Using data for local economic policy

Sept 2019

what works centre for
local economic growth

00

Introduction

Open data offers many opportunities for better evidence-based policy making. This report illustrates how local industrial strategies can be informed by the use of new data sources to analyse sectors, links between firms and flows of knowledge. In data analysis, like in policy evaluation, risks can be anticipated and mitigated by starting early, bringing everyone on board and replicating examples of best practice. This guide contains case studies on a range of data initiatives that provide examples of best practice and highlight risks that can arise during implementation.

How to use this guide

This guide discusses ways in which data can be used to aid the development of effective local economic policy. We have used our 10 Principles for Developing a Local Industrial Strategy as a framework, but the advice applies to any Economic Development strategy. This guide is structured in order to support those local economic practitioners who have no or little experience of the data available, as well as those with a more advanced knowledge of using data and statistical analysis.

To complement the guide, case studies describing how UK local and combined authorities have applied these principles are referenced throughout, and detailed in the annex.

Section one

How to retrieve and use the data

01 What data is available and where to get it

02 Choosing the data to use – some of the risks attached to new and experimental datasets

03 Using standard industrial classifications – what to be aware of

04 Analysing the data – what to think about

Section two

Applying the data to economic strategy

This section looks in more detail at how to apply some of the secondary and primary datasets and mechanisms for their analysis to more specific elements of most economic strategies. It addresses the following three areas:

Case studies

01

What data is available and where can you get it?

Secondary datasets

How do we get data about a local economy? Most practitioners will be aware of some of the information available to them, particularly through governmental or administrative datasets. The examples throughout this guide however demonstrate that much of the information local areas need may be available through different datasets, often held by private organisations. These datasets will not always be free to use, but will be a fraction of the cost when compared to collecting primary data.

Available secondary data sources

Public Sector Administrative Data

Administrative data can either come from central government or from towns and cities themselves. These datasets are increasingly available in raw form (the 'micro-data') and can often be linked together. Furthermore, they are generally free to use for policymakers and academic researchers and 'administrative big data' on firms is already widely used by academics.

Commercial/Privately held data

Business-built datasets are available for free or low cost to researchers and government, typically through their application-programming interface (API).

APIs provide mechanisms for customers to access and manipulate data stored by the API provider. The user makes a "request" to a webserver, and that webserver accesses a database, with customer data, and returns it to the requester in a "response".

Some data requests can be provided as 'raw' output from sites like Flickr, Twitter or Yelp; some data can be packaged up, as with Google Trends.

Paid-for offerings often use machine learning to process raw data, for example mapping firms or industries as in case studies 2 and 5.

Primary datasets and experimental datasets

This will most likely be data about companies or individuals scraped from the web. Some of this may have been pre-processed by commercial organisations but would still be experimental in its nature.

Practitioners may also require bespoke information that is not otherwise available in an existing dataset. In those scenarios, an option would be to collect the data themselves; however, this should be approached with caution as collecting primary data is both costly and time consuming.

Tips for building primary datasets

When resorting to primary data, be realistic about what can be achieved. Simple surveys (e.g. of specific types of firms) may be relatively easy to implement. Complicated surveys (e.g. to map supply chains) are a much bigger challenge. On the basis of our case studies, and our discussions with local areas, we would urge caution in taking forward this type of analysis.

The best choice of dataset will depend on the question. Often analysis will require information from more than one dataset so it is crucial to collect unique identifiers that will allow datasets to be merged.

- For firms, Company Record Numbers (CRNs) allow matching to information in Companies House, administrative data resources like the Business Structure Database (BSD) and to many commercial data sets. It is even better if this data is supplemented by accurate address and postcode information.

- For individuals, it is harder to systematically gather data by unique customer IDs. Doing so would be a big step forward, although not without its challenges in terms of privacy. The rewards are large however, as linked datasets will often open up opportunities for other projects and lead to further opportunities in the medium term.

Examples of secondary datasets useful for economic policy

Below are suggestions for datasets that can help to enhance local understanding of the industrial structure, the links between firms, and the strengths and weaknesses in an economy with regards to employment and productivity.

Secondary administrative dataset: Business Structure Database (BSD) and Inter-Departmental Business Register (IDBR)

ONS's Inter-Departmental Business Register (IDBR) is a comprehensive list of firms used by the Government for statistical purposes. It covers 2.6 million VAT- or PAYE-registered firms in all sectors of the UK economy. The BSD is a point-in-time snapshot of the IDBR made available to researchers.

For each of the 2.6 million firms, the IDBR holds information on:

- Turnover
- Employment
- Standard Industrial Classification (SIC) (from 2007)

The data is available from 1997 at postcode level and Lower Layer Super Output Area (LSOA) level – a statistical geography with approximately 1,500 residents in each area.

Why use this database

For most industries the BSD covers over 99 per cent of all UK economic activity, so it has greater coverage than any other alternative.

The SIC information is fully specified – compared to other datasets – and the plant-level location data is fully geocoded. These offer big advantages for local and sectoral analysis by giving practitioners a more complete understanding of their local industrial structure and sectoral strengths.

How to access the database

Access to BSD data is via remote access to a secure lab. Analysts need to pass a free, one-day training course on data security and privacy delivered by ONS and the UK Data Service. The process from beginning to end takes several months and should be factored in to project timelines.

What to be aware of when using this data

The following issues may be relevant to practitioners as they conduct analysis using these datasets. However these problems need to be traded off against the clear advantages of this data.

- BSD requires firms to be VAT or PAYE registered which means it under-samples early-stage start-ups (firms that are pre-employment and pre-revenue). This may cause particular problems for understanding certain sectors heavily skewed toward microbusinesses.

- It is also hard to interpret firms that disappear from the BSD. They may close down but they may also drop below BSD inclusion thresholds temporarily. This is problematic for analyses of openings and closures.

- There are some issues with the timeliness of the data: the publication of the data lags by one to two years, depending on the variable. In addition, large companies will typically be registered only at their head office, when they have offices elsewhere.

Secondary administrative dataset: Annual Business Survey (ABS)

The ABS – previously known as the Annual Business Inquiry, or ABI – is an annual survey of around 50,000 firms, taking a census of larger companies – those with 250 or more employees – and a stratified sample of smaller firms that change between years.

Data on productivity and other detailed firm-level characteristics are asked for in the ABS, although the detail of data varies depending on the questionnaire answered.

Why use this database

- The ABS can provide detailed information for larger firms, including a breakdown of employment costs from pension contributions to severance pay.

