LDC index methodological note

Constructing the index of local digital capital – How do we move from the model to a unit-free index?

Many socioeconomic indexes, such as the Human Development Index, convert a raw input figure into a unit-free index (typically on a scale from 0-1 or 0-100), using the formula below.

x the raw figure in the data

a the "minimum" figure for that variable among the nations and regions

b the "maximum" figure for that variable among the nations and regions

$$x \text{ index} = \frac{x - a}{b - a}$$

But to ensure within component index scores that are reflective of a nation's/region's true performance and not just their comparative performance, we decided to set *a=0.* As such, a nation or region will only score zero on any individual LDC measure if the data for that measure is indeed zero.

Where there is a natural maximum of 100%, such as for infrastructure variables like superfast broadband and gigabit coverage we set *b=1*. This allows us to benchmark the nations/regions against one another and against an achievable long-term benchmark of full coverage.

For other variables that are measured in percentage terms we have taken a more flexible approach.

For R&I expenditure, we took the OECD target of 3.0% of GDP as the benchmark "maximum" (i.e. *b* in the equation above). Since the data was scaled by per labour hours, we took 3.0% of UK GVA scaled by total UK labour hours worked as this figure. All the nations/regions, except the East of England, were below this benchmark so their index was calculated automatically using the formula above. To avoid an index score greater than one, we manually set the East of England's index score for R&I expenditure to exactly one, to indicate that this region has already achieved the 3.0% benchmark.

Where the maximum benchmark was less clear and achieving 100% on a measure is an implausible target, such as percentage employment in tech jobs, we set *b* equal to the actual maximum data point for the nations/regions. One borderline case was the percentage of firms using basic digital skills. We set *b* equal to one (100%) here as we thought it would be a plausible long-term target for all UK firms to utilise basic digital skills in some way.

For raw data, such as total researchers (scaled by labour hours) or SME bank lending (scaled by labour hours), we set *b* to the raw maximum of the nations and regions. The individual *within* component index here for a nation/region is therefore essentially a percentage compared to the leading nation/region.

Variables that needed to be scaled in per capita or per hours worked terms?

The economic variable of interest used in our econometric model was GVA per hour worked, which is essentially a measure of labour productivity and very highly correlated to GDP per capita (and therefore average living standards) since labour hours per capita vary little between nations/regions. Therefore, the data added to the model needed to be scaled in per capita or in per hours worked in order for the index to make economic sense.

Data that is as a percentage of the population should strictly be rescaled in per hours terms (rather than per capita), however since the model is set up in "log-log functional form" the economic effects of a percentage change in each independent variable (such as percentage of the properties that have superfast broadband coverage) on labour productivity are unaffected by the multiplicative scaling difference between per capita and per labour hours worked. As such, percentage data is not scaled.

The same economic logic applies to variables such as F&I. F&I improves either total productivity or is used to deepen a firm's capital stock, both of which improve labour productivity (which is a per labour hours measure).

'Between component' weightings

The econometric model helped us understand which components of LDC had a greater positive impact on labour productivity. Consistent with previous empirical work in other similarly developed countries we found a larger economic impact of digital skills and a slightly lower (or no clear picture*) impact for the other components and therefore weighted skills highly (30% of the overall index). This also reflects the findings of the European Commission's Digital Economy and Society Index (DESI), which weights 'human capital' (including multiple measures of digital skills) joint highest of their components.

'Within component' weightings

A higher weighting was generally given to the LDC measure within each component that produced the most precise (i.e. lower standard deviation and result robust to different exact specification of the model) results, as we have greater confidence that variation in measure of LDC is driving differences in labour productivity between the nations/regions.

Where various within component measures produced similar model results, we took a balanced 'within component' weighting (such as 50-50 or 30-30 and 40 to the various other measures within that same component). In most cases, with the exception of R&I, changing the 'within component' components does not noticeably alter the LDC index and ranking for the overall component, so we can say the index and rank for these components is robust to particular weighting choices.

*This does not mean the model indicates no positive economic impact, but rather it indicates a high variation in point estimates (which were still positive but did not cluster around a particular number depending on the exact model specification used).

Component breakdown

Digital skills

Basic skills: This data is from the Lloyds Consumer Digital Index and measures the percentage of the population with basic or foundation digital skills (according to Lloyds definition). This is benchmarked against a maximum figure of 100%.

