2, is to capture what activities (e.g. an industry, product or technology) exist in different geographies (e.g. a country or region). This gives rise to a useful two-way relationship: places are complex because complex activities occur there, and actives are complex because they occur in complex places.

Formally, and following (Hidalgo and Hausmann 2009), I express this in the following way. Let the complexity, K, of a place, c, be K_c, and the complexity,

K, of an activity, p, be K_p. Further, let M_{cp} be a matrix summarising the activities, p, present in location, c. M_{cp} is defined as $M_{cp} = 1$ when a place's presence in an activity is larger than what is expected for a place of the same size and an activity with the same total output.

Let the matrix, M, known as the presence matrix, by defined with reference to a location's Revealed Comparative Advantage (RCA) or, as is often the case when computing sub-national complexity measures, a Location Quotient (LQ). The presence matrix, M, is defined as:

$$M_{cp} = \begin{cases} 1 & \text{if } R_{cp} \geq 1 \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where R_{cp} is the measure of RCA, or LQ, as in the sub-national case. It is given as:

$$R_{cp} = (X_{cp}X)/(X_c X_p), \tag{3}$$

where X is the measurement statistic of interest – in our case employment and exports – and is summed across places and industries, such that:

$$\begin{aligned} X_c &= \sum_p X_{cp} \\ X_p &= \sum_c X_{cp} \\ X &= \sum_{cp} X_{cp}, \end{aligned} \tag{4}$$

where X_c is therefore the total number of people employed or the value of exports in a particular place and X_p is the total number of people employed or the value of exports in a particular sector.

Utilising M_{cp}, we define the diversity, D, and ubiquity, U, vectors. Diversity, D_c, represents the number of industries in which a place, c, holds a comparative advantage, given as:

$$D_c = \sum_p M_{cp} \tag{5}$$

Ubiquity, U_p, quantifies the number of places in which a sector, p, is prevalent, given as:

$$U_p = \sum_c M_{cp}{}^{4} \tag{6}$$

Taking D and U, I construct diagonal matrices for each, denoted as D^{-1} and U^{-1}, inverting their values to create the modified matrix $\widetilde{M}$:

$$\widetilde{M} = D^{-1} M U^{-1} M' \tag{7}$$

$\widetilde{M}$ captures the shared industrial strengths among places, adjusted for each industry's ubiquity. An equivalent way to express this matrix, suggested in (Mealy and Coyle 2022), is as the matrix $\widetilde{M} = D^{-1}S$ where $S = MU^{-1}M'$ is a symmetric matrix, with elements S_{ij} reflecting the common competitive industries between places i and j, inversely weighted by industry ubiquity. This weighting scheme elevates the significance of unique industrial strengths. To ensure $\widetilde{M}$ is row-stochastic, I normalize it, which also ensures it has the property of a stochastic matrix, where each row distributes probability across places based on their shared industrial strengths.

The ECI vector is therefore given by the eigenvector corresponding to the second-largest right eigenvalue of $\widetilde{M}$.[5] By symmetry, PCl is given by transposing the M matrix and finding the eigenvector associated with the second-largest right eigenvalue of an $\widehat{M}$ matrix, given by:

$$\widehat{M} = U^{-1} M' D^{-1} M \tag{8}$$

Both ECI and PCI are normalised, such that:

4. To avoid division by zero in subsequent calculations, a small constant (ϵ) is added to each element.

5. The use of the eigenvector corresponding to the second-largest right eigenvalue to compute the ECI follows from the fact that the matrix $\widetilde{M}$, being normalised as row-stochastic, can be interpreted as representing a Markov chain. In a Markov chain, each element m_{ij} of a stochastic matrix represents the probability of transitioning from state i to state j. The largest eigenvalue of such a matrix is always 1 due to the Perron-Frobenius theorem, which applies to non-negative matrices and guarantees a positive eigenvalue that is at least as large as the absolute value of any other eigenvalue. The corresponding eigenvector, often called the stationary distribution, reflects a state where, if the system were to continue indefinitely, the proportions of time spent in each state would reach equilibrium.

$$
\begin{aligned}
ECI &= \frac{K_c - \tilde{K_c}}{\sigma(K_c)} \\
PCI &= \frac{K_p - \tilde{K_p}}{\sigma(K_p)},
\end{aligned} \tag{9}
$$

where K_c is the eigenvector corresponding to the second-largest right eigenvalue of $\widetilde{M}$, $\tilde{K_c}$ is the average of K_c, $\sigma(K_c)$ is the standard deviation of K_c, K_p is the eigenvector corresponding to the second-largest right eigenvalue of $\widehat{M}$, $\tilde{K_p}$ is the average of K_p and $\sigma(K_p)$ is the standard deviation of K_p.

In conceptual terms, computing ECI can be thought of in the following way. Imagine a network where each node represents a region, and the nodes are connected if they share industrial strengths. Not all industries are equally distributed — some are more common (e.g. agriculture might be widespread), while others are unique and specialised (like aerospace engineering). The goal is to understand how similar or different each region is, based on the unique and common industries they have.

The $\widetilde{M}$ matrix modifies the initial presence matrix, M, using the diversity, D, and ubiquity, U, of certain industries. Connections between regions are *stronger* if they share industries that are rare (less ubiquitous). This also means places are more similar if they share unique industries. Each region's connections are adjusted by how diverse it is. This means if a place is very diverse (has many industries), it won't unfairly influence the similarity scores just because it has more industries. $\widetilde{M}$ can therefore be thought of as a map of similarities between regions, considering the uniqueness of their industries and not just the quantity. The eigenvector associated with the second-largest eigenvalue allows us to collapse the information contained in $\widetilde{M}$ into a one dimensional vector. This eigenvector, the ECI, ranks each region based not just who has more or unique industries, but who has a more complex, interconnected economic structure.

3.2.2 Digression: Defining the presence matrix, M - alternative approaches

The use of RCA / LQ as the function to define the presence matrix, M, is well established in the complexity literature (Hidalgo and Hausmann 2009) and is employed throughout this paper. It does, however, have some weaknesses when applied to understanding the network of production in a place. For one, it tends to under-report the sophistication required to produce common industries alongside other industries. Imagine a scenario involving two distinct regions within a country. Region A specialises exclusively in wheat production, while Region B produces both wheat and automobiles in equal measure. Region B demonstrates a more sophisticated economic structure, as it encompasses the skills required for both agriculture and manufacturing. Nevertheless, the LQ for wheat would be below 1, indicating that wheat constitutes a smaller share of its economy compared to the national average. This occurs because the presence of additional, less common products (like automobiles) dilutes the significance of more widespread goods (such as wheat) within the economy. Thus, the metric inadvertently penalises regions with a broader production spectrum by assigning them lower LQ for commonly produced goods, even though these regions are, in reality, more economically diverse and potentially more sophisticated as a result of being able to operate at this level of diversity.

One simple alternative is to just use the raw LQ matrix. This would in theory preserve all of the information contained in each place's sectoral specialisation. It would also avoid the threshold effect, whereby places with a sectoral LQ just below 1 receive no credit for that sector in the M_{cp} matrix. The problem here is that the degree of specialisation implied by the LQ does not scale linearly. A place with an $LQ = 8$ is not twice as specialised as somewhere with an $LQ = 4$.

Another alternative, proposed in (Fritz and Manduca 2021), is to define the presence matrix as exactly that; whether a sector is present in a place, such that:

$$M_{cp} = Presence_{cp} = \begin{cases} 1 & \text{if } X_{cp} \geq n \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

where n is a given employment threshold.[6] The same principle would apply if using sub-national export data. This avoids the under-reporting of common industries in well-diversified areas. (Fritz and Manduca 2021) cite the example

6. (Fritz and Manduca 2021) propose using $n = 1$.

of gas stations (NAICS code 447110) to make this point. In 2015, there were 13,972 gas station workers in the New York MSA. Because local services scale sublinearly with region size (Youn et al. 2016), this means that despite the high number of employees in the sector – and by implication having the capabilities to operate gas stations alongside everything else New York does – gas stations in New York had an $LQ < 1$. New York would therefore receive a 0 in the MP_{cp} matrix for gas stations, implying it does not possess the capability to run gas stations effectively, despite there being a large sector that does exactly that.

Aside from the problem of deciding the appropriate n cut-off, the main issue with this approach for the purposes of this paper is that it tends to compress differences between somewhat and more complex places. Exactly because the presence matrix, M, includes more common industries, it also in effect dilutes the effect of specialising in complex industries. It also allows places to get credit for having a potentially small number of employees in a given sector, boosting its perceived complexity artificially.

This point reveals why RCA-type measures are more appropriate when trying to identify industrial policy and growth opportunities. The paper is concerned with doing the opposite of trying to work out which places have the capabilities to do more things, like operate gas stations alongside software engineering firms. The question here is what *complex* sectors can industrial policy usefully be directed towards. Therefore, in the trade-off between accurately capturing the full spectrum of activities a place is capable of doing, versus maximising the fidelity with which advanced sectoral opportunities can be identified, I choose the latter. When trying to assess where the capabilities exist to produce advanced products, and potential adjacent products, it matters that there are tens of thousands of employees in software engineering in the Bay Area, who could plausibly shift into adjacent industries, but that only a handful of the same people in rural Missouri. In the presence matrix, M, using the approach from (Fritz and Manduca 2021), both places would get a 1 for software engineering. By comparison, the LQ matrix, by definition, identifies sectors where the place is truly strong, attracting a disproportionate amount of the resources, in this case labour, needed to produce a product.

3.3 Data

3.3.1 Sub-national data used in complexity measures

Sub-national analysis in the UK has long been a challenge due to the availability of good sub-national data covering a range of indicators. Where it has existed, it often lacked timeliness, was limited to a few geographic scales and had reliability issues. Sub-national data is often less available than national data because most national statistical agencies collect information at the national level. Therefore, assigning observations to sub-national units can be challenging. This can be especially the case at lower geographic levels, despite these being of economic relevance. This paper utilises a suite of new sub-national data recently made available in the UK, in addition to existing sub-national data. This is made possible due to the broader public debate on UK spatial disparities, which prompted the UK Government Statistical Service (GSS) and Office for National Statistics (ONS) to recently emark on a subnational data strategy (ONS 2022) to improve the quality of sub-national data in the UK.

I use data from the ONS's Business Register and Employment Survey (BRES). The survey publishes employment estimates at the geographic-industry level. I use the full-time employment series, by the 3 digit Standard Industry Classification (SIC) 2007 level and International Territorial Level (ITL) 2 level (geographic level discussed further below). A full list of the SIC 3 industries for which data is used is included in Appendix A. BRES is a large sample survey, of approximately 85,000 businesses collected annually. The survey covers all businesses in Great Britain registered for VAT and/or PAYE (the payroll tax system in the UK).[7] Northern Ireland data is collected independently by the Department for Finance and Personnel Northern Ireland (DFPNI). Importantly, the data is allocated according to workplace location, rather than business registration address, so it accurately maps where jobs are in a given sector and region. It is available for the years 2015-2022, all of which are used in this paper.[8] BRES is favoured over the Annual Population Survey (APS) and Labour Force Survey

7. This includes public sectors jobs but the self-employed, armed forces and government-supported trainees are excluded.

8. Data for 2023 is not available until October 2024. BRES has a historical series that is available from 2008-2014, but the pre-215 series does not include businesses registered for PAYE but not for VAT (the UK's sales tax). The ONS therefore do not treat the post-2015 series as continuous with pre-2015 BRES datasets.

(LFS) because APS/LFS only produce data at the sub-national level by occupation, using the Standard Occupational Classification (SOC) 2010, and does not have a SIC industry breakdown. BRES is also the ONS recommended source of data on employment, by geography and industry.

For export data, I exploit new data on sub-national trade published by the ONS. This is experimental data that includes the estimated value of exports, imports and balance of goods and services for 2019-2021 for ITL1, ITL2, ITL3, and city regions, split by industry. Data is published annually with a 1.5 year lag. The ONS collects goods export estimates are from HM Revenue and Customs (HMRC) using customs data. Goods exports are apportioned to regions using the local unit of a business (like a shop or office), separate to the reporting unit (like a head office or registered business address). Trade in services estimates have been derived from surveys and other sources, including the International Trade in Services (ITIS) survey, which covers 27,200 businesses annually. The ITIS Survey covers total exports and imports of services broken down by service types (52 in total), the country of origin, and country of destination. This is supplemented by the International Passenger Survey (IPS), which is used for estimates of travel services imports. Bank of England data is also used to break down financial services data to the regional level. Services exports are apportioned to a geographic unit using the same business unit approach as goods exports. One shortcoming of this data is that while it is available at lower geographic scale, the industry disaggregation is relatively high-level, with 7 sectors broken out by goods and services exports. This means that product analysis is not as practically useful as that produced using employment data, which is available at the SIC 3 level. As we examine in Section 4, the export-based measure still provides useful additional insights into the operation of regional economies.

One initial insight, available from just the descriptive data presented in Figure 12, is that the clustering of total exports in London and the South East is even more pronounced than the GVA and GDP per capita examined in Section 2.1. Furthermore, London is the only region of the UK that has shown any significant growth in exports since the pandemic. Many places are still exporting less than they were pre-pandemic.

Total trade exports by International Territorial Level (ITL) 1 region, UK, 2018 to 2021

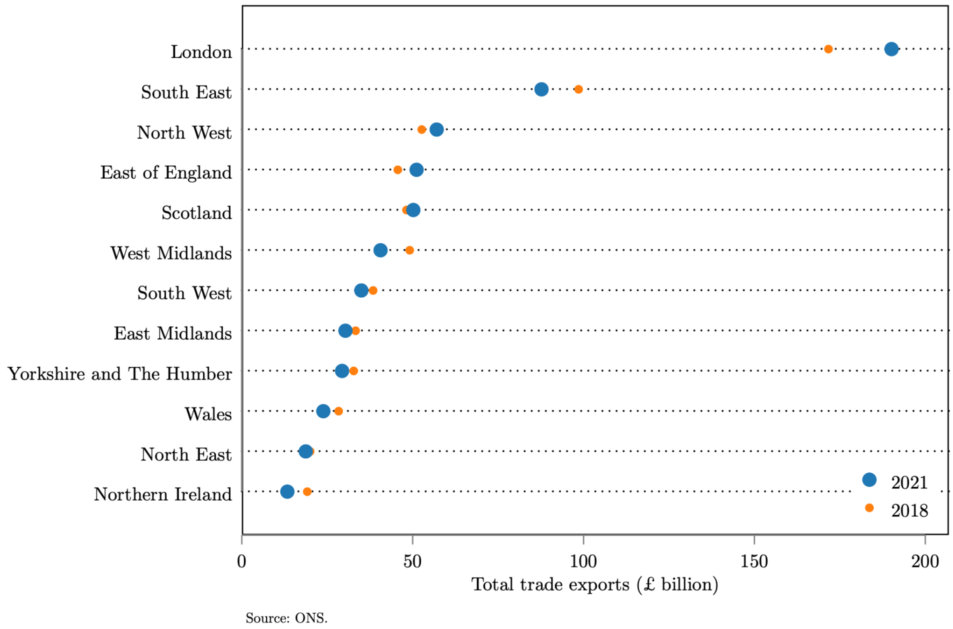


Figure 12: UK sub-national export performance clusters in London and the South East. Post-pandemic, most regions are still below their pre-pandemic level.