- As an administrative dataset, there are fewer risks around the quality of the data.

- Although the ABS is only a sample, the data can be matched with the BSD.

How to access the database

National- and regional-level datasets are available for free online. Lower level geographies are available but may come at a small cost. Data on individual firms is accessed in the same way as the BSD.

What to be aware of when using this data

- The ABS has the same shortfalls as the BSD, in particular the timeliness of publication.

- As it provides a sample of smaller firms that can change year-on-year, the dynamic picture may not always be accurate.

Secondary commercially-held dataset: Open Corporates platform using Companies House data

Certain companies are required to register with Companies House and provide information about their firm. This is published raw at firm level. The information entered includes:

- Registered address
- Date of incorporation
- Current and resigned officers
- Document images
- Mortgage charge data
- Insolvency information

Open Corporates is a platform with business-level information, which is sourced and cleaned up using Companies House data.

Why use this database

As with the BSD, Companies House is a fairly comprehensive database, with most firms required to provide their information. Companies House provides live, online, and geocoded data. The advantage of Companies House data is that, unlike the BSD and ABS, data is wholly in the public domain and there are no confidentiality restrictions on use, even at the level of individual companies.

How to access the database

Practitioners can access information from both Companies House and Open Corporates using their API. They can also pay for direct bulk downloads.

What to be aware of when using this data

There are some limitations to Companies House data:

- Companies are required to self-assign their SIC using a pre-defined list, which can lead to inaccuracies and inconsistencies in the data because of misreporting by firms.

- The SIC self-registration in Companies House is not cleaned and double-checked as it is in the BSD.

- Approximately 30 per cent of firms in Companies House are classified as 'Other'.[1] Some sectors are particularly prone to using this code, for example more than a half of 'Staffing' and 'Health & Wellness' firms classify themselves as 'Other'.[1]

- While self-assignment of a sector is mandatory since 2016, 5.7 per cent of registered Companies House firms did not provide a SIC code.[1]

- Companies House have minimal reporting requirements for small- and medium-sized enterprises and as a result data available for analysis can be incomplete.

- The location of registered companies in Companies House data can also lead to inaccuracies. Large companies will typically be registered only at their head office, when they have offices elsewhere. Some companies provide a 'registered address' which may not be where business activity takes place or where staff are employed.

- Employment information derives from company accounts and is therefore patchy in coverage and dated.

Other commercial alternatives take Companies House data in the UK, clean the data, and combine it with regulatory and other sources (e.g. annual reports).

BvD Data

BvD data (Orbis / FAME) combines data from regulatory and other sources, including 160 information providers, with software to allow users to manipulate data for a range of research needs and applications.

---

1   https://medium.com/@glassAI/using-web-content-to-better-understand-business-activities-7c9e91f3a1b9

Access to these data sets is not free, but can be useful for multinational enterprises and larger firms. They remain imperfect in their coverage of smaller firms. Fame contains financial information on over 2 million public and private companies from the UK and Ireland. It can be used to find company contact information, company financials, company directors and key employees, competitor analysis, ownership data and stock prices for listed companies.

Secondary Commercial Datasets LinkedIn and Crunchbase

Private companies such as LinkedIn provide a platform for companies seeking an online presence leading to connections with both firms and individuals. These networks also provide a wealth of data on firms, often with more detail than administrative datasets will provide.

Data on these datasets can include:

- Team size
- Funding rounds
- Financial settlements

Why use this database

- These types of datasets are much more timely than the administrative sets mentioned above.
- Companies of different sizes will be present on these networks, offering a stronger cross-section of firm size than in the ABS.

What to be aware of when using this data

Some of the challenges around commercial datasets such as this one are explained in more detail in Chapter 2, but in summary:

- They often have limited sampling frames, as a firm must be present on a network like LinkedIn to be counted.
- Individual users on networks like LinkedIn tend to be of a specific demographic, which is not representative of the population as a whole. [2]
- As a commercial dataset, it will not adhere to strict quality assurance as in an administrative dataset.

---

2   https://www.marketing-mojo.com/infographic/infographic-guide-demographics-linkedin-users/

02

Choosing the data to use: what are the risks?

There are important considerations that apply when using data, particularly experimental or commercial datasets:

Data relevance/timeliness

This primarily applies to administrative datasets where coverage and quality is higher, but where the latest available data may already be more than a year old, or to Census data which may be more granular, but older still. The advantage of other data sources, including online platforms as Crunchbase, is their timeliness.

Data quality

Both secondary and primary datasets will have data quality issues with regards to their completeness and any biases.

Commercial secondary datasets

Commercial big data often are not as complete as they claim to be and are only as representative as their user base. This is often skewed towards younger, better-off social groups and particular ethnicities. This makes it important to check statistics about the coverage of data and ask how this sampling might influence results. Case study 8 using LinkedIn data to look at the local supply of skills provides an example of the issues.

Open data and proprietary datasets are often not validated to the standard of administrative sources, which means that the overall quality and reliability of the data may be questionable. Cross-validating exercises with online and offline datasets should be undertaken to attempt to correct and estimate the size of the selection bias. We discussed some of the possibilities in case studies 4 and 6.

Primary datasets using web scraping

Data scraped from the web will suffer from similar problems because of incomplete coverage of online data. For example, many companies do not have a website and not all websites can be successfully scraped. As Nathan et al. (2015)[3] show, "while non-scrapability is likely random, having a website is not". The type of firms captured by online web scraping are systematically different from other types of firms in ways that might bias findings: this is an example of "selection bias".

Given these concerns, whenever possible, analysis based on new datasets should be used as a complement to administrative datasets, rather than a substitute.

A second set of concerns relate to algorithms used to process data. The application of undisclosed algorithms to private data – for example when classifying industries (as in case study 3) - raises questions about transparency, reliability and replicability. When the processing of data is central to the analysis it is important to be as clear as possible on the way in which algorithms work and to understand the potential for systematic bias. Do not underestimate the risks of errors in using these methods or the costs of minimising them through due diligence.

Data Security, Data Sharing and Data Ownership

Being aware of these issues will apply to any primary data sources and any instances where commercial or public datasets are shared or merged.

Maintaining data privacy when sharing data between agencies

Local areas should ensure that enough resources are devoted to protecting the privacy of any data they use or collect. New data partnerships must come with strict data-sharing protocols to keep the data secure and protect the privacy of individuals. This is especially true with the introduction of GDPR regulation. These protocols can be a lengthy and time-consuming process that local areas need to anticipate well in advance.