Skills proxy: This data is Eurostat internet usage data and measures the percentage of the population who have accessed public services online. This is benchmarked against the maximum figure achieved by any nation or region in any year measured. This is benchmarked against a maximum figure of 100%.

Internet Users: This data is the ONS Internet Users data, and measures the percentage of the population who have used the internet in the previous three months (most frequent measure). This is benchmarked against a maximum figure of 100%.

Digital adoption

Tech jobs as a percentage: This is data from the Technation Jobs and Skills Report and measures the number of tech jobs as a percentage of the total number of jobs in the region. This is benchmarked against the highest percentage achieved by any nation or region in any year measured.

Tech firm employees: This is data from the Technation Annual Reports, and measures the number of tech firm employees as a percentage of total employees in the region. This is benchmarked against the highest percentage achieved by any nation or region in any year measured.

Businesses with essential digital tasks: This is data from the Lloyds Business Digital Index and measures the percentage of businesses in a region undertaking at least one 'essential digital task' (as defined by Lloyds). This data is benchmarked against a maximum figure of 100%.

People with essential digital skills for work: This is data from the Lloyds Consumer Digital Index and measures the percentage of people with 'Essential Digital Skills for Work' (as defined by Lloyds). This data is benchmarked against a maximum figure of 100%.

Information and communication jobs: This data is from the ONS Workforce jobs by region dataset, and it measures number of information and communication jobs as a percentage of total jobs in a region. This is benchmarked against the highest percentage achieved by any nation or region in any year measured.

Tech jobs (broad definition): This data is from the ONS Workforce jobs by region dataset, and it measures the number of information and communication, finance, insurance, scientific and technical jobs as a percentage of total jobs in a region. This is benchmarked against the highest percentage achieved by any nation or region in any year measured.

Digital infrastructure – connected nations

Superfast Broadband: This data is from the Ofcom Connected Nations reports, and it measures the percentage of premises with superfast (30Mb-300Mb) broadband coverage (the available coverage, not the take-up). This data is benchmarked against a maximum figure of 100% coverage.

Ultrafast Broadband: This data is from the Ofcom Connected Nations reports, and it measures the percentage of premises with ultrafast (300Mb-1Gb) broadband coverage (the available coverage, not the take-up). This data is benchmarked against a maximum figure of 100% coverage.

Gigabit: This data is from the Ofcom Connected Nations reports, and it measures the percentage of premises with gigabit broadband coverage (the available coverage, not the take-up). This data is benchmarked against a maximum figure of 100% coverage.

Research and Innovation

R&D Spending: This is Eurostat Data , which measures the total amount of R&D spending in a region by all sectors (Government, Higher Education, businesses and non-profits). This is scaled to per hours worked in each region. This data is benchmarked against the OECD target of 3.0% GDP spend on R&D.

Number of researchers: This is Eurostat Data on the total number of researchers working in all sectors in a nation or region. This is scaled to per hours worked in each region. This is benchmarked against the highest number achieved by any nation or region in any year measured.

Finance and Investment

SME Lending: This is cumulative data, gathered from postcode lending data by the five biggest lenders, which are Barclays, NatWest Group (RBS), Santander, Lloyds Bank, and HSBC. In Northern Ireland, NatWest Group and Lloyds are not used as they do not publish data for NI, instead data from Danske Bank and Bank of Ireland is used. This data is scaled to per hours worked in each region. This is benchmarked against the highest cumulative figure in any nation or region in any year measured.

Equity Finance: This is data is from the British Business Bank Equity Tracker and it measures the percentage of equity finance value invested to each nation or region. This data is scaled to per hours worked in each region. This is benchmarked against the highest percentage achieved in any nation or region in any year measured.

Venture Capital: This data is from Technation Data Commons, and it measures the total value of venture capital investment in tech companies in a region, in USD. This data is scaled to per hours worked in each region. This is benchmarked against the highest figure achieved in any nation or region in any year measured.

Trade support

Goods exports: This is HMRC Regional Trade in Goods data, and it measures the total value of exports in goods in nations and regions on an annual and quarterly basis. This is scaled to per hours worked in each region. This is benchmarked against the highest figure achieved in any nation or region in any year measured.

Services Exports: This is experimental ONS data, and it measures the international trade in services by subnational areas of the UK. This is scaled to per hours worked in each region. This is benchmarked against the highest figure achieved in any nation or region in any year measured.