Both the employment and exports data cover both goods and services, which is an advancement on the typical ECI measures produced using just goods export data. As noted above, these measures have to infer the existence of services complexity by assuming that they are needed to produce complex goods. No such assumption is required here, facilitating the construction of a network map for the whole economy.

3.3.2 Geographic level

Analysis throughout the paper is conducted at the ITL2 level. ITLs replaced the previous EU classification system, the Nomenclature of territorial units for statistics (NUTS), after the UK left the European Union, but the system remains substantially the same, with both following standards set by the OECD. ITL1 regions are major socio-economic regions, like London and the North East of England. These regions are useful units for some sub-national analysis but, in most cases contain multiple functional economic units. The North West of England, for instance, contains urban areas like Manchester and Liverpool but also large rural areas like Cumbria. ITL2 regions are therefore preferable as they better proxy functional economic areas. They are groups of counties in England, groups of council or local enterprise companies (LECs) in Scotland and group of unitary authorities in Wales. In total, there are 41 ITL2 regions in the UK (versus 12 ITL1 regions and 179 ITL3 regions), each with a population range of 800,000 to 3 million.

ITL2 regions are also reasonable proxies for the newly established Combined Authorities (CAs). CAs (and the counter-parts that also have directly elected mayors, Mayoral Combined Authorities (MCAs)) were established in 2014 as a new governance structure in England, that sought to allow local authorities to pool some of their powers in exchange for central government granting devolved areas more powers and spending autonomy. Each place that wants a CA negotiates a bespoke structure and set of powers with the central government. Greater Manchester was the first such MCA agreed in 2014. As of the start of 2024, there are now 11 combined authorities, of which 10 have directly elected mayors (Sandford 2023). This includes nearly all of England's major cities, including Birmingham (West Midlands), Liverpool, Leeds (West Yorkshire), Sheffield (South Yorkshire), North East (Newcastle), Nottingham (East Midlands) and Bristol (West of England). As well as mapping the political and economic geography of the UK well, ITL2 regions also have a wider range of sub-national statistics available, enabling a wider range of analyses.[9]

3.3.3 Digression: Traded vs. non-traded employment in economic complexity analysis

One question raised in the economic complexity literature is whether to include local industries or just traded goods and services (Michael E. Porter 2003). The national measures computed using exports by definition only capture traded services. Employment data, included that utilised in this paper, of course, captures all sectors of work. In principle, industrial and growth policy should only be concerned with traded services. Those are the sectors capable of export and are an important source of inter-regional income. An important finding is that it

9. For instance, sub-national services export data is available at ITL2 but not at the local authority level, and sector breakdowns at not available at the ITL3 level.

is these industries support employment in nontradable sectors as well. (Moretti 2010) found that, for instance, for each additional manufacturing job in a given city, 1.6 jobs are created in the nontradable sector in the same city in the US, with the effect being significantly larger for skilled jobs, where the estimate is 2.5 jobs in local goods and services. As wages rise in the tradable sector, the demand for local goods and services rises.

One response is that local industries can vary in complexity across places, presenting differing potential future opportunities. SIC code 910, "Libraries, archives, museums and other cultural activities", was present (meaning it had an $RCA > 1$) in Inner London West and the Highlands and Islands of Scotland in 2019. The range and sophistication of the museums and libraries of West Central London are among some of the best in the world. This is a source of economic strength for the people who live there, even if the sector isn't directly tradable. SIC 910 also had an average presence in 2019 of 0.3, lower than the overall overage presence across all industries of 0.34. This means that people who live in areas with museums and libraries have access to products that the average person in the UK does not. This is not only a direct benefit to the living standards of those who live there but could also present a future opportunity to develop related (and potentially tradable), industries, like education technology. Furthermore, insofar as there is some discrete effect from traded industries, the export-based measure of ECI will capture this effect, which is examined further in Section 4.

3.3.4 Control and dependent variables

The regression analysis in Section 4 exploits a number of sub-national data sources. The main dependent variables of interest in the analysis are real Gross Value Added (GVA) and Gross Domestic Product (GDP) per person. I use GVA data from the ONS, who have recently made sub-national estimates available, covering the period 2004-2021. The ONS combines estimates from income (GVA(I)) and production (GVA(P)) approaches to produce a single, balanced estimate of GVA, GVA(B). The UK is the first country to produce a balanced measure of regional GVA. The data is available at the per hour and per job level, with the per hour measure used as the primary variable of interest in the regression analysis. Regional GDP per person covers the period 1998-2021 and is constructed from the GVA series by aggregating the chained-volume measure industry level estimates of regional GVA, and adding back in taxes net of product subsidies.

As an additional variable of interest, I supplement GVA per hour and GDP per person estimates with Gross Disposable Household Income (GDHI), also from the ONS. GDHI is the amount of money that all people in the household sector have available for spending or saving after direct and indirect taxes are paid, and any direct benefits are received. The per head measure is used throughout the analysis. Using GDHI has some weaknesses. Ideally, a direct measure of income would be used that doesn't include the effect of redistribution. Unfortunately, the Annual Survey of Hours and Earnings (ASHE), and the Labour Force Survey (LFS) do not produce earnings or income data at the ITL2 level. The regression analysis that follows in Section 4 should therefore be seen alongside the GVA and GDP per person analysis as corroborating the effects observed. The results from both variables are consistent across regression models.

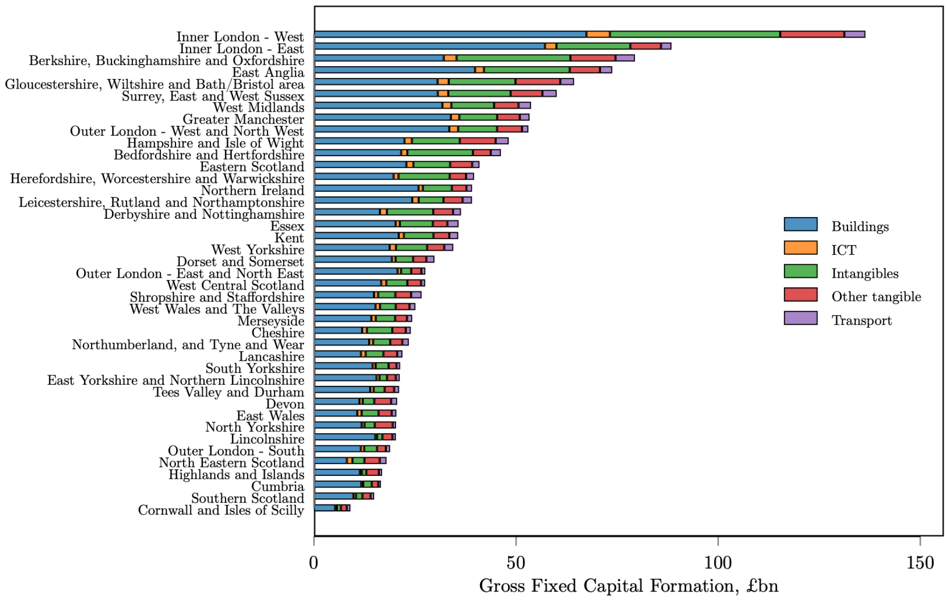
GDHI is also useful because it is the ONS's choice of measure for 'material welfare' within the household sector.[10] It is natural for policymakers to be concerned with the effect of complexity on material outcomes for citizens, as well as direct productivity gains. This is particularly the case given the lack of a one-to-one relationship between productivity improvements and increasing income (Grossman and Oberfield 2022).

The main controls used in the analysis are Gross Fixed Capital Formation (GFCF) and the share of the working population with at least tertiary education. GFCF measures the flow of capital over a given period of time. Where GFCF exceeds the depreciation of existing capital, net capital formation will be positive, resulting in an increase in the net capital stock. Regional GFCF is another new dataset from the ONS. It covers the period 1997-2020 at the region-industry level. It is available down to the ITL3 and 2-digit SIC level and, for the first time, has an accompanying asset breakdown covering, buildings and structures,

10. The Scottish Government also use GDHI as one of the main indicators of economic performance, including in its formal Economic Strategy. The Welsh and Northern Irish governments follow a similar approach.

transport equipment, ICT equipment, other tangible assets and intangible assets (Martin and Becker 2023). Human capital is proxied by the share of the working population with at least tertiary education. This data is provided by the OECD at the ITL2 level for the period 2015-2019.

Gross Fixed Capital Formation, by asset and ITL2 region, 2017-2020 total



Source: ONS.

Figure 13: GFCF is heavily clustered in London and the South East.

3.3.5 Initial application of economic complexity analysis

With the appropriate data in hand, the following section sets out how the economic complexity computations are executed. The resulting measures are then applied to the analysis contained in Section 4. First, I visualise the X_{cp} matrix set out in Section 3.2 and Equation 4 using the data for 2019. The plot is ordered by diversity and ubiquity, i.e. places with employment in more sectors are at the top and sectors with more employment are on the right. The heatmap shows the triangular pattern we would expect to see in the data, with places like London exhibiting relatively high and concentrated employment in a number of sectors clustered in the top right of the plot.

Figure 15 them shows the M_{cp} computed using Equations 2 and 3. This matrix is ordered in the same way as the X_{cp} plot and also exhibits the triangular pattern. A different ordering of regions emerges, however, with places one might

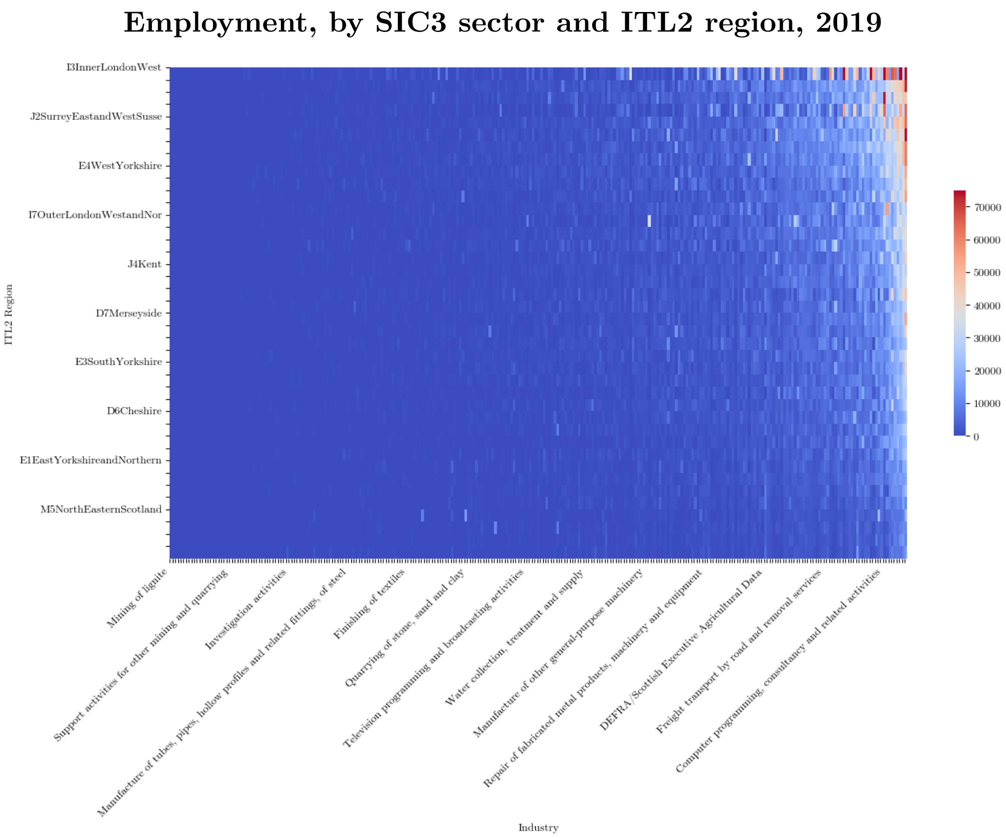


Figure 14: The X_{cp} matrix shows sectors of regional employment strength.

less expect showing a high number of sectors with an $RCA > 1$. As we will come on to discuss in Section 4, this pattern is a function of complexity being the combined product of having competitive advantage in *rare* sectors, rather than many sectors, per se.

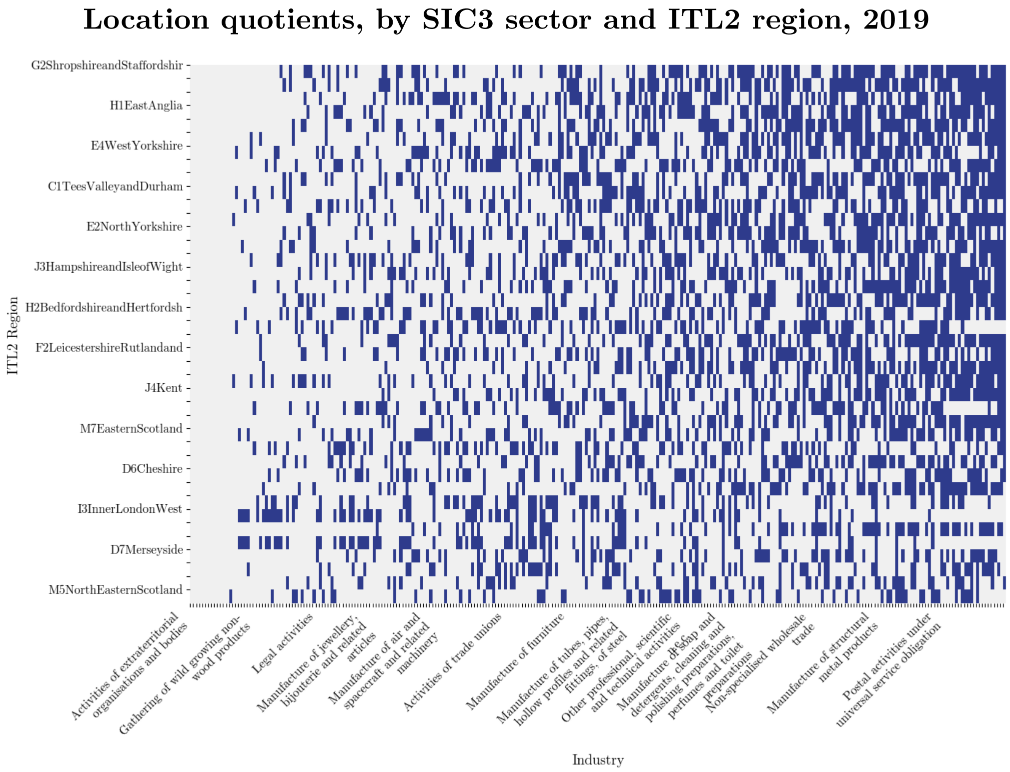


Figure 15: The M_{cp} matrix shows the revealed comparative advantage of regions in sectors.