A major pitfall can be the data-sharing agreement itself: the IT departments of the respective organisations that hold the datasets need to come together and agree on a set of clauses regarding data privacy and data security.

---

3   http://cep.lse.ac.uk/pubs/download/occasional/op044.pdf

Local areas that are new to this process should seek help and advice from organisations that have experience with this process. This comes at a cost, but the payoff can be substantial. In the case study of Greater Manchester (case study 9), they initiated data sharing protocols with local providers such as social workers and police departments to better understand the demands and needs of their local citizens. The result was greater productivity throughout the combined authority.

Some key principles for data sharing are set out in the Data Protection Act 2018.

Sharing local data with other agencies or data analysts

Local areas can effectively only share content if the local information governance team decides it can (Symons, 2016). In addition, any information collected on individuals can be used if local areas have the "informed consent" of data owners. An assessment is then required on the impact that data sharing may have on the individuals whose data are shared (ICO, 2011).[4] Most organisations need to de-identify the data before sharing them but, as the volume of data grows, "deductive disclosure" becomes a major risk. The ONS provides training on this, which shows how computer scientists have been able to re-identify "de-identified" data (Ohm, 2010).

When it comes to data ownership, local areas need to question their rights when sharing their data with private organisations, such as the right to access downstream products or services developed using that data. Local areas may not have much negotiating power with providers, however, and need to consider which asks are negotiable and which are not.

---

4   https://onlinelibrary.wiley.com/doi/full/10.1002/poi3.141

03

Using standard industrial classifications (SIC)

What to be aware of

For any economic strategy, and in particular industrial strategies, understanding the sectoral composition of a local economy can be very helpful. However, this kind of analysis often requires us to categorise industries, for example, by sectors or by their stage of development.

Standard classifications will largely suit most local areas and their industrial structure, but those places that are trying to capture the performance of some industries, particularly new or emerging sectors, will find these classifications are less likely to capture the activity accurately. Alternative data sources can help and we consider some of the possibilities here.

What is the Standard Industrial Classification (SIC)?

The SIC is the workhorse classification used to label industry – and sector – level firm activities in all OECD countries, including the UK. This official taxonomy is used, for example, to measure sector size and trends.

Administrative and commercial data sources often provide information on UK firms classified according to SIC codes. The coverage of these data sources – both in terms of which firms appear and what information is available – varies across sources. We detail some of the commonly used sources in Chapter 1.

Local practitioners should first think about how well a SIC-based analysis will capture the sector or 'industry space' being analysed; and then think about whether to combine SIC-based analysis with alternative approaches.

Why continue to use the SIC

- No fully validated alternative to the SIC system exists. In certain industries, such as mining, manufacturing or public services, the SIC might be sufficiently representative.

- Administrative data classified according to SIC often provides better data on other outcomes of interest (e.g. employment or turnover).

Limitations of the SIC

- The SIC system can be slow to catch up with the transformation of the economy due to the need for international comparability. This means that at times it can provide an inaccurate reflection of the state of a sector or industry, particularly for nascent sectors and technologies.

- Key firms usually have cross-sectorial activities and thus self-classify under several SIC codes (e.g. the BBC).

- There is some disagreement about which SIC codes should be used to delineate certain industries, for example the information economy.

Moving beyond the SIC

Researchers have developed methods to offer complementary classifications for those sectors that can be inaccurately captured by the SIC. These methods prove to be particularly interesting for the analysis of new sectors such as creative enterprise, fintech, biotech or artificial intelligence.

Treatment of the creative industry provides an example of ways in which SIC codes can be used to assess newer activities. SIC codes were established long before the creative industries evolved into their current form. As a result, SIC codes often do not reflect the creative industries very well. One solution is to weight certain SIC codes to allow for the fact that each code might combine a mix of creative and 'non-creative' activities. Currently, the creative industry is converted into a combination of SIC codes, which are used across institutions to analyse data from official government surveys such as the Annual Business Survey or the Labour Force Survey.

Open databases can also allow analysts to verify activities of given firms and validate whether company identities are properly represented by their assigned SIC code. The open database *Open Corporates*, a platform with firm-level information, undertakes additional cleaning of the basic firm-level data provided by Companies House. For local authorities or local enterprise partnerships worried about the accuracy of sector- and sub-sector analysis for their local economy, the data in Open Corporates or reports on the ONS website can help identify whether large local employers fall within industries with ambiguous SIC codes.

Examples of developing alternative classifications using primary, experimental and secondary data sources

There are a range of options available to complement analysis based on the SIC classification. Using datasets that do not use SIC codes can help identify local strengths, which might not show up in SIC-based data. These examples involve primary and experimental data sources. Chapter 2 discusses some of the challenges associated with these datasets.

Experimental data combined with secondary administrative data: Using publicly available data to provide an alternative classification

Open data can be used to identify and classify firms. The most popular examples involve the scraping of web data to better understand a pool of firms.

Nesta's work mapping the UK Games Industry (case study 3) started by scraping video games website companies and web directories, without relying on SIC codes. They identified games companies through the products they release, not the information they supplied when they registered at Companies House. They found 226,302 unique games titles, developers and publishers. They then used Companies House to identify which of these companies were UK-based and compared it to the SIC codes with which these companies were registered. Finally, they compared their results to IDBR data to compare industry sizes. The project was able to identify additional industry hot spots, as well as many firms (two-thirds of the total) that would not have been classified as part of the games industry based on their SIC classification alone.

Nesta's research on Immersive Technology (case study 4) started from a similar point – using web scraping and textual analysis to find UK based organisations that talk about immersive technologies on their web pages. These firms were then surveyed and the resulting data was used to assess the level of project funding available to these organisations and to build a map of clusters and of their links to other industries.

This type of web scraping can be done in house or can be contracted out. For example, tech start-ups, such as Glass AI or Gi, offer algorithms to assign sectors to firms based on the descriptions, webpage content, links on the homepage and other attributes.5

Data sets generated in this way usually need combining with other data or sources of information to cross-validate the resulting classification or to provide additional information on firm outcomes of interest.

---

5   Glass AI is one of many tech companies offering web-scraping services. They have been contacted for the purposes of this report and advise that for customized research like case study number 4 (e.g. looking at sectors, topics or regions), you would expect the budget to be £10k. This figure can be lower or higher depending on the requirements.

Secondary commercial dataset: Using privately held data to provide an alternative classification

About 75 per cent of available data lies outside the control of public authorities,[6] within companies such as Facebook, Twitter and YouTube. Arrangements for accessing this data will vary by source. Once accessed, it can be used to develop alternative classifications of firms.

For example, Nesta used data from Meetup.com, a service used to organise online groups that host in-person events for people with similar interest, to identify tech start-ups that had attended local meetings (case study 5). Information on the topics of these meetings was also used to identify a smaller set of tech fields containing inter-related topics and networks of topics often mentioned together. This data provided a new source of information on north-south disparities in tech start-ups as well as helped identify new hotspots with high levels of meet-up activity.

Academics have used innovative datasets to refine models of sectors and product spaces for many years. Nathan et al. (2015)[7], for example, used Growth Intelligence (Gi) data to get a finer understanding of the ICT sector compared to the BSD. Variables were modelled from unstructured sources such as raw text fragments derived by Gi from company websites and machine-learning techniques were used to identify the ICT product spaces. The authors found non-trivial differences between their new estimate of the ICT production space and the SIC-based estimates: their 'ICT production space' is 42 per cent larger than the SIC-based estimates, with at least 70,000 more companies. They also find ICT employment shares of total employment over twice as large as those found using conventional datasets, from around 3.5 per cent to nearly 12 per cent of all jobs from 2008 to 2012.

Examples of combining data sources

Analysis will likely need to combine more than one dataset to be useful. Effectively combining datasets requires an identifier that can be used for matching.

In the UK Games Industry study (case study 3), this involved matching online data on firms scraped online to data from Companies House. In the Welsh innovation mapping study (case study 2), this involved combining data from firm datasets; open data published by Research Councils and Innovate UK; and web data including company websites, professional meeting platforms and Twitter accounts.

Combined qualitative and quantitative sources

Combining quantitative data with qualitative data can also help build a finer understanding and enlarge datasets that might be otherwise incomplete.

---

6   'Big data: Are you ready for blast-off', BBC, March 2014. www.bbc.co.uk/news/business-26383058

7   http://cep.lse.ac.uk/pubs/download/occasional/op044.pdf

Most quantitative data will not provide information on management practices, how new knowledge is accessed and shared by firms, or on firm investment and supply chain links – all of which may inform economic policy development. Qualitative data may fill these gaps, but it is not easy or cheap to collect.

Greater Manchester Combined Authority's (GMCA) work on the textile industry provides an example of the way in which quantitative and qualitative data can be combined to inform a supply chain analysis for local industrial strategies (case study 6).

In this example, domain experts reviewed and adjusted the findings from survey responses weekly. GMCA gathered online information on over 200 firms and complemented this information with 130 site visits of retailers during 2013 and 2014 to understand the supply- and demand-side barriers to growth. This survey and methodology allowed a micro level understanding of the entire textile supply chain in the UK. Most importantly, it established networks of firms and identified which were most central to the development of the supply chain.

Essex County Council also provides a good example (case study 10) of how to combine insights from qualitative and quantitative findings and how qualitative insights can be structured and mapped out against quantitative results.

Limitations of qualitative and quantitative combined datasets

Unstructured qualitative interviews can result in, among other challenges, over-representation or under-representation of particular demographics, and biased feedback due to vested interests.

Given these pitfalls, it is worth considering the following questions when combining quantitative and qualitative data:

- Is the sample representative of the broader population?
- Is there a financial incentive to participate? Does this induce any selection or availability biases?
- How do you balance qualitative findings against the results from quantitative data?

Example of building primary data sources: Generating new data through surveys

If the data needed is not available, then primary research may be required. For example, in the Immersive Technology study (case study 4), Nesta surveyed all 2,000 firms that were classified under a specific industry via web-scraping.

Other approaches can be used to identify the right pool of firms to survey. For example, working from an initial list of relevant firms, a 'snowball' survey can be rolled out by asking surveyed firms to name other firms that should be interviewed.

This method can be particularly informative for informal activities, which would not show up in administrative datasets, but it can be time consuming. For example, when Nesta tried to snowball a survey to complement their mapping of innovation in Wales based on existing datasets, they struggled to deliver the survey on time (case study 2).

Warning: sectoral analysis and its limitations

Given our current state of knowledge about the underlying drivers of economic growth, local policymakers should be wary about trying to achieve a particular sectoral composition – shares of businesses or employment in specific industries – in their area. The sectoral composition of a local economy is an outcome of a large number of decisions taken by both firms and workers about where they want to locate and what they want to produce. Furthermore, the link between sectoral composition and local economic performance is complex. The supply chains – or 'value chains' – for today's industries typically stretch well outside a given local economy, and sometimes across multiple countries. They are driven by factors that are not straightforward to understand, and will change in the future in ways that are hard to predict.

Yet analysing the sectoral composition of a local economy can be helpful for several reasons. Economic conditions can vary across different sectors and sector-specific knowledge can help design 'horizontal' policies – such as skills and employment training programmes – more effectively. Sectoral analysis can also help local policymakers understand the local spatial footprint of a given industry and how their area is plugged into larger value chains. In turn, this can help local areas co-ordinate their own policies with national interventions – especially sector deals, but also cross-sector 'grand challenges' – and potentially justify funding and support for these.

For some questions, simply understanding sectoral shares in the local economy, and the likely pattern of development over time will be helpful – for example, when trying to predict demand for specific skills or types of commercial provision. Shift-share analysis (an estimate of local growth attributable to local, industry and national factors) might help understand current trends. However, simple sectoral shares and shift-share analyses do not identify areas of relative strength, or help make the case for policy.

As discussed above, identifying relative strengths needs appropriate comparisons on appropriate indicators. Comparison with national industry shares, typically through location quotients, is a good first step. Herfindahl-Hirschman indices, which effectively measure the market share of individual firms within a given industry, may also help. But these measures generally need supplementing with analysis of sectoral *productivity* to understand whether relatively large employment shares represent strength or weakness. Even then, none of this analysis makes the case for policy to intervene. Similar arguments apply for skills, infrastructure and other key growth drivers.

04

Analysing the data – what to think about

Some privately-held or experimental datasets require advanced data analysis skills. Similarly, when combining datasets – other than data security – there may be other considerations to ensure this works as well as possible.

In some cases, analysis of the data will involve working with local partners instead of sourcing the skills in-house. Local practitioners need to think about the different mechanisms for making the most of, and analysing, data. Some new datasets may be unreasonably expensive to acquire and fail to provide the level of insight to offset the cost.

Striking data partnerships with local providers

Manchester City Council's development of a data warehouse provides an example of how to integrate data from multiple agencies (case study 9). Such activities can also offer insights from beyond local authority boundaries relating to the broader benefits of local economic activity.