With the method and data established, I now move to the analytical portion of the paper.

4 Analysis and Results

4.1 Applying economic complexity to the UK's regions

4.1.1 Descriptive analysis of economic and product complexity across the UK's regional economies

Figure 16 shows the ECI computed using employment data at the ITL2 level and 3-digit SIC code for regions with an ECI > 1.75, < -0.5 and the Core Cities. The Core Cities are an advocacy group of large regional cities in the UK outside London and Edinburgh. Formed in 1995, it includes the city councils of eleven city councils: Belfast, Birmingham, Bristol, Cardiff, Glasgow, Leeds, Liverpool, Manchester, Newcastle, Nottingham, and Sheffield. Due to data availability, Belfast is often not included in the analysis in this section.

As the ECI has the property $ECI \sim \mathcal{N}(0, 1)$, the metric lends itself to straightforward interpretation. London's regions are all in the far-right tail of the distribution, $> 2\sigma$ from the mean. By comparison, the UK's Core Cities all cluster around, and in some cases are some way below, the mean. Perhaps particularly concerning is the state of the UK's smaller former industrial cities, like Newcastle, Cardiff, Sheffield, and Nottingham. As Figure 16 shows, these places exhibit ECIs comparable to places with significantly fewer economic assets, including a number of rural areas like Devon, Eastern Scotland and the Highlands, all of which lack major urban areas. Looking geographically, Figure 17 shows that the UK's economic complexity clusters in London and the South East, with a sharp drop-off in neighbouring regions like the West Midlands and the East of England. There are a couple of bright spots outside London and the South East. West Central Scotland, which contains Glasgow, is relatively complex, supported by its specialisation in motion picture, video and television programme activities. This tracks Glasgow's reputation as a global location for film and TV production.[11] Manchester is similarly strong in the creative industries – this tracks the BBC moving a substantial portion of its operations to Salford – but also advertising, market research, and sound recording and music publishing.[12]

11. See, for instance, that the Flash, Indiana Jones, Bat Girl and Tetris were all filmed there in 2021: https://www.glasgow.gov.uk/index.aspx?articleid=27932

12. This is also a well-reported and long-standing strength of Manchester, as the historic home

Employment-based ECI, by ITL2 region, 2015-2022 average, selected places

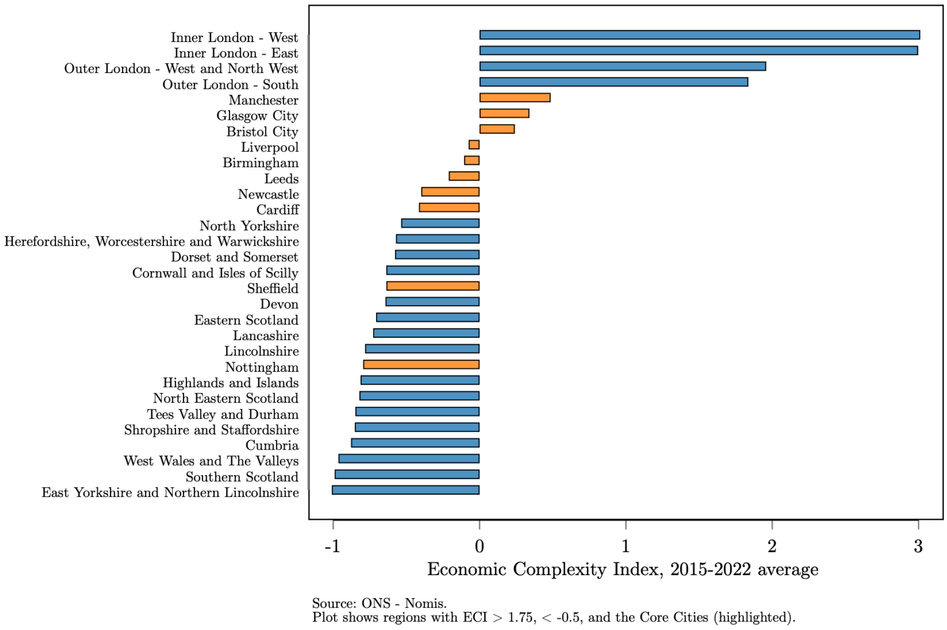


Figure 16: The UK's cities significant lag London in economic complexity.

Employment-based ECI, by ITL2 region, 2015-2022 average

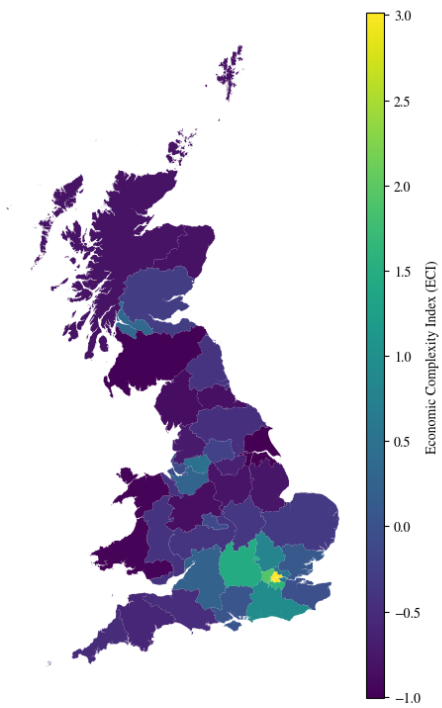
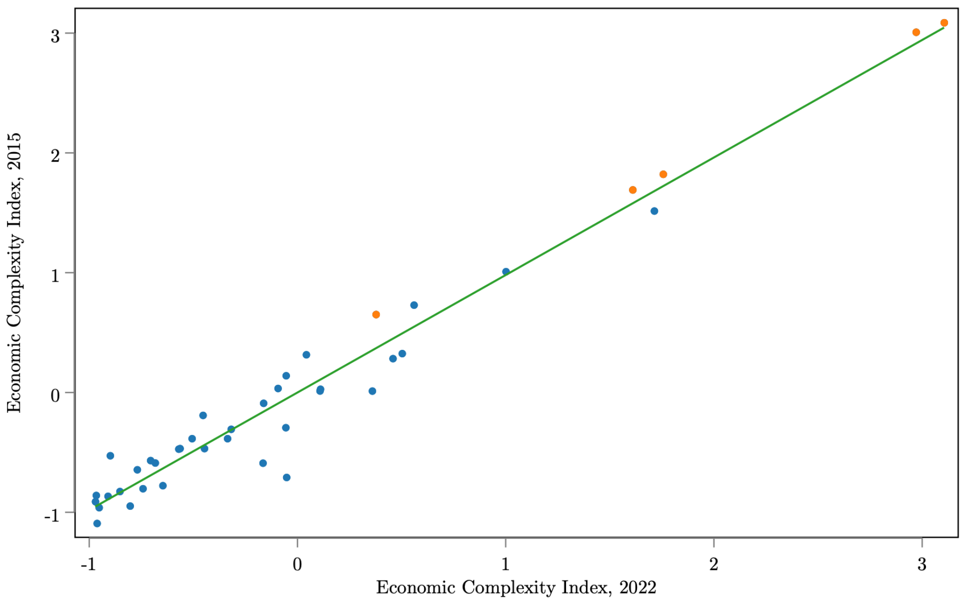


Figure 17: The UK's economic complexity clusters in the South East.

of UK punk rock in the 70s, the 'Madchester' indie dance scene of the 80s and Britpop in the 90s.

ECI at the sub-national level also exhibits substantial persistence over time. Figure 18 shows the ECI of ITL2 regions in 2022 and 2015. It has a correlation $\rho = 0.98$. This is consistent with other national economic complexity analyses. For instance, (Weber et al. 2021) found reasonable persistence in productive capabilities, proxied by economic complexity and export diversification, over a century. This is perhaps not surprising, given the complexity of production networks and the difficulties associated with changing industrial structure. Relevant to the policy implications below, this is also indicative of the level of effort and resources required if regional economies are to become more complex over time.

Employment-based ECI, by ITL2 region, 2015 and 2022



Source: ONS - Nomis. London and surrounding areas highlighted.

Figure 18: High-value services exhibit very high levels of product complexity.

Turning to Product Complexity Index (PCI), Figure 19 shows PCI scores for the top and bottom 15 industries, averaged over the period 2015-22. As with ECI, PCI is a normalised statistic, with $PCI \sim \mathcal{N}(0.26, 2.39)$. Given this, the range of complexity across sectors is significant. The most advanced sectors are $> 2\sigma$ from the mean, consistent with the ECI measures. Given the UK's reported strength in services (De Lyon et al. 2022), it is perhaps not surprising that these sectors, ranging from financial services to consultancy and the creative industries, dominate the list of the most complex sectors. By comparison, the least complex sectors are made up of mostly low-value manufacturing sectors. Two sectors are worth highlighting because they speak to the potential misallocation of current industrial policy efforts; forging, pressing and roll-forming of metal, and the manufacture of basic iron and steel and of ferro-alloys. Despite their low measured complexity, huge political and financial effort has been put towards supporting the British steel industry in recent years (Hutton and Rhodes 2021) (see also the discussion above in Section 2.1 of the UK's intereventions to support Tata steel at Port Talbot). Recalling the discussion above in Section 3.3.3, while most complex sectors are traded, it is not by any means exclusively the case, with local services like travel agency and event catering both exhibiting relatively high complexity (3.1 and 3.2 average PCI, respectively).

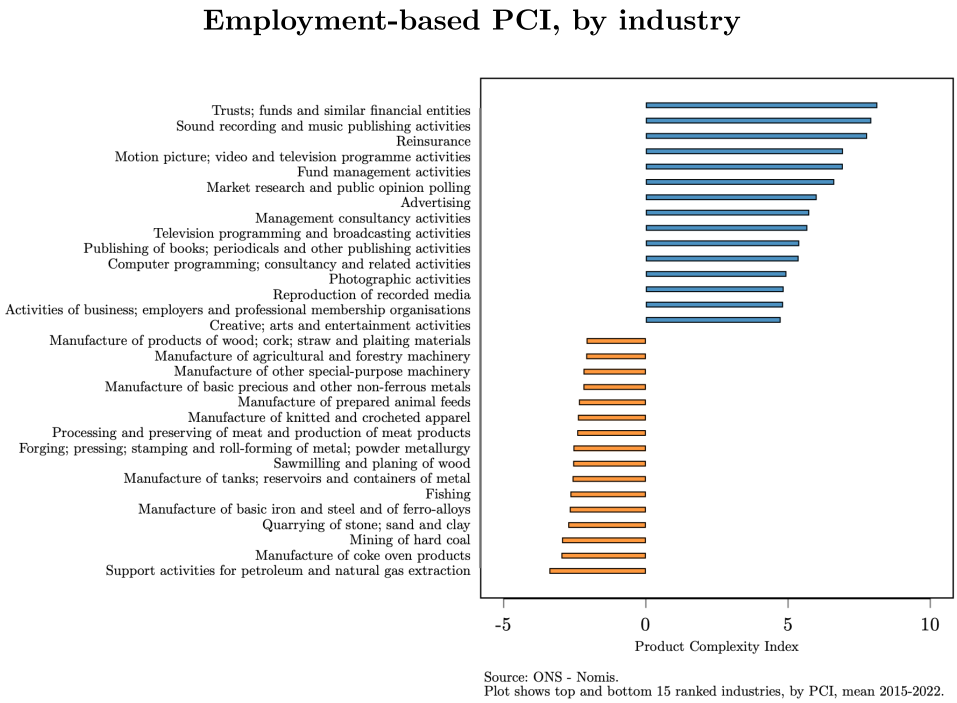


Figure 19: High-value services exhibit very high levels of product complexity.

ECI isolates complexity as distinct from having a lot of comparative advantage, as discussed in Section 2.2. A problem with traditional measures like RCA is that places can have a high RCA in several sectors but not be complex. As per Equation 5, diversity, D, is the sum of sectors a place has a comparative advantage in. One might expect those places to be the most complex. What the ECI measures show us is that this is not the case. The most diverse place in the data used here is West Wales, with $D = 117$. The challenge for West Wales is that a number of these 117 sectors where it has a comparative advantage are also very common i.e. they have high ubiquity, U, scores, like building completion and finishing, sale of motor vehicles, and manufacture of bakery and farinaceous products.

This result holds when looking at the whole data set, as in Figure 20, where we see that there is a negative relationship between sectoral diversity and ECI. The two most complex regions, Inner West and East London, have comparative advantage in many fewer sectors than West Wales but have nearly 4σ higher complexity scores. This is because London specialises in more scarce, and by implication, more complex, services like sound recording and music publishing activities, and television programming and broadcasting activities, which are only present in a few regions. This runs contrary to many popular intuitions about economic sophistication, i.e. that places with specialisation in a greater number of sectors should be more complex per se, but is not borne out in the complexity literature (Mealy, Farmer, and Teytelboym 2019).

ECI and sectoral diversity, by ITL2 region, 2015-2022

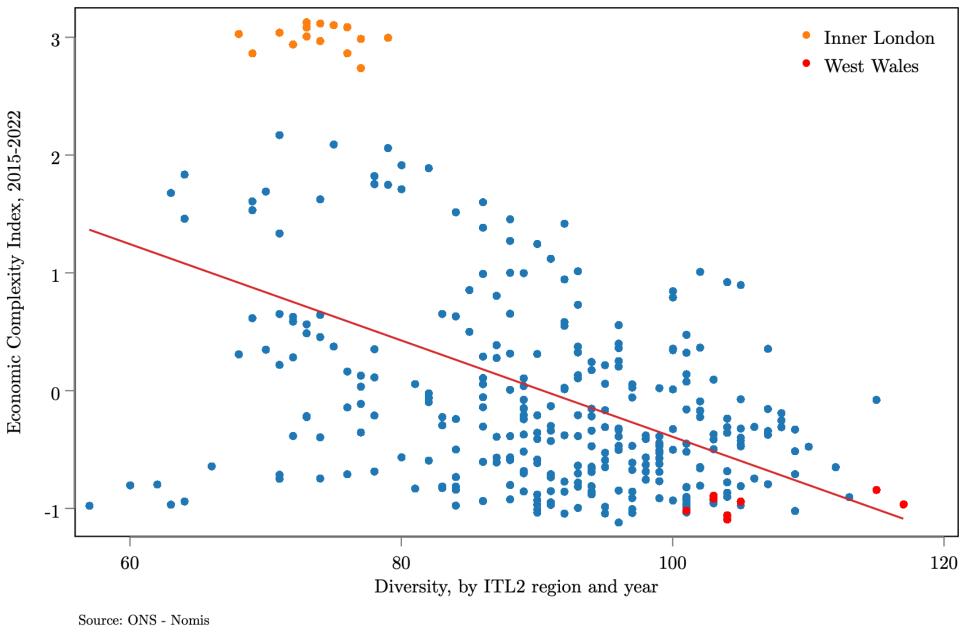


Figure 20: The higher the number of sectors a place has an RCA, the lower its complexity, in expectation.