Striking data partnerships with national providers

National providers such as the Open Data Institute or the Office for National Statistics are proactive in finding data solutions for local areas. The ONS have, for instance, released the first experimental balanced regional gross value added (GVA) estimates, by bringing together existing income- and production-based measures into single estimates for each region.

Data partnerships with private businesses

These partnerships can mean access to richer, more granular data. The creation of new kinds of mutually beneficial public and private partnerships, or "data collaboratives", can help policymakers acquire more timely and granular data. Such public-private exercises are not risk free. Working with private companies can raise new challenges (see some of the risks highlighted in Chapter 2). Private sector partners are unlikely to engage in this kind of activity unless they see a business case. Clarity on the incentives and risks of data sharing is essential.

Data partnerships with universities and other not-for-profits

These partnerships can primarily help with access to data analysis skills. Local areas can collaborate with colleges and universities that are keen to work on policy-relevant data, and that might seek work placements for data science and economics students. Local areas could also tender to tech start-ups with a good track record of data privacy or get in contact with groups like the Open Data Institute and Nesta.

05

Developing a granular understanding of the local economy

Poor data has been flagged many times by local practitioners as a limiting factor in understanding their local economies.

In this section, case studies where existing secondary data and classifications have been used to get more detailed information on local industries and on the links between different parts of the local economy. As well as making better use of existing data, the options for more in-depth analysis using merged or experimental resources are discussed.

Making better use of secondary datasets

For practitioners who tend to provide fairly simple comparisons of local and national averages using administrative datasets, a simple way to improve their analysis is to access the underlying business-level micro-data instead. This can offer a more detailed picture of the industrial structure of a place as a whole, as well as firm-level detail on employment levels and turnover.

This data is publicly administered, which reduces some of the risks associated with commercial datasets such as quality or bias.

Recent work in Greater Manchester[8] and West Midlands[9] has drawn on microdata to better understand differences in productivity across firms with a focus on the performance of the 'long tail' of low productivity firms.

---

8   https://www.greatermanchester-ca.gov.uk/media/1132/gm_prosperity_review_baseline_report___evidence_review___
        november_2018.pdf
9   https://www.wmca.org.uk/media/2227/productivity-and-skills-commission_full-report.pdf

Combining more than one secondary dataset

Merging different micro-datasets can provide a further level of detail for understanding local economies. Case Study 1 shows how the ONS combined business-level data from the IDBR[10] with detailed data on domestic and foreign fleet landings by port. The purpose of bringing these datasets together was to help assess the economic and social impact of the changes to the UK fishing industry on local coastal communities.

More detail about this specific case is available, which includes links to access the code and build-up visualisations that were used to display the combined data sources. In this example, some of the data privacy risks of using more detailed datasets are avoided by communicating the data through maps and other visualisations.

What data is available and how they are best combined to reach conclusions may not be clear to those undertaking this kind of analysis for the first time. Making the most of data partnerships, as discussed above, could be of benefit here.

Combining and mapping micro-level and experimental datasets

Visualisations and maps of existing micro-data are a useful tool for gaining deeper insights into a data source than more simple analysis tools. In case study 2, mapping is used alongside a combination of administrative and experimental data sources to establish a detailed picture of research and innovation across Wales. Nesta used interactive visualisations designed to identify innovation system failures where people with complementary capabilities (e.g. firms with a problem and researchers with a solution) do not necessarily talk to, or even know, each other.

The maps aimed to encourage collaboration by increasing the visibility of the existing research activities, and classifying creative industries into four circles according to commonalities in business models and supply chains. This type of visualisation can also help local areas understand where firms fit within national supply chains.

An exercise of this type requires a strong data partnership with a trusted organisation, an understanding of the risks associated with the quality of some primary or experimental datasets, and an awareness of privacy and security challenges.

Alternative data sources can be used to try and better understand the links between firms. In this section, we review some of the simple analytical tools available, as well as those which use more advanced mapping and analysis skills.

---

10   ONS's Inter-Departmental Business Register (IDBR) is a comprehensive list of businesses used by the government that covers 2.6 million businesses in all sectors of the UK economy. See Annex A.

06

Understanding the links between firms

Mapping supply chains is difficult and it may not provide the insight different local areas need. This is because, outside of input-output tables, building up an understanding of the links between firms requires primary datasets as the available secondary data is not adequate.

Input-output tables

Traditional input-output tables describe the sale and purchase relationships between producers and consumers within an economy. They provide a picture of the flows of products and services in the economy each year and how these are used to produce further products and satisfy final demand. They are one way of understanding the links between firms. The Office for National Statistics traditionally published these once every five years, although they now aim to produce them annually.

Building on an understanding of sectors, these tables can provide some insight into how firms in different sectors could interact. Input-output tables use a version of SIC, which enables them to be matched with other datasets.

Limitations of input-output tables

- While they use SIC codes, which can in themselves be limited, they are also aggregated, so some detail is lost.

- They provide national, rather than local, data, so it cannot be understood from the data whether or not these firm links are happening within a local area. Any information derived from input output tables against an understanding of industrial structure should be used as a proxy.

Mapping Supply Chains

There is some academic literature[11] that has been able to improve on traditional input-output tables using value added data. To the best of our knowledge such analysis is not feasible in the UK, at least based on data available to researchers. This means that efforts to better map supply chains have had to rely on primary data instead.

Greater Manchester Combined Authority's work on the textile industry (case study 6) used primary data sources to build a picture of the links between firms in a single industry. The primary sources were both quantitative gathering information over a two-year period, and qualitative via surveys with firms.

By combining quantitative and qualitative data, the combined authority was able to build a detailed understanding of the local supply chain and an analysis of how central different firms are to the sustainability of the industry. This in turn informed their decision about which firms to invest in as well as the funding conditions of their Regional Growth Funds.

Greater Manchester was able to apply this primary analysis to policy and make evidenced decisions about which firms to offer support to, but it was a resource-intensive approach and may not be available to smaller places.

This case also highlights uses of alternative industrial classifications, see Chapter 3.

Mapping knowledge flows

Policymakers will be interested in knowledge flows between industries as well as the flow of goods and services.

The case study mapping innovation in Wales by Nesta used collaborations, citations and professional meeting platform discussions to build visualizations of local networks and knowledge transfers (case study 2). This helped the Government of Wales identify areas of potential future partnerships between firms of a given network.

Sometimes other datasets are used as proxies for knowledge flows. In case study 5, Nesta used Meetup.com data to explore how knowledge exchange between tech start-ups in the UK varies. They mapped the correlations and associations in the topics discussed between meet-ups across the country and used this to understand how knowledge transfers and technological innovation takes place, and in which parts of the country this happens more or less often.