4.2 Understanding the relationship between living standards and economic complexity: Going beyond traditional input relationships

It is well-established that economic complexity measures are more effective predictors of national economic growth than traditional predictors like education levels, institutions, and current GDP per person (Hausmann and Hidalgo 2014). This section develops an analysis of key economic metrics, like GVA per hour and job, and GDP per person, as functions of ECI. The results indicate that complexity measures perform well at predicting productivity and output per person, even after controlling for traditional input variables.

4.2.1 Regression framework

I examine the impact of employment and export-based economic complexity indices on key economic outcomes, specifically real GVA per hour and GDP per person. The core of the empirical strategy is articulated through two regression models. The first model is a cross-sectional analysis, while the second extends the framework into a panel analysis, incorporating urban fixed effects to control for unobserved heterogeneity across area types.

The initial model can be expressed as follows:

$$y_i = \beta_0 + \beta_1 \mathrm{ECI}_i^{\mathrm{emp}} + \beta_2 \mathrm{ECI}_i^{\mathrm{exp}} + \boldsymbol{\beta}_3^T \mathbf{X}_i + \epsilon_i, \tag{11}$$

where y_i denotes the dependent variable for observation i, capturing either the natural logarithm of GVA per hour ($\ln(\text{GVA per hour in pounds})_i$) or the natural logarithm of GDHI per person ($\ln(\text{GDP per person in pounds})_i$). The independent variables of interest are $\mathrm{ECI}_i^{\mathrm{emp}}$ and $\mathrm{ECI}_i^{\mathrm{exp}}$, representing the ECIs calculated using employment and exports data, respectively. $\mathbf{X}_i$ is a vector of control variables, including the natural logarithm of Gross Fixed Capital Formation ($\ln(\mathrm{GFCF}_i)$) and the share of the working age population with at least tertiary education ($\mathrm{HC}_i^{\mathrm{tertiary}}$). $\boldsymbol{\beta}_3$ denotes the coefficients for these control variables, and ϵ_i is the error term.

To account for unobserved geographic characteristics that may influence the economic outcomes, I extend the analysis through a fixed effects regression, given by:

$$y_{it} = \alpha_i + \beta_1 \text{ECI}_{it}^{\text{emp}} + \beta_2 \text{ECI}_{it}^{\text{exp}} + \boldsymbol{\beta}_3^T \mathbf{X}_{it} + \gamma_t + \epsilon_{it}, \tag{12}$$

where y_{it} reflects the dependent variable for region, i, at time, t. The model introduces α_i, representing place-specific fixed effects, and γ_t, capturing year-specific effects to control for any temporal trends or shocks that could uniformly affect all regions. Due to the shorter nature of the panel $\mathbf{X}_{it}$ is the vector of control variables adjusted for the panel structure, and the rest of the variables maintain their definitions as in the cross-sectional model.

Additionally, to explore the potential for both within and between variation, I employ a random effects model:

$$y_{it} = \gamma + \beta_1 \text{ECI}_{it}^{\text{emp}} + \beta_2 \text{ECI}_{it}^{\text{exp}} + \boldsymbol{\beta}_3^T \mathbf{X}_{it} + \upsilon_i + \epsilon_{it}, \tag{13}$$

where υ_i is the random effects associated with each region, capturing unobserved heterogeneity that varies across regions but is constant over time. This model accounts for the variability between regions while also considering the influence of time-invariant regional characteristics.

The fixed effects model incorporates fixed effects using urbanisation. This is an OECD measure, which categorises ITL3 regions into either 'predominantly urban', 'urban with significant rural' or 'predominantly rural' based on the share of people living in local rural units. Using fixed effects at the urban-mixed-rural level is preferred to using the ITL2 regions themselves. This approach is justified by the literature on agglomeration discussed in Section 2.1.2 that emphasise the role of urbanisation in economic development, innovation, and productivity. Urban fixed effects is also a practical specification given the size of the panel. In the fixed effects models, $n = 198$, and $n = 66$ respectively. With 40 ITL2 regions, this would significantly reduce the degrees of freedom (DoF) for the purposes of estimation. Controlling for urbanisation effects therefore strikes a balance between capturing the relevant effects of place, separate to the parameters in the model, while maintaining a robust estimation strategy.

The empirical investigation proceeds in two stages. Initially, I estimate the cross-sectional model to identify the baseline relationships between economic complexity and the outcome variables, controlling for relevant economic and demographic factors. After this, I leverage the panel data, controlling for unobservable regional characteristics that remain constant over time, thus mitigating potential omitted variable bias. A series of robustness checks can be found in Appendix A.

4.2.2 Regression results

Figures 21 and 22 show the correlation between ECI and key economic outcomes for ECI generated using sub-national employment and export data. The correlates are log of GVA per job, log of GVA per hour, and log of GDHI per person and log GDP per person. Across the panel there is a strong positive association, consistent with the wider global and national literature (Hausmann and Hidalgo 2014; Tacchella et al. 2012; Hidalgo 2021; Gao and Zhou 2018; Chávez, Mosqueda, and Gómez-Zaldívar 2017). More complex regions are associated with higher productivity, output, and incomes. The export-based measure has lower explanatory power compared to the employment-based measure, but still performs well, despite the smaller sample and sector set ($R^2 = 0.65$ vs $R^2 = 0.30$ for GVA per hour, respectively).

Turning to the regression models, Table 2 contains the cross-section estimates of the power of the two ECI measures to explain productivity (measured as the log of GVA per hour at the ITL2 level) and incomes (the log of GDP per person at the ITL2 level). The models using just the ECI^{emp} cover the period 2015-2022, whereas the ECI^{exp} models cover 2019-2021 (as discussed above in Section 3.3. To compare the performance of the ECI measures against traditional metrics, I include measures of capital and labour as controls in the form of the log of Gross Fixed Capital Formation at the ITL2 level and the share of workers with tertiary education qualifications at the ITL2 level. For the purposes of interpretation, recall that ECI measures are standardised units and the tertiary share variable is in units, $(0-100)$ so a one percentage point change is associated with a change in Y by $\beta \times 100)$ percent.

The GVA per hour results show a consistent and significant effect of both ECI measures. The employment-based measure shows that a σ increase in ECI^{emp} is associated an approximately 10.1% increase in GVA per hour, after controlling

Employment-based ECI and key economic outcomes, 2015-2022

Figure 21: ECI is highly associated with key regional economic outcomes.

Export-based ECI and key economic outcomes, 2015-2022

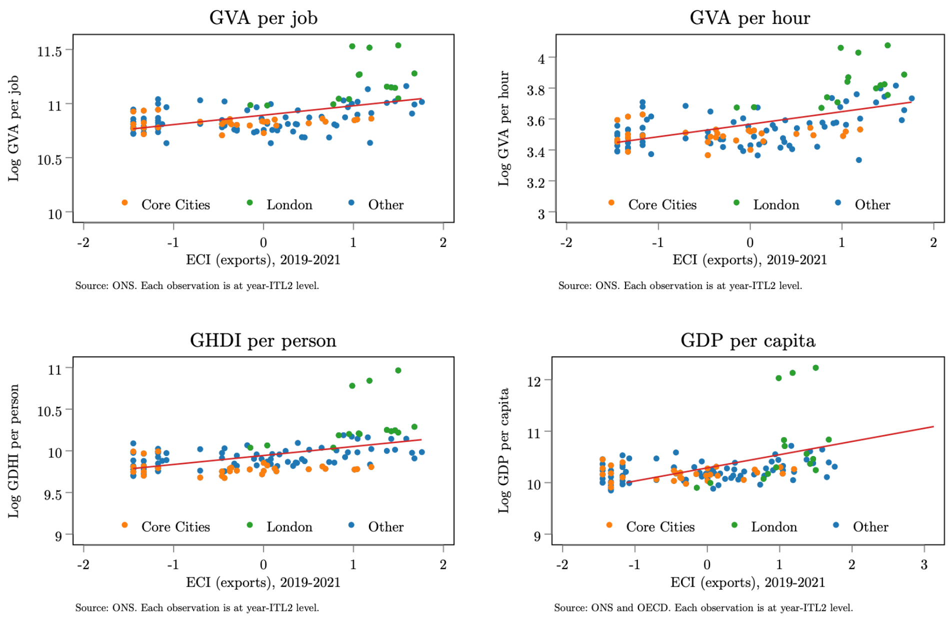


Figure 22: The export-based measure of ECI is also highly associated with key regional economic outcomes.

for capital investment (model 1). This estimate falls after the inclusion of the human capital control to 3.4% (model 3). Across the board, the ECI estimate is expected to fall after the inclusion of controls, especially the human capital control in the case of the employment-based ECI measure as the measure itself is computed using employment shares, thereby partially capturing the underlying human capital capability. Thought of in that sense, the ECI measure with controls represents the residual 'capabilities' requires to produce a certain level of productivity or output. The export-based measure has a significant positive relationship with GVA per hour. A σ increase in ECI^{exp} is associated with a 5.23% increase in GVA per hour, controlling for capital investment (model 2). When the human capital control is added to the model (model 4), the coefficient remains positive but losses significance.

The GDP per person results follow a similar pattern. The employment-based measure shows that a σ increase in ECI^{emp} is associated an approximately 19.6% increase in GDP per person, after controlling for capital investment (model 7). The export-based measure is also significant at the 5% level, showing a σ increase in ECI^{exp} is associated with a 4.9% increase in GDP per person, controlling for capital investment (model 8). Both measures lose significance once human capital controls are included, as is also the case in the mixed models, with the exception of the employment measure in model 11. The sample size of these models is perhaps a barrier to clearer identification of the effects in these models.

Appendix A, Table 5, contains the results for household income. The estimates indicate higher complexity is also positively and significantly associated with higher incomes. The effect of ECI^{emp} is generally larger than in the GVA models, with coefficients of 0.166 and 0.059 for GDHI (models 7 and 9) compared to 0.109 and 0.0338 for GVA (models 1 and 3). ECI^{exp} exhibits a similar pattern to the corresponding GVA models. Across both dependent variables, the export-based models are inhibited throughout by the smaller sample size within the cross-section of ECI^{exp} and the human capital control ($n = 80$ vs $n = 39$).

Table 3 expands the analysis to the panel specifications. The first 3 models for each dependent variable for random effects models, and the last 3 contain urban fixed effects. Comparing Model 1 in Tables 3 and 2 compares the standard OLS regression to random effects (RE) computed with generalized least squares (GLS). Recall that, OLS assumes independence of observations and does not account for unobserved heterogeneity across panels. It treats the regression as if pooling all cross-sectional units, without considering the within-unit correlation over time. Random effects, on the other hand, acknowledges that data are collected across both time, (t), and entities, (i), and allows for unobserved effects that vary across entities but are constant over time. These unobserved effects could capture intrinsic characteristics of the regions that are not included as explanatory variables. Therefore, β^{OLS} assumes there's no omitted variable bias or unobserved heterogeneity affecting the estimates. β^{RE}, on the other hand, assumes unobserved heterogeneity across entities (regions) but that these effects are uncorrelated with the regressors. β^{RE} provides a weighted average of the within and between estimates, providing an insight into the effect controlling for unobserved, entity-specific factors. As such, the model fit statistics also vary. In RE, the R^2 is decomposed into within, between, and overall, indicating how well the model explains variation over time within entities, across different entities, and overall, respectively.

The ECI^{emp} coefficient is lower in the RE model compared to OLS for both the GVA and GDP specifications, suggesting that when accounting for unobserved regional characteristics that are constant over time, the impact of economic complexity on GVA per hour is smaller than suggested by OLS. It is also worth noting that the $GFCF$ coefficient is significantly higher in the RE model, indicating that capital investment's impact on GVA per hour is more pronounced when controlling for unobserved regional heterogeneity. This suggests that there are some region-specific factors not incorporated into the model, which could include factors like institutional quality or market integration and access, that are associated with productivity independently of ECI. That said, the relatively small difference in coefficients suggests that the substantial share of variation is captured by differences in ECI across regions.

Models 4-6, and 10-12 in Table 3 introduce urban fixed effects. The fixed effect models control for any time-invariant characteristics of urban, semi-urban, and rural categories that might affect the dependent variable, effectively removing omitted variable bias associated with those fixed characteristics. FE models do not assume the effects are random and allow for correlation between the unobserved urbanity characteristics and the regressors.

The inclusion of fixed effects does not change the positive and significant ECI coefficients, showing an association between GVA per hour and both ECI^{emp}

and ECI^{exp} of between 11.9% and 5.2%, for a σ increase in complexity (models 4 and 5). In the GDP per capita panel, the coefficient on ECI^{emp} in model 10 remains significant and higher than in both the RE and cross-section models, indicating an association between a σ increase in ECI and GDP per capita of 21.5%. The coefficient on ECI^{exp} however, loses significance in the fixed effects specification (model 11).

The role of region type appears to have a mixed effect across the models. The F-statistic for the joint significance of the urban fixed effects in model 5 of GVA per hour is 1.02, $P = 0.36$. It is also not significant for model 6. This suggests that factors captured by economic complexity and capital investment are more predictive of regional productivity, and that urbanisation is not a significant marginal explanatory factor. This is consistent with findings elsewhere in the literature, that shows economic complexity is higher in urban areas (Balland et al. 2020; Balland and Rigby 2017). We should therefore not expect to see further significant urban-rural variation once ECI is incorporated into a model explaining regional productivity patterns. There is however, a more significant effect in the GDP per person models. The F-statistic for the joint significance of the urban fixed effects in model 10 of GVA per hour is 20.36, $P = 0.000$. It is similarly significant in the mixed model 12. Further investigation of this point and the specific effects of urbanisation would be required to establish the underlying effects driving this result.

4.3 Applying economic complexity to industrial policy

With the understanding of economic complexity's power in predicting key economic outcomes at the sub-national level, I now turn to policy applications. As discussed in Section 2.1.3, UK industrial policy has suffered from a lack of coordination that has inhibited its success. A significant challenge for policy-makers is targeting interventions efficiently. There is disagreement about how to determine where to focus interventions, which leads to further coordination issues between national and local government. Moreover, the economic geography literature discussed in Section 2.1 and ECI analysis discussed in Section 4, suggests that there is significant geographic heterogeneity in industrial structure and that interventions may therefore need to be tailored carefully to the differ-

Table 2: Cross-section results

	ln GVA per hour						ln GDP per capita					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ECI Emp	0.109***		0.0338***		0.0995***	0.0302	0.196***		0.0176		0.188***	0.00805
	(0.00638)		(0.00755)		(0.0117)	(0.0186)	(0.0287)		(0.0260)		(0.0567)	(0.0578)
ECI Exports		0.0523***		0.0101	0.0263***	0.00944		0.0488**		−0.0392	−0.000523	−0.0394
		(0.00999)		(0.0106)	(0.00990)	(0.0114)		(0.0228)		(0.0283)	(0.0246)	(0.0285)
ln GFCF	0.0511***	0.134***	0.0635***	0.0853***	0.0304	0.0620**	0.163***	0.372***	0.177***	0.225***	0.175***	0.219***
	(0.0133)	(0.0262)	(0.0103)	(0.0170)	(0.0226)	(0.0256)	(0.0315)	(0.106)	(0.0320)	(0.0616)	(0.0596)	(0.0772)
Sh Tertiary			0.0101***	0.0113***		0.00938***			0.0272***	0.0287***		0.0282***
			(0.000732)	(0.00124)		(0.00167)			(0.00352)	(0.00793)		(0.00840)
Constant	3.039***	2.349***	2.482***	2.275***	3.282***	2.568***	8.831***	6.929***	7.575***	7.043***	8.697***	7.121***
	(0.119)	(0.234)	(0.101)	(0.148)	(0.205)	(0.272)	(0.280)	(0.931)	(0.384)	(0.808)	(0.534)	(0.977)
Observations	240	80	195	39	80	39	240	80	195	39	80	39
R^2 adjusted	0.694	0.549	0.831	0.841	0.750	0.852	0.567	0.426	0.747	0.731	0.545	0.724

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses.

Table 3: Panel results

VARIABLES	ln GVA per hour (1)	(2)	(3)	(4)	(5)	(6)	ln GDP per capita (7)	(8)	(9)	(10)	(11)	(12)
ECI Emp	0.0776***		0.0994***	0.119***		0.113***	0.0940***		0.134**	0.215***		0.225***
	(0.0112)		(0.0159)	(0.00659)		(0.0132)	(0.0363)		(0.0679)	(0.0257)		(0.0532)
ECI Exports		0.0189*	0.0123		0.0519***	0.0183*		0.00866	−0.00689		0.0255	−0.0413*
		(0.0102)	(0.00912)		(0.0109)	(0.0106)		(0.0287)	(0.0303)		(0.0214)	(0.0242)
ln GFCF	0.100***	0.0342	−0.00234	0.0535***	0.153***	0.0406**	0.109***	0.403***	0.287***	0.246***	0.505****	0.281*****
	(0.0210)	(0.0324)	(0.0198)	(0.0125)	(0.0309)	(0.0180)	(0.0292)	(0.111)	(0.0661)	(0.0437)	(0.128)	(0.0764)
Constant	2.601***	3.248***	3.575***	3.005***	2.166***	3.176***	9.312***	6.653***	7.687***	8.058***	5.697***	7.706***
	(0.190)	(0.281)	(0.179)	(0.112)	(0.279)	(0.163)	(0.245)	(0.976)	(0.589)	(0.387)	(1.133)	(0.681)
Observations	240	80	80	198	66	66	240	80	80	198	66	66
Urban FE	NO	NO	NO	YES	YES	YES	NO	NO	NO	YES	YES	YES
R^2 within	0.064	0.011	0.006				0.0418	0.255	0.231			
R^2 between	0.740	0.571	0.748				0.582	0.437	0.557			
R^2 overall	0.670	0.5413	0.736				0.568	0.430	0.545			
R^2 adjusted				0.788	0.592	0.826				0.662	0.501	0.652

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Robust standard errors in parentheses.

ing needs of places. Policymakers therefore need a measure that allows them to grapple with both geographic and industrial variations when forming policy. As discussed in Section 3.2, one of the advantages of economic complexity measures over statistics like the Krugman Index, is that they allow us to 'get under the hood' and examine the underlying drivers of industrial composition.

Recall, that economic complexity builds on the idea of industrial 'relatedness' i.e. that places are likely to develop into industries that are similar to their existing industrial strengths. This relatedness is defined in reference to the overall network of production. The idea rests on an intuitive notion. Imagine that if place, A, produces jet engine parts, landing gear and fuselage components, one could guess, without any further analysis, that it is likely to diversify into some form of aircraft assembly next. This process can be determined for the whole production system by looking at places, B, C, D, that produce these three goods as well as also specialising in aircraft assembly to confirm the intuition.

If these diversification opportunities are systematically identified, they could be used to inform industrial policy. Places are likely to have the greatest success and return on investment by focusing on sectors adjacent to existing strengths. Notwithstanding the challenge of determining these sectoral opportunities quantitatively, the idea sounds obvious but often runs contrary to popular debates about industrial policy. One common theme is a form of 'industrial nostalgia', that romanticises a return to Britain's industrial heyday (Tomlinson 2019; De Lyon et al. 2022).[13]

To conduct this analysis, I define a proximity matrix ϕ_{pq}, which contains the conditional probability of two industries p and q existing in the same place. Since conditional probabilities are not symmetric, I take the minimum probability. (Hidalgo 2023) provide a nice example to illustrate why: suppose that 17 countries export wine, 24 export grapes and 11 export both, all with $RCA > 1$. The resulting proximity between the wine and the grapes is $11/24 = 0.46$, using the 24 instead of the 17 to reduce the probability that the relationship is false. Proximity, ϕ, in place, c, and industries, p, and, q, are therefore given by:

13. This is not just the case in Britain, see, for instance in the US: https://www.cfr.org/article/curse-nostalgia-industrial-policy-united-states

$$\phi_{pq} = \min\left(\frac{\sum_c M_{cp} M_{cq}}{\sum_c M_{cp}}, \frac{\sum_c M_{cp} M_{cq}}{\sum_c M_{cq}}\right) \tag{14}$$

The proximity matrix, ϕ, is then used to construct a measure of 'density'. Density captures the average presence, defined by the M_{cp} matrix, of industries around other industries. This is then weighted by the proximity matrix. Per (Hidalgo et al. 2007), the density measure ω_{cp} computes the average proximity of a new sector, q, to all existing sectors in a place, c, given by:

$$\omega_{cq} = \frac{\sum_p M_{cp} \phi_{pq}}{\sum_p \phi_{pq}} \tag{15}$$

The distance, d, between a place's current sectoral composition and a new sector is therefore just the inverse of this:[14]

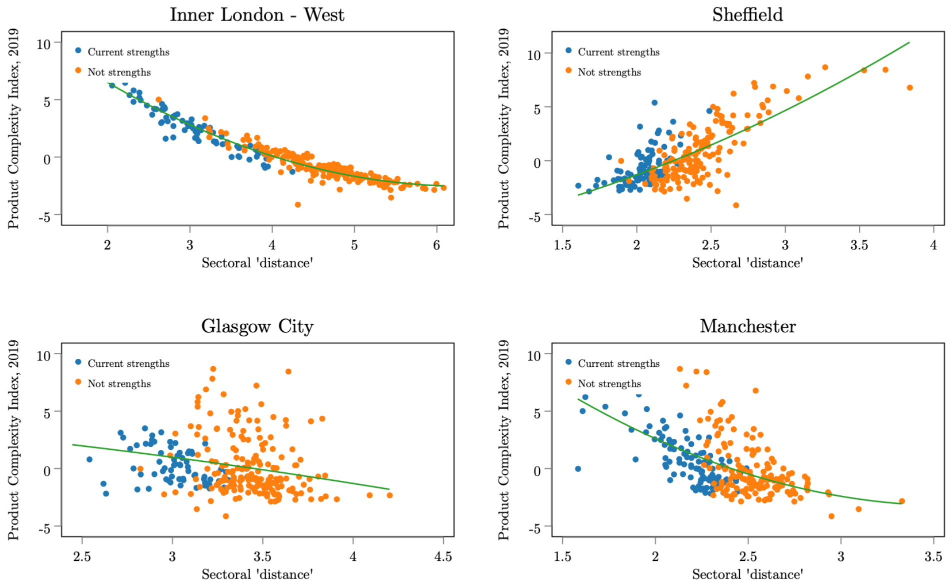
$$d_{cq} = \frac{1}{\omega_{cq}} \tag{16}$$

Plotting distance, d, against PCI gives a useful picture of the overall economy in a place. Figure 23 does this, highlighting current strengths (defined as where the place has an $RCA > 1$). The shape alone of these plots can tell us a lot about the industrial prospects of a place. Ideally, a place should want to look like West Inner London. Several points emerge. First, West Inner London is specialised in high complexity industries, shown by the cluster of blue plots in the top left. While it has a presence in lower complexity sectors (shown in orange), it is not 'overcommitting' resources to these sectors. Second, the plot is tightly clustered and has a negative correlation. This shows that London's sectors get more complex as they are more central to the production network, a trait all economies should want. Finally, it has potential opportunities (the orange plots) in the top left, which is important given our interest in places diversifying into industries that are both complex and proximate to existing strengths.

By comparison, the general positive shape of Sheffield's economy is a cause for concern. Its existing strengths are clustered in the bottom left, with all of its potential opportunities in higher complexity sectors a long distance away from its existing industrial structure. It has a long way to travel before it can access

14. Sometimes, distance is defined as $1 - \omega_{cq}$ but this choice doesn't substantively affect the analysis.

Comparing PCI and sectoral 'distances', UK cities, 2019



Source: ONS - Nomis. Strengths are sectors where the place has an RCA > 1.

Figure 23: The product structure of local economies forms an overall picture of local economic strength.

many of the higher complexity industries. Part of this shape, it should be said, is a function of how the measures are produced. High ECI places will tend to have downward slopes as a mechanical result of how they are constructed. Their differential shape nevertheless can tell us important things about the challenges faced by the UK's regions. Low ECI places will tend to have the positive correlation of Sheffield but a steeper slope, for instance, would suggest that the returns to diversification would be higher for the other place.

This is what can be seen comparing Manchester and Glasgow. Manchester's slope is not only negative but steeper than Glasgow's, suggesting it stands to enjoy higher returns to diversification efforts, at least in the short term. In general, we should expect Manchester to access those returns more easily than Glasgow as well, given the average distance of its opportunity sectors is considerably lower than Glasgow's. Glasgow does have an advantage over Manchester through its greater number of 'foothold' sectors, though. These are sectors where a place has an existing strength in a high PCI area, that is close to other opportunity sectors. These are the small number of orange plots among the blue plots to the middle-left. If Glasgow can bridge from these existing strengths into adjacent sectors, it may be able to avoid having to fully 'traverse the ladder' to get to a particular high-value sector.

I now turn to the specific insights this analysis can offer places that seek to better target their industrial policy. By focusing on the current strengths and production networks in places, potential opportunities for future diversification can be isolated, and then targetted through industrial policies. Figures 24, 26, and 25 do this for Manchester, Newcastle and Leeds respectively, using the 2019 vintage of the data. The plots show the same PCI-distance space as Figure 23, but I have further overlaid them with the green 'opportunity' markers. These are sectors that are complex ($PCI > 1$), proximate to the place's existing sectoral composition (d is less than the 50th percentile of the area-year distances), but not an existing strength ($RCA < 1$ i.e. $M_{cp} = 0$).[15]

This analysis has several implications. First, the opportunity sectors are grounded in the existing (often less complex) industrial structure of a place. This offers the advantage of making them plausible targets for industrial diversification, separating them from many popular debates around industrial policy, that often focus on moonshot attempts to have places leapfrog up the value chain. Newcastle could, per this analysis, focus on its legal services industry. This may be not as exciting or headline-grabbing as trying to restart the semiconductor manufacturing on the Tyne[16] but is a much more viable option for an industrial policy interested in actually improving living standards. Sense-checking this result shows what we would hope to see too. Newcastle already has some promising legal services firms, including offices of national firms like Womble Bond Dickinson.

These opportunities, by the nature of the underlying complexity analysis, are often themselves related. Manchester has opportunities in computer programming, data processing, specialised design activities and other information service activities.[17] The same is true in Liverpool where there are a series of opportu-

15. The choice of RCA and d threshold is somewhat arbitrary, and any real-world policy application of this approach would require testing different calibrations.

16. There once was a nascent semiconductor manufacturing industry on Tyneside, which started in the 1990s as US and Japanese firms sought to broaden their footprint in Europe. While the sector struggled for much of the period, the financial crisis was the final deathknell. Aptly, the equipment from one of the last plants to close, the US-owned Atmel Fab 9, was sold to Taiwan Semiconductor Manufacturing Company (TSMC): https://www.ft.com/content/bf3ec360-769d-11dc-ad83-0000779fd2ac

17. Of course, some discretion is required when interpreting the model. Manchester, for example, has an opportunity in sea and coastal freight transport, despite not being on the coast (there was actually a Port of Manchester until 1982, but it was on the Manchester Ship

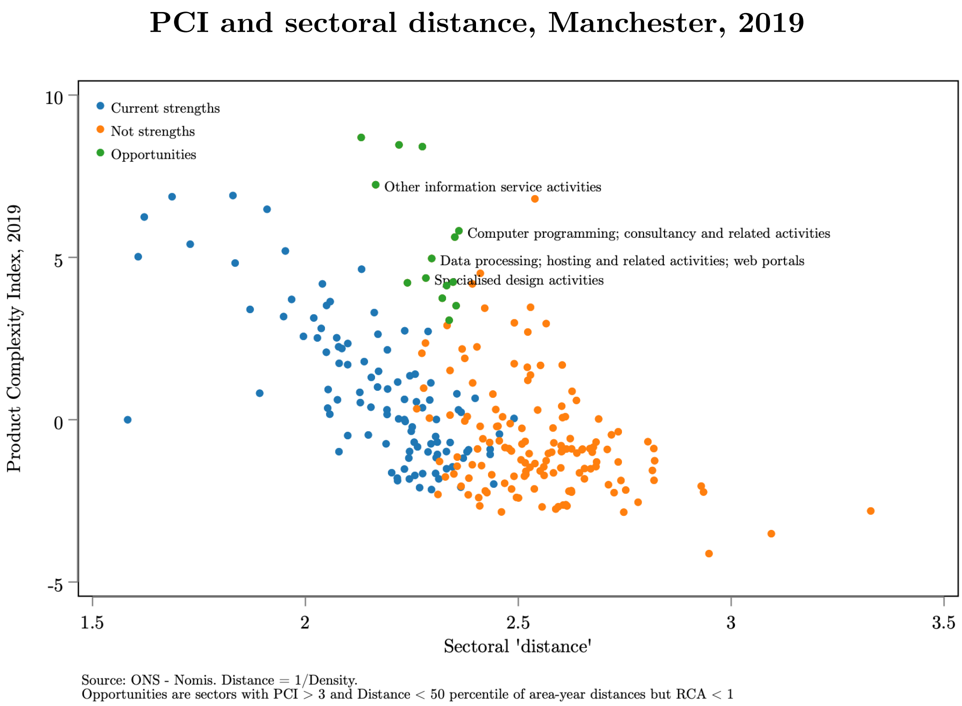


Figure 24: Manchester could diversify into new data and computing industries.