---

11   Bernard, Moxnes, Saito (2015) Production Networks, Geography and Firm Performance, NBER Working Paper No.21082 http://www.nber.org/papers/w21082; Bernard, Sweets, Warzynski (2016) http://cep.lse.ac.uk/pubs/download/dp1423.pdf

07

Supply-side or demand-side?

A key challenge in local economic strategies is to assess whether interventions are needed on the demand- or supply-side. In this final section, we review two examples that go beyond relying on traditional market signals – e.g. prices of commercial and residential property or the premium paid to graduates – and provide alternative data sources to complement the picture of local supply and demand.

Using job vacancy data

Vacancies data from the Employer Skills Survey or online jobs postings can be used to identify demand side skills shortages. Work by the Centre for Cities offers an example of the innovative use of data to better understand skills demand (case study 7). They used job advert data from Burning Glass to produce skills maps showing skills that are most frequently requested by occupation group, as well as skills that have experienced the fastest and slowest growth in mentions. This can help inform forecasts of skills demand.

Using LinkedIn data

On the supply side, an example of the innovative use of data is provided by Greater Manchester's work with LinkedIn (case study 8) to examine the profiles of the 614,000 LinkedIn users in the geographic area. The report identified strengths in technology and media and weaknesses in medicine, science, banking as most graduate students were identified as leaving the area for London. These findings were fed into the Local Digital Skills Plan but also used by LinkedIn to match jobs and people where the local gaps were.

When considering supply, products such as Whythawk's Pikhaya analyse business rates data to understand areas where demand for business space is high or low. This type of data can provide a means of understanding firm dynamism and competition in a local area.

ANNEX

Case studies

Case study 1: The state of the UK fishing industry

Office for National Statistics

Context and objective

This case study shows how the Office for National Statistics (ONS) collected local data to get a finer picture of the UK fishing industry. The project explores how visualisation can be used to paint a more detailed picture of industry and geographical data without disclosing sensitive information.

Problem

How to assess the economic and social impacts of the UK fishing industry on coastal communities – an issue on which there is little data.

Methodology

ONS linked the Inter-Departmental Business Register (IDBR) to a very detailed dataset on UK fleet landings and foreign fleet landings into the UK, by port.[1] This includes details on the volume and value of fish and marine species landed at UK ports by port and vessel nationality. This allowed them to construct composite indicators from IDBR variables (e.g. turnover and employment) for each Lower Layer Super Output Area and fishing sub-sector.

---

1   https://www.gov.uk/government/statistical-data-sets/uk-fleet-landings-and-foreign-fleet-landings-into-the-uk-by-port by the Marine Management Organisation (2015)

ONS explored two types of visualisation which allow users to compare port-level activity, including number of sales and total value of sales, across the UK. These were sankey diagrams and interactive maps. Sankey diagrams allow users to compare port-level activity by constructing a chart by either clicking ports on the map, or by using the look-up box in the controls.

Results

The prototype uses information on the UK fishing industry to explore its impact on small coastal regions and also to generate summary statistics for the UK fishing industry, derived from the IDBR.

Limitation

There is a hard limit to the granularity of the data visualisation, which is dictated by privacy requirements.

Why does it matter for local industrial strategy?

This is a good example of how visualisation can help communicate useful information at a granular level, without being disclosive. The code and steps required to create the industry dashboard and visualisation are available as open source code online. This is one example of open source tools for local authorities and local enterprise partnerships to use when looking to understand the relative dynamism of a national industry in a local area.

Case study 2: Innovation mapping in Wales

Welsh Government and Nesta

Context and objective

This case study offers an example of how web data can be scraped to highlight the interdependence of local businesses and how they interact and fit within national supply chains. In addition, the maps might encourage local businesses to collaborate and share information with businesses located nearby.

Problem

Innovation can be captured by various proxies (e.g. patents, trademarks, innovation surveys), but even when taken together, they may fail to present a complete picture of the state of innovation in an area.

Methodology

1. Data was combined from different sources including business datasets from Office of National Statistics; open data published by Research Councils and Innovate UK, and web data (e.g. company websites, professional meeting platforms, Twitter accounts, among others).

2. A novel data platform was created that gathered and automatically synchronised this data.

3. Visualisations and maps of local networks were built.

Results

- Maps showing the connectivity and cross-pollination between different research institutions and tech communities in areas across Wales

- Organisations are more likely to look for collaborators nearby: a third of the research collaborations identified were inside the same area

- Networks were mapped to help identify areas of potential partnerships with neighbouring local authorities and local enterprise partnerships.

Limitations

One of the data pilots used company website data to 'snowball' a survey on a network of innovative companies; however, it failed to deliver on time. Such data projects can take longer than anticipated.

Why does it matter for local industrial strategy?

Identifying and supporting clusters is likely to be an important part of local industrial strategy. Nesta's interactive visualisations of the business network map, available in their online report, increases the visibility of research activities already taking place, which could help local authorities and local enterprise partnerships identify burgeoning clusters.

This case study offers an example of how web data can be scraped to highlight the interdependence of local businesses and how they interact and fit within the national supply chains. In addition, the maps might encourage local businesses to collaborate and share information with businesses located nearby.

Case study 3: The gaming industry in the UK

Nesta

Context and objective

This case study illustrates how to collect new data on an emerging industry made up of relatively smaller companies, and how to compare and contrast the results with more conventional data from the Inter-Departmental Business Register (IDBR) that uses the Standard Industrial Classification (SIC) system.

The objective of the study was to map an industry for which high quality and timely data is sparse.

Problem

The SIC system has been slow to catch up with the transformation of the economy and is often an inaccurate reflection of the nature of current business activities. In this example, SIC codes capture gaming companies poorly. The gaming industry SIC code only appeared in 2007 and has failed to adequately label many firms, as gaming companies straddle different sectors and SIC codes (e.g. educational games, games app, and developers). Given the price of carrying out new surveys, Nesta opted instead to exploit free online data sourced from public websites. This data is real time and is not confined to SIC codes.

Methodology

1. Online data was sourced from 226,302 unique games titles, their developers and publishers.

2. Companies House – an open database of companies and company data – was used to identify which of these companies are UK-based.

3. DueDil validated company identities through web search and used decision-tree methods to identify non-validated companies that have a high probability of being gaming companies.