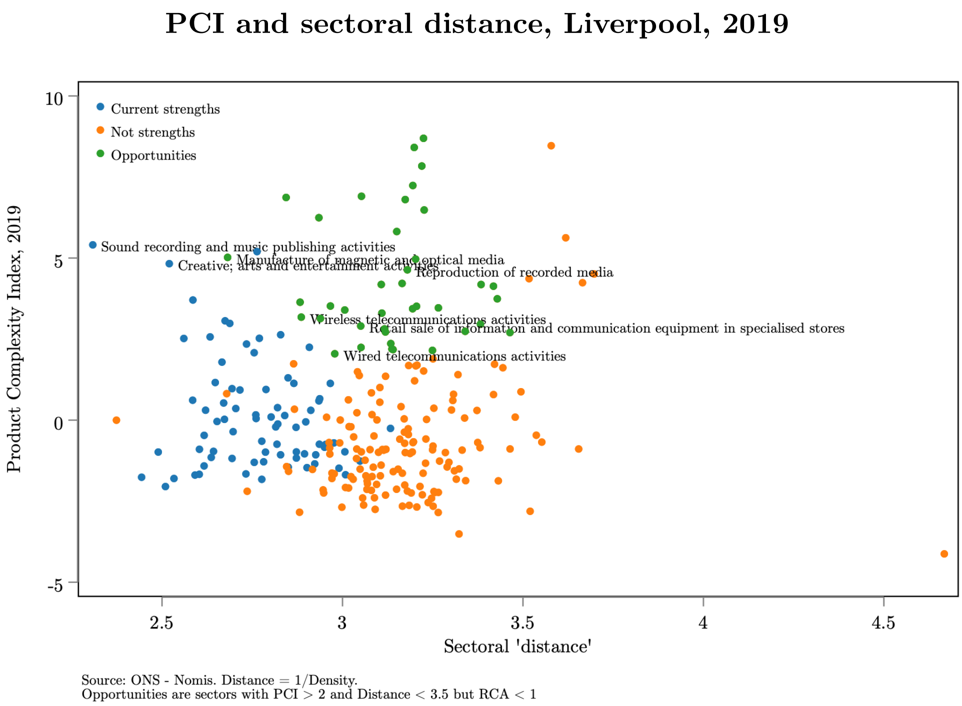


Figure 25: Liverpool can build on its existing strength in the creative arts.

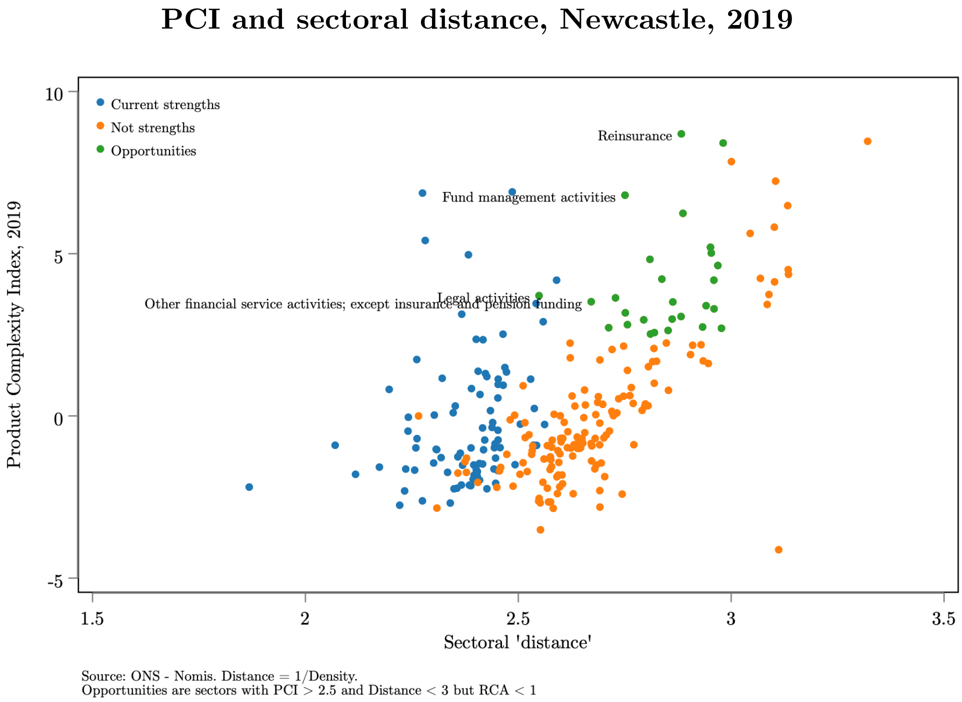


Figure 26: Newcastle could become a northern hub for financial and legal services.

nities in wireless and wired telecoms. There is also a related retail opportunity, in the sale of information and communication equipment. In Liverpool, opportunities relate to existing strengths. Liverpool has a strong history in music and the creative arts. This can be seen in the labelled blue plots in the Figure 25, in sound recording and music publishing, and arts and entertainment activities.

Because the opportunities are functions of the specific industrial composition of a place, they vary significantly from place to place. In Newcastle, there are opportunities in financial services, specifically, reinsurance, and fund management. This demonstrates that industrial policy cannot be an exclusively national operation. If one were just to look at the national level, even using this method, the prescriptions would not be appropriate for many of the places in need of development. This can be seem by looking at the Observatory for Economic Complexity's national export-generated measures. The relatedness measure recommends the UK focus on a top 10 sectors that include; inorganic chemicals, vaccines, blood, antisera, toxins, medical instruments and railway maintenance vehicles (*Observatory of Economic Complexity (OEC)* 2024). These should obviously be part of a national industrial strategy (not least given one of the UK's

Canal in nearby Salford, which fell into decline after the growth of containerisation.)

standout global export strengths is in pharmaceuticals), but exclusively pursuing such a strategy, with a place-based lens, would leave huge potential gains unrealised and do little to close regional divides.

Once opportunity sectors are identified, further analysis can be undertaken to identify what kind of support is needed to help diversification. As mentioned above, tools like this can help different places coordinate their strategies as well. Liverpool and Manchester will almost certainly have overlapping opportunities and they should look to combine efforts when such opportunities are identified. There may be specific sectoral frictions hampering progress, like a skills-shortage, or a 'horizontal' barrier, like inadequate transport infrastructure. Resource and capacity constraints, and the nature of the particular frictions will naturally limit how many sectors are ultimately pursued, but the approach can usefully help guide the process.

Economic complexity gives us two more tools that can inform the process of industrial policy formation; the Complexity Outlook Index (COI), and Opportunity Outlook Gain (OOG). The COI is a composite measure, defined at the place level, capturing how many complex industries are near a place's existing sectoral mix. COI has been found to be positively associated with ECI over 5 and 10 year horizons (Hausmann and Hidalgo 2014). Because of this, OOG can be thought of as a forecast measure, indicating how easy (or hard) a place is likely to find moving into more complex sectors. This is a useful indicator because, like the 'shape of the economy' plots in Figure 23, they give a sense of the relative resource level needed to drive progress in a place.

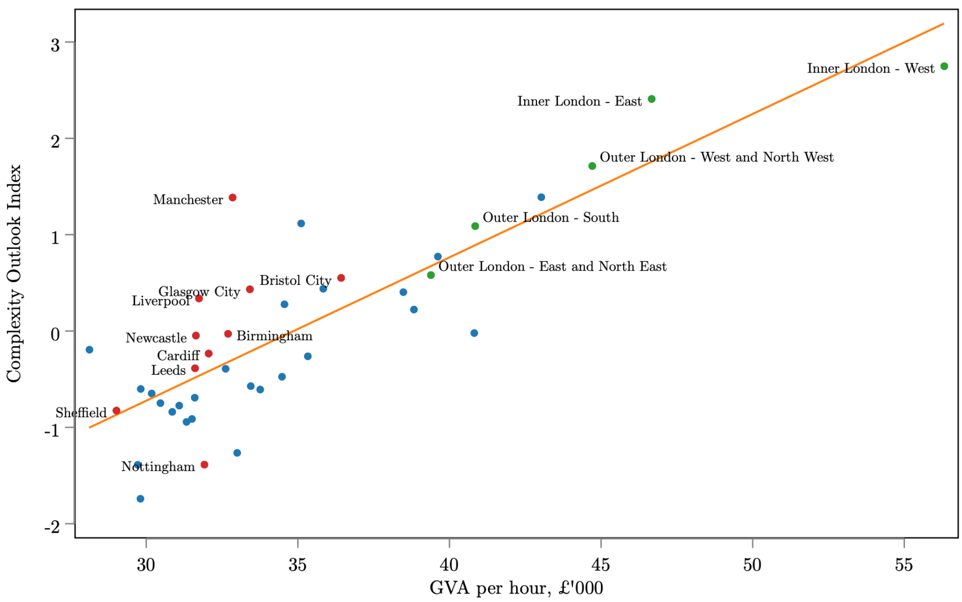
I compute COI using the distance, d_{cp}, measure outlined above, normalised so that $d_{cp} \in [0, 1]$. $1 - d_{cp}$, call it 'closeness', is then summed over all of the products, and weighted by the complexity of the sectors, such that:

$$\mathrm{COI}_c = \sum_p (1 - d_{cp})\,(1 - M_{cp})\,\mathrm{PCI}_p, \tag{17}$$

where PCI is the PCI of sector p and $1 - M_{cp}$ ensures only sectors where the place is not currently competitive are included.

Figure 27 shows the expected pattern, that places with higher current productivity are more likely to move into more complex sectors. Success begets success. But the distribution of places either side of the conditional mean gives us useful information. Places above the line, like Manchester, have a lot of potential 'catch-up' growth. They are proximate to a lot of complex industries relative to their current productivity. This is consistent with the encouraging negative shape of its sectoral distribution in PCI-distance space, as per Figure 23. By contrast, Nottingham, while being as productive as many of its Core City peers, faces significant challenges in trying to improve the aggregate complexity of its economy.

GVA per hour and Complexity Outlook Index, Core Cities and London, 2019



Source: ONS - Nomis. London and Core Cities highlighted.

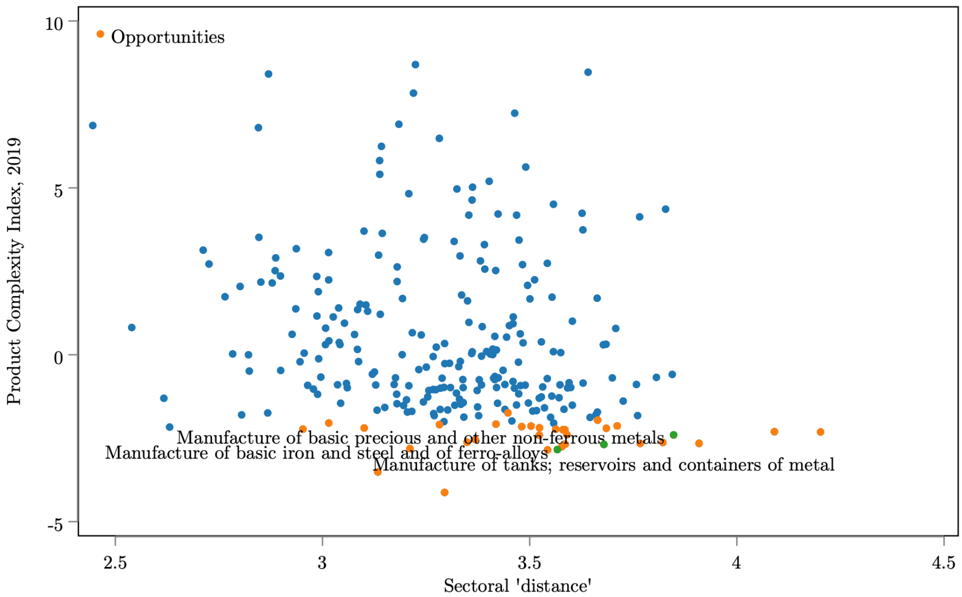
Figure 27: Complex places tend to have more opportunities in other complex industries.

Another way to think about economic opportunities is to look at the OOG. The OOG measures how much a place could benefit from future diversification in a particular product. If COI is the overall outlook of a place, the OOG is the gains from realising that outlook, in a given sector. Specifically, it quantifies how diversification into a sector could unlock adjacent products, also of higher complexity. OOG takes into account the complexity of products not currently being produced in a place and the distance to the existing capabilities a places possesses, i.e. existing sectoral strengths, and is given by:

$$\mathrm{OOG}_{cp} = \left[\sum_{p'} \frac{\Phi_{p,p'}}{\sum_{p''} \Phi_{p'',p'}} \left(1 - M_{cp'}\right) \mathrm{PCI}_{p'} \right], \tag{18}$$

where PCI is the PCI of product p' and $1 - M_{cp'}$ counts only products not currently being produced, i.e. those where the presence matrix $M_{cp} = 0$, as above for the COI. The higher the OOG, for a given product, implies that the product is closer to more products that are more complex.

PCI and sectoral distance, with Opportunity Outlook Gain, Glasgow, 2019



Source: ONS - Nomis. Opportunities are sectors with COG > .5.

Figure 28: The OOG measure allows places like Glasgow to identify opportunities that could unlock other complex sectors.

OOG can be thought of as an alternative way of identifying the opportunity sectors in Figures 24, 26, and 25. Whereas the opportunity measure from above will prioritise complex sectors in the short-term, the OOG usefully takes into account the dynamic effect diversification could have by unlocking connected opportunities. If we reproduce the PCI-distance plots, but now use the OOG to identify opportunities, similar inferences can be made. Figure 28 shows one such plot. As with those Figures, we see emergent patterns of related sectors, here in manufacturing basic precious and other non-ferrous metals, manufacturing iron and steel, and manufacturing of tanks, reservoirs and containers of metal. These sectors also illustrate the difference between OOG the and the original opportunity measure. OOG won't necessarily select the highest complexity industries first, instead it will select industries needed to unlock other complex industries. The predominance of basic, not particularly complex, manufacturing sectors in this measure for Glasgow shows this. One can imagine though that once Glasgow developed these capabilities, that could unlock more advanced manufacturing opportunities, that are currently too far from its existing capabilities and sectoral mix.

5 Conclusion

To follow

References

Agrawal, Sarthak, and David Phillips. 2020. *Catching up or falling behind? Geographical inequalities in the UK and how they have changed in recent years.* Technical report.

Andrews, Richard B. 1953. “Mechanics of the Urban Economic Base: Historical Development of the Base Concept.” *Land Economics* 29 (2). ISSN: 00237639. https://doi.org/10.2307/3144408.

Arthur, W. Brian. 2021. *Foundations of complexity economics,* 2. https://doi.org/10.1038/s42254-020-00273-3.

Axtell, Robert L., Omar A. Guerrero, and Eduardo López. 2019. “Frictional unemployment on labor flow networks.” *Journal of Economic Behavior and Organization* 160. ISSN: 01672681. https://doi.org/10.1016/j.jebo.2019.02.028.