4. Quality assurance was conducted using additional data from *DueDil* and through conversations with domain experts, *Ukie*. Experts removed 17 companies from the list and added 184 companies that had no web presence

Results

The final master dataset is composed of 1,902 firms. The project found that different industry "hotspots" were identified using online scraped data as opposed to BSD data. Conversations with domain experts revealed that extra hubs identified by the web analysis may be more plausible than those detected through the Business Structure Database (BSD); namely Liverpool and Cardiff over Hull and Reading. They also found that only a third of businesses were in the official SIC codes because of classification issues.

Limitations

Web data captures information only on firms or individuals, which have left a "digital footprint". In this example, domain experts added 184 companies to the web analysis results, as those companies had no web presence. More generally, the sample of firms captured through online data might present a biased picture of the actual pool of firms of interest.

Lessons learnt

- Web data is not a substitute for but a complement to other more-traditional data sources.
- Automated data collection can and should be verified and cross-checked with domain experts.
- The risk of errors or the costs of minimising them through due diligence should not be under-estimated.

Why does it matter for local industrial strategy?

Some booming industries that might be important to consider for local industrial strategy are hard to measure. This case study provides an example of a way to map and estimate the size of such an industry. On the basis of that evidence, local authorities and local enterprise partnerships can consider whether to support those industries according to the principles outlined in our previous report.

Case study 4: Immersive technology mapping

Immerse UK, Innovate UK and Nesta

Context and objective

This case study uses mixed methods to obtain a clearer picture of an emerging industry. It provides an example of how to scrape websites to shortlist and target firms for a survey, and collect feedback and visualise barriers and opportunities.

The objective of the study was to provide hard data about the size of the sector, its performance, its geography, the drivers of success and its barriers to growth.

Problem

Despite it being a booming technology, the data on immersive technology in the UK is limited. As there is no industrial code for immersive technology, it is impossible to measure the sector using official statistics. This study used open source web data to target a set of firms and map the industry.

Methodology

- Nesta partnered with a tech start-up, Glass AI, to collect web data on hundreds of thousands of UK organisations

5. It scraped the web for organisations that talk about immersive technologies and found 2,000 organisations.

6. Partnered with MTM London, a creative research agency, who set up an online survey targeted at the 2,000 organisations identified in the previous stage

7. Immersive tech developers and creative companies producing immersive content were mapped, as well as organisations in other sectors using immersive technology to deliver new products and services.

8. The survey also worked as a 'labelled dataset' that was used to generate predictions about the level of immersive engagement, size and turnover of those organisations that did not respond to the survey

9. Several models were trained for size, turnover and engagement class labels with cross-validation (i.e. testing the robustness of the predictions on a held-out set) and the best performing model was chosen.

Results

1. It was dentified that this technology may create value in other industries (e.g. creative, education, training, architecture, advanced manufacturing and energy).

2. A UK map of potential clusters was built – controlling for the size of local economies. It identified 14 locations in the UK that are relatively specialised in immersive technology.

3. It was estimated that funding devoted to immersive technology projects was nine times higher in 2016–2017 than in 2009–2010

4. The main drivers and barriers per geographic area were identified. Notably, it found that skills shortages and a weak tech ecosystem may be important obstacles to growth in the immersive sector across the UK

Limitations:

- Working with agencies such as MTM London and Glass AI can be costly
- Data analysis carried out for cross-validation and prediction can be time consuming

Why does it matter for local industrial strategy?

This study uses mixed methods to obtain a clearer picture of an emerging industry. It provides an example of how to scrape websites to shortlist and target firms for a survey, and collect feedback and visualize barriers and opportunities.

This method could help Local authorities and local enterprise partnerships rank themselves in comparison to other hubs across the country and perhaps seek out collaboration with other UK hubs. In addition, it highlights synergies between different industries and could help Local authorities and local enterprise partnerships build a rationale for policy intervention. You can find more information on how to think about this rationale in the design principles outlined in our previous report.

Case study 5: Mapping tech start-ups and communities in the UK

Nesta

Context and objective

Nesta tracked tech communities with *Meetup.com* data and mapped the location of 'Meetups' and their different technology specialisations. The objective of the study was to help local authorities and local enterprise partnerships track the trends and the supply chains in an emerging industry by mapping the geography of tech 'Meetups'.

Problem

Data on business networks and their spatial diffusion is not available in conventional datasets.

Methodology

1. The Meetup API for groups in the "tech" category in UK cities was queried. This returned a list of 1,391 businesses, after removing duplicates.

2. These topics were arranged into a smaller set of 'tech fields' containing inter-related topics.

3. A data-driven approach was followed based on scientometrics principles, i.e. the quantitative analysis of science and technology metrics such as academic papers and patents.

4. Maps of these associations in a 'topic network' were created where topics often mentioned together were pulled together.

5. Community detection algorithms were used to look for densely connected 'clusters' of topics within these wider clusters.

6. This was used to determine the tech specialisation of Meetup groups.

Results

- The north of the UK was found to have less tech Meetup activity than the south.

- High levels of tech Meetup activity was found in locations not previously identified, such as Belfast, Birmingham and Cardiff.

Limitations

Data completeness is a pervasive problem with websites such as Meetup as not all business-networking events are captured. This limitation highlights the importance of using complementary data to get a fuller picture.

Why does it matter for local industrial strategy

This type of data can help local authorities and local enterprise partnerships:

- Track the emergence of new fields and innovations

- Map the geography of business Meetups

- As a proxy for the relative specialisation of a region

- Measure industrial networking and knowledge spillovers in creative and tech clusters

This case study provides an example of way of measuring potential spillovers that might form the basis for policy intervention as discussed in our previous LIS report.

Case study 6: Textile industry survey

Greater Manchester Combined Authority

Context and objective

The combined authority gathered information on over 200 firms and several retailers to understand the supply- and demand-side barriers to growth alongside a desk review of key literature.

Problem

Quantifiable data is not readily available to Local authorities and local enterprise partnerships on barriers to growth and demand forecast.

Methodology

1. Manchester Local Authority gathered information on over 200 firms and several retailers during 2013 and 2014 to understand the supply and demand-side barriers to growth alongside a desk review of key literature.

2. They complemented this information with 130 site visits.

3. KPMG looked at the data and estimated the size of the growth opportunity.

4. Feedback that arose from the survey was discussed and analysed by domain experts at weekly industry-led board meetings and triangulated with quantitative data and national statistics.

5. Cluster and network analysis was carried out and centrality tested to understand which firms to invest in.

Results

This survey and methodology allowed a micro-level understanding of the entire textile supply chain in the UK. It identified barriers to growth and opportunities. Most importantly, it established networks of firms and identified which firms were most central to the development of the supply chain.

Limitations

Such a large qualitative endeavour comes at a high cost.

Why is it interesting for local industrial strategy?

The conditions for funding by the Regional Growth Funds were chosen on the basis of these survey findings. As such, the findings were directly used to guide policy intervention. Local authorities and local enterprise partnerships might similarly consider selecting firms to invest in on the basis of a detailed understanding of the local supply chain and how central firms are to the sustainability of an industry.

Case study 7: Using big data to map and forecast skills

Centre for Cities

Context and objective

This case study is an example of innovative use of data to forecast skills supply and demand in a local area. Centre for Cities processed data provided by Burning Glass, which collects online job advertisements.

Problem

It is hard to find data on individual skills as well as trends in skills supply and demand.

Methodology

Burning Glass processed the skills data.

Results

- A map shows skills that are most frequently requested by occupation group compared to the skills required for all occupations.
- The map shows the skills that have experienced the fastest and slowest growth in mentions between 2012-14 and 2014-16.

Limitations

1. Burning Glass only captures online job adverts, excluding many jobs that are not captured by this website. The framing sample is biased towards more digital and new economy jobs

2. Skills are at times mislabelled within job adverts, which can lead to measurement error.

3. Online job adverts are problematic in many ways as a proxy for skills but can form the foundation layer for a skills map.

Why does it matter for local industrial strategy?

Given the endemic problem of skills shortage in the UK, skills policy is likely to be a big component of local industrial strategies. This case study offers a useful and readily applicable example of a cross-regional skills comparison, as well as credible forecasts of skills demand and supply.

Resources

View a prototype, or for other resources on employer demand for skills and a skills map, Centre for Cities collected a range of useful tools and datasets at the local authority level.

Case study 8: Labour market characteristics

Greater Manchester Combined Authority and LinkedIn

Context and objective

This data provides a complimentary picture to the one offered by conventional datasets. In this example, data was fed into the Local Digital Skills Action Plan of Greater Manchester Combined Authority.

Methodology

1. 614,000 LinkedIn users in the geographic area were profiled.
2. Skills, hiring, migration and education were considered.

Results

- The report identified strengths in technology and media and weaknesses in medicine, science and banking as most graduate students were identified as leaving the area for London.

- Data was fed into the Local Digital Skills Action Plan

- Data was used by LinkedIn to target jobs and courses to the people where the local identified gaps exist.

Limitations

Data acquired by platforms like LinkedIn are not complete or comprehensive. Rather, the observations are often biased or over-represent a specific population as the platform has a higher coverage of some sectors than others. Equally, the coverage of certain skills or labour markets of more remote areas is almost absent. This potential bias must be acknowledged and taken into account when drawing higher-level conclusions.

Why is it interesting for local industrial strategy?

Skills demand and skills forecast is instrumental to the design of local industrial strategies. This data provides a complimentary picture to the one offered by conventional datasets. In this example, data was fed into the Local Digital Skills Action Plan of Greater Manchester Combined Authority.

More details and insights can be found here.

Case study 9: Data sharing agreements

Manchester City Council

Context and objective:

While not directly relevant to LIS, the methodology and steps taken by Manchester City Council to create a data warehouse and link up different datasets carries lessons for Local authorities and local enterprise partnerships. Manchester's integration of data from multiple agencies gives social workers rapid insight into family circumstances and needs during triage and assessment.

Methodology

1. 16 data-sets were integrated from multiple agencies into one system.

2. Each agency loaded its data in to an agreed schedule.

3. A range of statistical techniques was applied (e.g. decision trees, cluster analysis, spatial analysis) to forecast the impact of policies.

Results and lessons learnt

• Manchester sought legal advice on data security to allow key workers to access data without having to ask.

• It designed a Privacy Impact Assessment with the guidance of a barrister and legal experts to ensure the safeguarding of this data.

• The system records digital footprints of system users, which means that any inappropriate use can be detected.

- The greater quality, synchronisation and consistency of the information available to all case workers removes friction due to data sharing and data harmonising and ensures that the data is consistent among decision-makers.

- Nesta evaluated that the data system saved workers three to four hours per assessment. Given workers go through 40 assessments a year, this translates into saving two weeks of some workers' time, or a 4 per cent increase in the total amount of worker resource.

Why is it interesting for local industrial strategy?

More granular data is used to accurately estimate demand for services and thus informs the budget and resources needed for service provision. In addition, data-sharing agreements prove to be useful tools for local authorities and local enterprise partnerships for short-term local industrial strategy goals such as increased productivity, and longer-term goals such as reduced costs of data collaboration among users of this data.

Case study 10: School readiness programme

Essex County Council and Basildon Borough Council

Context and objective

While not directly relevant to local industrial strategy, this is a useful case study of a triangulation exercise by a local authority. The lessons learnt and methodology used can be readily transferred to triangulation exercises for industrial policy.

Essex County Council (ECC) wanted to better predict school readiness. The objective of this project was to predict whether children will be school ready before they start reception. More information can be found on their website.

Problem

In an effort to understand the needs of poorer students in their county, ECC wanted to go beyond quantitative analysis and triangulate the data with ethnographic studies.

Methodology

1. ECC identified hotspots in terms of level of risk of various neighbourhoods.

2. ECC conducted the quantitative and qualitative research separately.

3. The company TONIC was contracted by Basildon Borough Council and ECC to engage the community and collect insights on the levels of risk of the various neighbourhoods.

4. TONIC used strong community champions and volunteers to carry out some of the interviews and qualitative data collection. They also used social media to publicise engagement events.

5. They compared the results from quantitative and qualitative data and found that some communities were identified by both approaches, whilst some communities were only identified by one of the two.

6. This led to conversations on the strengths and weaknesses of the hotspots identified.

Results

The qualitative data helped identify new hotspots and new themes in the community.

Lessons learnt

1. ECC program managers said that qualitative data was very powerful to convince senior sponsors who might be otherwise wary of results purely based on predictive quantitative models.

2. They also mentioned it allowed for a more dynamic interaction with the population rather than using data alone.

3. Using local volunteers and community champions was seen as a good way to gather relevant information and for the community to take ownership of the project.

Why is it interesting for local industrial strategy?

This case study provides an example of how to triangulate quantitative and qualitative data. Local authorities and local enterprise partnerships could carry out ethnographic engagement work either in-house or through external agencies to inform their understanding of their business or industrial base.

what works centre for
local economic growth

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

HM Government