Balassa, Bela. 1965. “Trade Liberalisation and “Revealed” Comparative Advantage.” *The Manchester School* 33 (2). ISSN: 14679957. https://doi.org/10.1111/j.1467-9957.1965.tb00050.x.

Balland, Pierre Alexandre, Cristian Jara-Figueroa, Sergio G. Petralia, Mathieu P.A. Steijn, David L. Rigby, and César A. Hidalgo. 2020. “Complex economic activities concentrate in large cities.” *Nature Human Behaviour* 4 (3). ISSN: 23973374. https://doi.org/10.1038/s41562-019-0803-3.

Balland, Pierre Alexandre, and David Rigby. 2017. “The Geography of Complex Knowledge.” *Economic Geography* 93 (1). ISSN: 19448287. https://doi.org/10.1080/00130095.2016.1205947.

Battiston, Stefano, J. Doyne Farmer, Andreas Flache, Diego Garlaschelli, Andrew G. Haldane, Hans Heesterbeek, Cars Hommes, Carlo Jaeger, Robert May, and Marten Scheffer. 2016. “Complexity theory and financial regulation.” *Science* 351 (6275). ISSN: 0036-8075. https://doi.org/10.1126/science.aad0299.

Beatty, Christina, and Steve Fothergill. 2020. "The Long Shadow of Job Loss: Britain's Older Industrial Towns in the 21st Century." *Frontiers in Sociology* 5. ISSN: 22977775. https://doi.org/10.3389/fsoc.2020.00054.

Blanchard, Olivier J., and Lawrence H. Summers. 1986. "Hysteresis and the European Unemployment Problem." *NBER Macroeconomics Annual* 1. ISSN: 0889-3365. https://doi.org/10.1086/654013.

Blanchard, Olivier Jean, and Lawrence F. Katz. 1992. *Regional evolutions,* 1. https://doi.org/10.2307/2534556.

Bown, C P. 2023. "Modern industrial policy and the WTO."

Breach, Anthony. 2020. *The future of the planning system in England.* Technical report. Centre for Cities.

Breach, Anthony, and Guilherme Rodrigues. 2021. *Measuring up: Comparing public transport in the UK and Europe's biggest cities.* Technical report. Centre for Cities.

Chávez, Juan Carlos, Marco T. Mosqueda, and Manuel Gómez-Zaldívar. 2017. "Economic complexity and regional growth performance: Evidence from the Mexican economy." *Review of Regional Studies* 47 (2). ISSN: 15530892. https://doi.org/10.52324/001c.8023.

Committee for a Responsible Federal Budget. 2022. *CBO Scores IRA with $238 Billion of Deficit Reduction,* July. https://www.crfb.org/blogs/cbo-scores-ira-238-billion-deficit-reduction.

Conwell, Lucas, Fabian Eckert, and Ahmed Mushfiq Mobarak. 2023. "More Roads or Public Transit? Insights from Measuring City-Center Accessibility." *SSRN Electronic Journal,* https://doi.org/10.2139/ssrn.4334956.

Coyle, Diane, and Musaddiq Adam Muhtar. 2021. "The UK's Industrial Policy: Learning from the Past." *SSRN Electronic Journal,* https://doi.org/10.2139/ssrn.3973039.

De Lyon, Josh, Raif Martin, Juliana Oliver-Cunha, Arjun Shah, Gregory Thwaites, and Anna Valero. 2022. *Enduring Strengths.* Technical report. London: Resolution Foundation.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What causes industry agglomeration? Evidence from coagglomeration patterns." *American Economic Review* 100 (3). ISSN: 00028282. https://doi.org/10.1257/aer.100.3.1195.

Evenett, Simon. 2024. "The Return of Industrial Policy in Data." *IMF Working Papers* 2024 (001). ISSN: 1018-5941. https://doi.org/10.5089/9798400260964.001.

Forth, Tom. 2019. *Birmingham is a small city.*

Fritz, Benedikt S.L., and Robert A. Manduca. 2021. "The economic complexity of US metropolitan areas." *Regional Studies* 55 (7). ISSN: 13600591. https://doi.org/10.1080/00343404.2021.1884215.

Gao, Jian, and Tao Zhou. 2018. "Quantifying China's regional economic complexity." *Physica A: Statistical Mechanics and its Applications* 492. ISSN: 03784371. https://doi.org/10.1016/j.physa.2017.11.084.

Geary, Frank, and Tom Stark. 2018. "150 years of regional GDP: United Kingdom and Ireland." In *The Economic Development of Europe's Regions: A Quantitative History since 1900.*

Giuliano, Genevieve, Sanggyun Kang, and Quan Yuan. 2019. *Agglomeration economies and evolving urban form,* 3. https://doi.org/10.1007/s00168-019-00957-4.

Glaeser, Edward. 2010. *Agglomeration Economics.* Chicago, IL: University of Chicago Press.

Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman, and Andrei Shleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100 (6). ISSN: 0022-3808. https://doi.org/10.1086/261856.

Glaeser, Edward L., and William R. Kerr. 2009. "Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain?" In *Journal of Economics and Management Strategy,* vol. 18. 3. https://doi.org/10.1111/j.1530-9134.2009.00225.x.

Grossman, Gene M., and Ezra Oberfield. 2022. *The Elusive Explanation for the Declining Labor Share.* https://doi.org/10.1146/annurev-economics-080921-103046.

Hausmann, Ricardo, and Cesar A. Hidalgo. 2014. *The Atlas of Economic Complexity: Mapping Paths to Prosperity.* 2nd ed. Cambridge, MA: MIT Press.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. 2007. "What you export matters." *Journal of Economic Growth* 12 (1). ISSN: 1381-4338. https://doi.org/10.1007/s10887-006-9009-4.

Hausmann, Ricardo, and Dani Rodrik. 2003. "Economic development as self-discovery." *Journal of Development Economics* 72 (2). ISSN: 03043878. https://doi.org/10.1016/S0304-3878(03)00124-X.

Heilburn, James. 1981. "The Urban Economic Base and Economic Policy." Chap. 7 in *Urban Economics and Public Policy,* 153–169. New York, NY: St Martin's Press.

Hidalgo, C. A., B. Winger, A. L. Barabási, and R. Hausmann. 2007. "The product space conditions the development of nations." *Science* 317 (5837). ISSN: 00368075. https://doi.org/10.1126/science.1144581.

Hidalgo, César A. 2021. *Economic complexity theory and applications,* 2. https://doi.org/10.1038/s42254-020-00275-1.

———. 2023. "The policy implications of economic complexity." *Research Policy* 52 (9). ISSN: 00487333. https://doi.org/10.1016/j.respol.2023.104863.

Hidalgo, César A., and Ricardo Hausmann. 2009. "The building blocks of economic complexity." *Proceedings of the National Academy of Sciences of the United States of America* 106 (26). ISSN: 00278424. https://doi.org/10.1073/pnas.0900943106.

HM Government. 2017. *Industrial Strategy: Building a Britain fit for the future.* Technical report. June.

———. 2021. *Build Back Better: Our plan for growth.* Technical report. March.

HM Government. 2022. *Levelling Up the United Kingdom.* Technical report. London: Secretary of State for Levelling Up, Housing and Communities.

Hommes, Cars. 2021. *Behavioral and experimental macroeconomics and policy analysis: A complex systems approach,* 1. https://doi.org/10.1257/jel.20191434.

Hutton, Georgina, and Chris Rhodes. 2021. *UK steel industry: statistics and policy.* Technical report. London: House of Commons Library.

Jones, Charles I., and John C. Williams. 2000. "Too much of a good thing? the economics of investment in R&D." *Journal of Economic Growth* 5 (1). ISSN: 13814338. https://doi.org/10.1023/A:1009826304308.

Juhász, Réka, Nathaniel Lane, and Dani Rodrik. 2023. "The New Economics of Industrial Policy." *SSRN Electronic Journal,* https://doi.org/10.2139/ssrn.4542861.

Kline, Patrick, and Enrico Moretti. 2014. "People, places, and public policy: Some simple welfare economics of local economic development programs." *Annual Review of Economics* 6. ISSN: 19411391. https://doi.org/10.1146/annurev-economics-080213-041024.

Krugman, P. 1992. *Geography and Trade.* Cambridge, MA: The MIT Press.

———. 1991. "Increasing returns and economic geography." *Journal of Political Economy* 99 (3). ISSN: 00223808. https://doi.org/10.1086/261763.

LeBaron, Blake, and Leigh Tesfatsion. 2008. "Modeling macroeconomies as open-ended dynamic systems of interacting agents." In *American Economic Review,* vol. 98. 2. https://doi.org/10.1257/aer.98.2.246.

Lindgren, Kristian. 1997. "Evolutionary Dynamics in Game-Theoretic Models." In *The Economy as an Evolving Complex System II,* edited by W. Brian Arthur, Steven N. Durlauf, and David A. Lane, 337–367. Reading: Addison-Wesley.

Martin, Josh, and Michael Becker. 2023. "New insights on regional capital investment in the UK, 1997 to 2019."

Martin, R. 1988. “The political economy of Britain’s north-south divide.” *Transactions - Institute of British Geographers* 13 (4). ISSN: 00202754. https://doi.org/10.2307/622738.

Martin, Ron, Ben Gardiner, Andy Pike, Peter Sunley, and Peter Tyler. 2021. *Levelling up Left Behind Places.* https://doi.org/10.4324/9781032244341.

Mealy, Penny, and Diane Coyle. 2022. “To them that hath: economic complexity and local industrial strategy in the UK.” *International Tax and Public Finance* 29 (2). ISSN: 15736970. https://doi.org/10.1007/s10797-021-09667-0.

Mealy, Penny, J. Doyne Farmer, and Alexander Teytelboym. 2019. “Interpreting economic complexity.” *Science Advances* 14 (2). ISSN: 23752548. https://doi.org/10.1126/sciadv.aau1705.

Moretti, Enrico. 2010. “Local multipliers.” In *American Economic Review,* vol. 100. 2. https://doi.org/10.1257/aer.100.2.373.

Nathan, Max, and Henry Overman. 2013. “Agglomeration, clusters, and industrial policy.” *Oxford Review of Economic Policy* 29 (2). ISSN: 0266903X. https://doi.org/10.1093/oxrep/grt019.

Observatory of Economic Complexity (OEC). 2024.

Ohlin, B, and E Heckscher. 1991. *Heckscher-Ohlin Trade Theory.* Edited by J Flanders and H Flam. Cambridge, MA: MIT Press.

ONS. 2022. *Government Statistical Service (GSS) subnational data strategy.*

Overman, Henry G., and Diego Puga. 2008. “Labour pooling as a source of agglomeration: an empirical investigation.” *SERC Discussion Paper 6.*

Palan, Nicole. 2010. “Measurement of Specialization -The Choice of Indices.” *FIW – Working Paper* 62 (December).

Pérez Pérez, Jorge, Felipe Vial Lecaros, and Román D Zárate. 2022. “Urban Transit Infrastructure: Spatial Mismatch and Labor Market Power.” *Nber.*

Perroux, François. 1955. “Note sur la notion de ”pôle de croissance”.” *Économie appliquée* 8 (1). ISSN: 0013-0494. https://doi.org/10.3406/ecoap.1955.2522.

Porter, M. E. 1998. "Clusters and the new economics of competition." *Harvard business review* 76 (6). ISSN: 00178012.

Porter, Michael E. 2003. "The economic performance of regions." *Regional Studies* 37 (6-7). ISSN: 00343404. https://doi.org/10.1080/0034340032000108688.

Reeves, Rachel. 2023. *Securonomics,* April. https://labour.org.uk/updates/press-releases/rachel-reeves-securonomics/.

Reynolds, Christian, Manju Agrawal, Ivan Lee, Chen Zhan, Jiuyong Li, Phillip Taylor, Tim Mares, Julian Morison, Nicholas Angelakis, and Göran Roos. 2018. "A sub-national economic complexity analysis of Australia's states and territories." *Regional Studies* 52 (5). ISSN: 13600591. https://doi.org/10.1080/00343404.2017.1283012.

Rice, Patricia G., and Anthony J. Venables. 2021. "The persistent consequences of adverse shocks: How the 1970s shaped UK regional inequality." *Oxford Review of Economic Policy* 37 (1). ISSN: 14602121. https://doi.org/10.1093/oxrep/graa057.

Rodrik, Dani. 2009. "Industrial Policy: Don't Ask Why, Ask How." *Middle East Development Journal* 1 (1). ISSN: 1793-8120. https://doi.org/10.1142/s1793812009000024.

Rodrik, Dani, and Charles Frederick Sabel. 2020. "Building a Good Jobs Economy." *SSRN Electronic Journal,* https://doi.org/10.2139/ssrn.3533430.

Rolls-Royce. 2024. *2023 Full Year Results,* February. https://www.rolls-royce.com/investors/2023-full-year-results.aspx#section-overview.

Romer, Paul M. 1994. "The Origins of Endogenous Growth." *Journal of Economic Perspectives* 8 (1). ISSN: 0895-3309. https://doi.org/10.1257/jep.8.1.3.

Rosser, J. Barkley. 1999. *On the Complexities of Complex Economic Dynamics,* 4. https://doi.org/10.1257/jep.13.4.169.

Saito, Hisamitsu, and Munisamy Gopinath. 2009. "Plants' self-selection, agglomeration economies and regional productivity in Chile." *Journal of Economic Geography* 9 (4). ISSN: 14682702. https://doi.org/10.1093/jeg/lbp010.

Saito, Hisamitsu, and Matsuura Toshiyuki. 2016. "Agglomeration Economies, Productivity, and Quality Upgrading Agglomeration Economies, Productivity, and Quality Upgrading *." *RIETI Discussion Paper Series.*

Sandford, M. 2023. *English devolution deals in the 2023 Autumn Statement.* https://commonslibrary.parliament.uk/english-devolution-deals-in-the-2023-autumn-statement/.

Simon, Herbert A. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106 (6): 467–482. ISSN: 0003049X. http://www.jstor.org/stable/985254.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations.* London: W. Strahan / T. Cadell.

Solow, Robert M. 1956. "A contribution to the theory of economic growth." *Quarterly Journal of Economics* 70 (1). ISSN: 15314650. https://doi.org/10.2307/1884513.

———. 1957. "Technical Change and the Aggregate Production Function." *The Review of Economics and Statistics* 39 (3). ISSN: 00346535. https://doi.org/10.2307/1926047.

Stansbury, Anna, Dan Turner, and Ed Balls. 2023. "Tackling the UK's regional economic inequality: binding constraints and avenues for policy intervention." *Contemporary Social Science* 18 (3-4). ISSN: 2158205X. https://doi.org/10.1080/21582041.2023.2250745.

Sweney, Mark. 2023. *Tata Steel seals £500m UK support package but big job losses feared,* September. https://www.theguardian.com/business/2023/sep/15/tata-steel-seals-500m-uk-support-package-but-big-job-losses-feared.

Tacchella, Andrea, Matthieu Cristelli, Guido Caldarelli, Andrea Gabrielli, and Luciano Pietronero. 2012. "A new metrics for countries' fitness and products' complexity." *Scientific Reports* 2. ISSN: 20452322. https://doi.org/10.1038/srep00723.

Tomlinson, Jim. 2019. "The Rise and Fall of the British Nation. A Twentieth Century History. By David Edgerton." *Twentieth Century British History* 30 (4). ISSN: 0955-2359. https://doi.org/10.1093/tcbh/hwy052.

Weber, Isabella, Gregor Semieniuk, Tom Westland, and Junshang Liang. 2021. "What You Expor ou Exported Matters: P ted Matters: Persistence in Pr ersistence in Productiv oductive Capabilities across Two Eras of Globalization."

Youn, Hyejin, Luis M.A. Bettencourt, José Lobo, Deborah Strumsky, Horacio Samaniego, and Geoffrey B. West. 2016. "Scaling and universality in urban economic diversification." *Journal of the Royal Society Interface* 13 (114). ISSN: 17425662. https://doi.org/10.1098/rsif.2015.0937.

Appendix

5.1 List of SIC 3 industries

Industry name	SIC2007 Group
DEFRA/Scottish Executive Agricultural Data	10
Support activities to agriculture and post-harv...	16
Hunting, trapping and related service activities	17
Silviculture and other forestry activities	21
Logging	22
Gathering of wild growing non-wood products	23
Support services to forestry	24
Fishing	31
Aquaculture	32
Mining of hard coal	51
Extraction of crude petroleum	61
Extraction of natural gas	62
Quarrying of stone, sand and clay	81
Mining and quarrying n.e.c.	89
Support activities for petroleum and natural ga...	91
Support activities for other mining and quarrying	99
Processing and preserving of meat and productio...	101
Processing and preserving of fish, crustaceans ...	102
Processing and preserving of fruit and vegetables	103
Manufacture of vegetable and animal oils and fats	104
Manufacture of dairy products	105
Manufacture of grain mill products, starches an...	106

Continued on next page

Industry name	SIC2007 Group
Manufacture of bakery and farinaceous products	107
Manufacture of other food products	108
Manufacture of prepared animal feeds	109
Manufacture of beverages	110
Manufacture of tobacco products	120
Preparation and spinning of textile fibres	131
Weaving of textiles	132
Finishing of textiles	133
Manufacture of other textiles	139
Manufacture of wearing apparel, except fur apparel	141
Manufacture of articles of fur	142
Manufacture of knitted and crocheted apparel	143
Tanning and dressing of leather; manufacture of...	151
Manufacture of footwear	152
Sawmilling and planing of wood	161
Manufacture of products of wood, cork, straw an...	162
Manufacture of pulp, paper and paperboard	171
Manufacture of articles of paper and paperboard	172
Printing and service activities related to prin...	181
Reproduction of recorded media	182
Manufacture of coke oven products	191
Manufacture of refined petroleum products	192
Manufacture of basic chemicals, fertilisers and...	201
Manufacture of pesticides and other agrochemica...	202

Continued on next page

Industry name	SIC2007 Group
Manufacture of paints, varnishes and similar co...	203
Manufacture of soap and detergents, cleaning an...	204
Manufacture of other chemical products	205
Manufacture of man-made fibres	206
Manufacture of basic pharmaceutical products	211
Manufacture of pharmaceutical preparations	212
Manufacture of rubber products	221
Manufacture of plastics products	222
Manufacture of glass and glass products	231
Manufacture of refractory products	232
Manufacture of clay building materials	233
Manufacture of other porcelain and ceramic prod...	234
Manufacture of cement, lime and plaster	235
Manufacture of articles of concrete, cement and...	236
Cutting, shaping and finishing of stone	237
Manufacture of abrasive products and non-metall...	239
Manufacture of basic iron and steel and of ferr...	241
Manufacture of tubes, pipes, hollow profiles an...	242
Manufacture of other products of first processi...	243
Manufacture of basic precious and other non-fer...	244
Casting of metals	245
Manufacture of structural metal products	251
Manufacture of tanks, reservoirs and containers...	252
Manufacture of steam generators, except central...	253

Continued on next page

Industry name	SIC2007 Group
Manufacture of weapons and ammunition	254
Forging, pressing, stamping and roll-forming of...	255
Treatment and coating of metals; machining	256
Manufacture of cutlery, tools and general hardware	257
Manufacture of other fabricated metal products	259
Manufacture of electronic components and boards	261
Manufacture of computers and peripheral equipment	262
Manufacture of communication equipment	263
Manufacture of consumer electronics	264
Manufacture of instruments and appliances for m...	265
Manufacture of irradiation, electromedical and ...	266
Manufacture of optical instruments and photogra...	267
Manufacture of magnetic and optical media	268
Manufacture of electric motors, generators, tra...	271
Manufacture of batteries and accumulators	272
Manufacture of wiring and wiring devices	273
Manufacture of electric lighting equipment	274
Manufacture of domestic appliances	275
Manufacture of other electrical equipment	279
Manufacture of general purpose machinery	281
Manufacture of other general-purpose machinery	282
Manufacture of agricultural and forestry machinery	283
Manufacture of metal forming machinery and mach...	284
Manufacture of other special-purpose machinery	289

Continued on next page

Industry name	SIC2007 Group
Manufacture of motor vehicles	291
Manufacture of bodies (coachwork) for motor veh...	292
Manufacture of parts and accessories for motor ...	293
Building of ships and boats	301
Manufacture of railway locomotives and rolling ...	302
Manufacture of air and spacecraft and related m...	303
Manufacture of military fighting vehicles	304
Manufacture of transport equipment n.e.c.	309
Manufacture of furniture	310
Manufacture of jewellery, bijouterie and relate...	321
Manufacture of musical instruments	322
Manufacture of sports goods	323
Manufacture of games and toys	324
Manufacture of medical and dental instruments a...	325
Other manufacturing	329
Repair of fabricated metal products, machinery ...	331
Installation of industrial machinery and equipment	332
Electric power generation, transmission and dis...	351
Manufacture of gas; distribution of gaseous fue...	352
Steam and air conditioning supply	353
Water collection, treatment and supply	360
Sewerage	370
Waste collection	381
Waste treatment and disposal	382

Continued on next page

Industry name	SIC2007 Group
Materials recovery	383
Remediation activities and other waste manageme...	390
Development of building projects	411
Construction of residential and non-residential...	412
Construction of roads and railways	421
Construction of utility projects	422
Construction of other civil engineering projects	429
Demolition and site preparation	431
Electrical, plumbing and other construction ins...	432
Building completion and finishing	433
Other specialised construction activities n.e.c.	439
Sale of motor vehicles	451
Maintenance and repair of motor vehicles	452
Sale of motor vehicle parts and accessories	453
Sale, maintenance and repair of motorcycles and...	454
Wholesale on a fee or contract basis	461
Wholesale of agricultural raw materials and liv...	462
Wholesale of food, beverages and tobacco	463
Wholesale of household goods	464
Wholesale of information and communication equi...	465
Wholesale of other machinery, equipment and sup...	466
Other specialised wholesale	467
Non-specialised wholesale trade	469
Retail sale in non-specialised stores	471

Continued on next page

Industry name	SIC2007 Group
Retail sale of food, beverages and tobacco in s...	472
Retail sale of automotive fuel in specialised s...	473
Retail sale of information and communication eq...	474
Retail sale of other household equipment in spe...	475
Retail sale of cultural and recreation goods in...	476
Retail sale of other goods in specialised stores	477
Retail sale via stalls and markets	478
Retail trade not in stores, stalls or markets	479
Passenger rail transport, interurban	491
Freight rail transport	492
Other passenger land transport	493
Freight transport by road and removal services	494
Transport via pipeline	495
Sea and coastal passenger water transport	501
Sea and coastal freight water transport	502
Inland passenger water transport	503
Inland freight water transport	504
Passenger air transport	511
Freight air transport and space transport	512
Warehousing and storage	521
Support activities for transportation	522
Postal activities under universal service oblig...	531
Other postal and courier activities	532
Hotels and similar accommodation	551

Continued on next page

Industry name	SIC2007 Group
Holiday and other short stay accommodation	552
Camping grounds, recreational vehicle parks and...	553
Other accommodation	559
Restaurants and mobile food service activities	561
Event catering and other food service activities	562
Beverage serving activities	563
Publishing of books, periodicals and other publ...	581
Software publishing	582
Motion picture, video and television programme ...	591
Sound recording and music publishing activities	592
Radio broadcasting	601
Television programming and broadcasting activities	602
Wired telecommunications activities	611
Wireless telecommunications activities	612
Satellite telecommunications activities	613
Other telecommunications activities	619
Computer programming, consultancy and related a...	620
Data processing, hosting and related activities...	631
Other information service activities	639
Monetary intermediation	641
Activities of holding companies	642
Trusts, funds and similar financial entities	643
Other financial service activities, except insu...	649
Insurance	651

Continued on next page

Industry name	SIC2007 Group
Reinsurance	652
Activities auxiliary to financial services, exc...	661
Activities auxiliary to insurance and pension f...	662
Fund management activities	663
Buying and selling of own real estate	681
Renting and operating of own or leased real estate	682
Real estate activities on a fee or contract basis	683
Legal activities	691
Accounting, bookkeeping and auditing activities...	692
Activities of head offices	701
Management consultancy activities	702
Architectural and engineering activities and re...	711
Technical testing and analysis	712
Research and experimental development on natura...	721
Research and experimental development on social...	722
Advertising	731
Market research and public opinion polling	732
Specialised design activities	741
Photographic activities	742
Translation and interpretation activities	743
Other professional, scientific and technical ac...	749
Veterinary activities	750
Renting and leasing of motor vehicles	771
Renting and leasing of personal and household g...	772

Continued on next page

Industry name	SIC2007 Group
Renting and leasing of other machinery, equipme...	773
Leasing of intellectual property and similar pr...	774
Activities of employment placement agencies	781
Temporary employment agency activities	782
Other human resources provision	783
Travel agency and tour operator activities	791
Other reservation service and related activities	799
Private security activities	801
Security systems service activities	802
Investigation activities	803
Combined facilities support activities	811
Cleaning activities	812
Landscape service activities	813
Office administrative and support activities	821
Activities of call centres	822
Organisation of conventions and trade shows	823
Business support service activities n.e.c.	829
Administration of the State and the economic an...	841
Provision of services to the community as a whole	842
Compulsory social security activities	843
Pre-primary education	851
Primary education	852
Secondary education	853
Higher education	854

Continued on next page

Industry name	SIC2007 Group
Other education	855
Educational support activities	856
Hospital activities	861
Medical and dental practice activities	862
Other human health activities	869
Residential nursing care activities	871
Residential care activities for learning disabi...	872
Residential care activities for the elderly and...	873
Other residential care activities	879
Social work activities without accommodation fo...	881
Other social work activities without accommodation	889
Creative, arts and entertainment activities	900
Libraries, archives, museums and other cultural...	910
Gambling and betting activities	920
Sports activities	931
Amusement and recreation activities	932
Activities of business, employers and professio...	941
Activities of trade unions	942
Activities of other membership organisations	949
Repair of computers and communication equipment	951
Repair of personal and household goods	952
Other personal service activities	960

Table 5: Cross-section regressions

	ln GVA per hour						ln GDHI per person					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ECI Emp	0.109***		0.0338***		0.0995***	0.0302	0.166***		0.0581***		0.149***	0.0425
	(0.00638)		(0.00755)		(0.0117)	(0.0186)	(0.0142)		(0.0171)		(0.0255)	(0.0398)
ECI Exports		0.0523***		0.0101	0.0263***	0.00944		0.0690***		0.00178	0.0299**	0.000804
		(0.00999)		(0.0106)	(0.00990)	(0.0114)		(0.0157)		(0.0195)	(0.0120)	(0.0187)
ln GFCF	0.0511***	0.134***	0.0635***	0.0853***	0.0304	0.0620**	0.0173	0.161***	0.0331**	0.0894**	0.00453	0.0567
	(0.0133)	(0.0262)	(0.0103)	(0.0170)	(0.0226)	(0.0256)	(0.0158)	(0.0517)	(0.0140)	(0.0336)	(0.0271)	(0.0365)
Sh Tertiary			0.0101***	0.0113***		0.00938***			0.0153***	0.0177***		0.0150***
			(0.000732)	(0.00124)		(0.00167)			(0.00183)	(0.00379)		(0.00439)
Constant	3.039***	2.349***	2.482***	2.275***	3.282***	2.568***	9.729***	8.489***	8.930***	8.349***	9.889***	8.760***
	(0.119)	(0.234)	(0.101)	(0.148)	(0.205)	(0.272)	(0.139)	(0.455)	(0.172)	(0.354)	(0.239)	(0.446)
Observations	240	80	195	39	80	39	240	80	195	39	80	39
R-squared	0.694	0.549	0.831	0.841	0.750	0.852	0.636	0.424	0.760	0.733	0.651	0.743

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses.

Table 6: Panel regressions

	ln GVA per hour						ln GDHI per person					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ECI Emp	0.0776***		0.0994***	0.119***		0.113***	0.0716***		0.0468**	0.189***		0.179***
	(0.0112)		(0.0159)	(0.00659)		(0.0132)	(0.0229)		(0.0192)	(0.0146)		(0.0276)
ECI Exports		0.0189*	0.0123		0.0519***	0.0183*		−0.00202	−0.00310		0.0694***	0.0161
		(0.0102)	(0.00912)		(0.0109)	(0.0106)		(0.00420)	(0.00433)		(0.0187)	(0.0133)
ln GFCF	0.100***	0.0342	−0.00234	0.0535***	0.153***	0.0406**	0.117***	0.0317**	0.0354**	0.0372*	0.205***	0.0256
	(0.0210)	(0.0324)	(0.0198)	(0.0125)	(0.0309)	(0.0180)	(0.0219)	(0.0157)	(0.0172)	(0.0214)	(0.0665)	(0.0360)
Constant	2.601***	3.248***	3.575***	3.005***	2.166***	3.176***	8.835***	9.645***	9.612***	9.544***	8.092***	9.694***
	(0.190)	(0.281)	(0.179)	(0.112)	(0.279)	(0.163)	(0.192)	(0.123)	(0.142)	(0.188)	(0.593)	(0.319)
Observations	240	80	80	198	66	66	240	80	80	198	66	66
R^2 within	0.064	0.011	0.006				0.093	0.021	0.000			
R^2 between	0.740	0.571	0.748				0.573	0.295	0.600			
R^2 overall	0.670	0.5413	0.736				0.554	0.290	0.596			
R^2 adjusted				0.788	0.592	0.826				0.720	0.414	0.712
Urban FE	NO	NO	NO	YES	YES	YES	NO	NO	NO	YES	YES	YES

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Robust standard errors in parentheses.