Disaggregating UK local authority-level gross value added to lower levels of geography:1998 to 2020

Office for National Statistics, UK

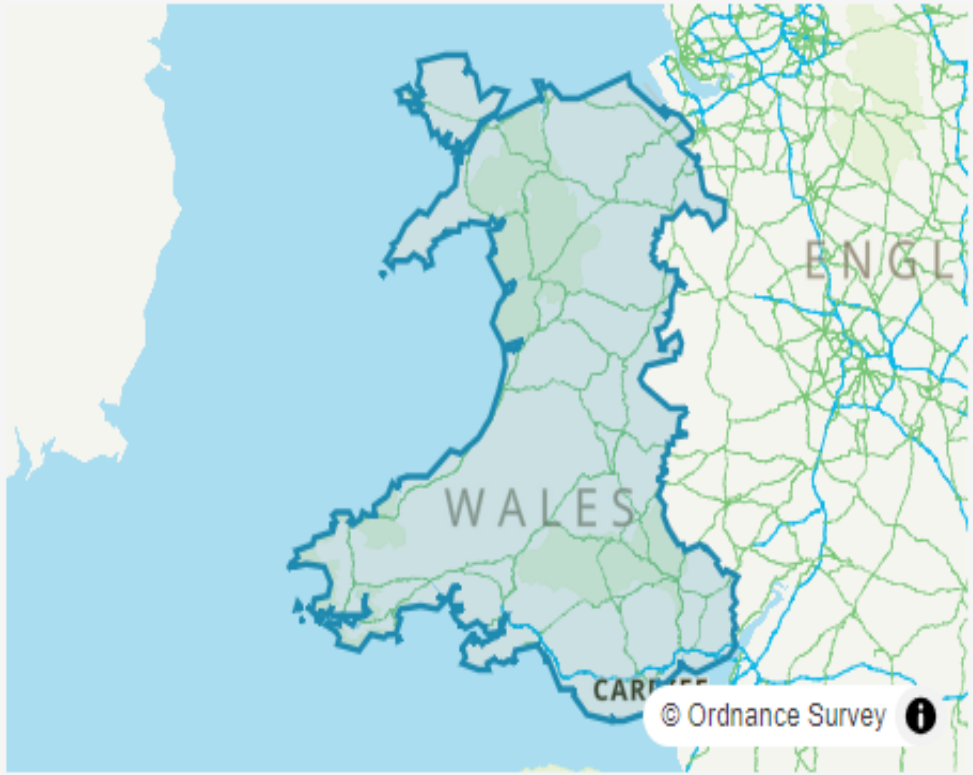
Blessing Chiripanhura

Presentation structure

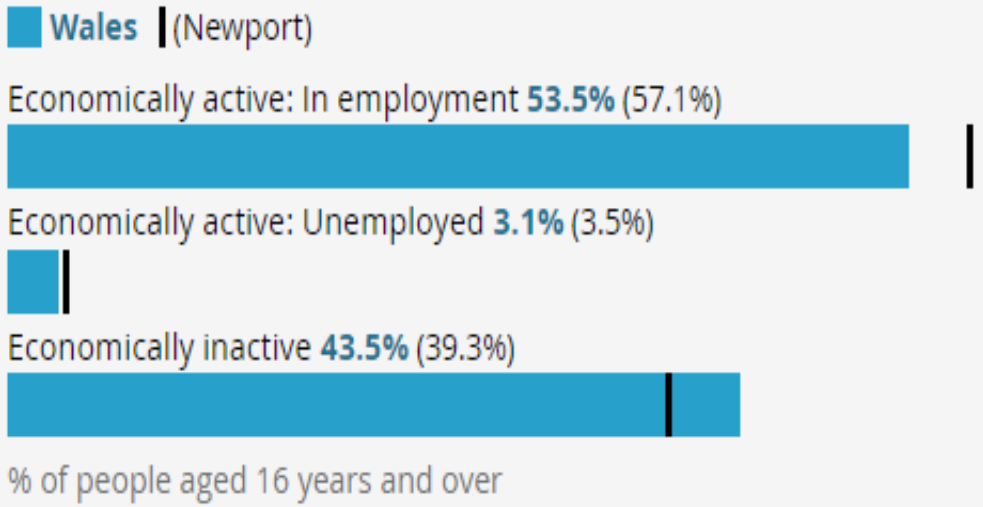
- Why subnational statistics?
  - Illustration: Wales
- Basic requirements for subnational statistics development
- Disaggregating UK local authority level GVA to small areas
  - Pre-processing of apportioning datasets
- Problems of producing subnational data
  - Statistical disclosure
    - Detecting and treating statistical disclosure
  - Outliers
- Key points to remember

Wales

Area map



Economic activity status



Source: ONS Census 2021 data visualisation

Why subnational statistics?

- There is a realisation that national (or higher-level geographical areas) data are too aggregated and tend to mask differences within countries.
- Subnational data provides useful and detailed contextual data that improves our understanding of the social and economic outcomes within regions. Policy makers use such data for resource allocation to where there is greatest need.
- Communities / regions are inter-dependent in many ways and such proximity creates synergies for growth and development. Lack of subnational statistics reduces our understanding of the synergies.

Basic principles for subnational statistics development

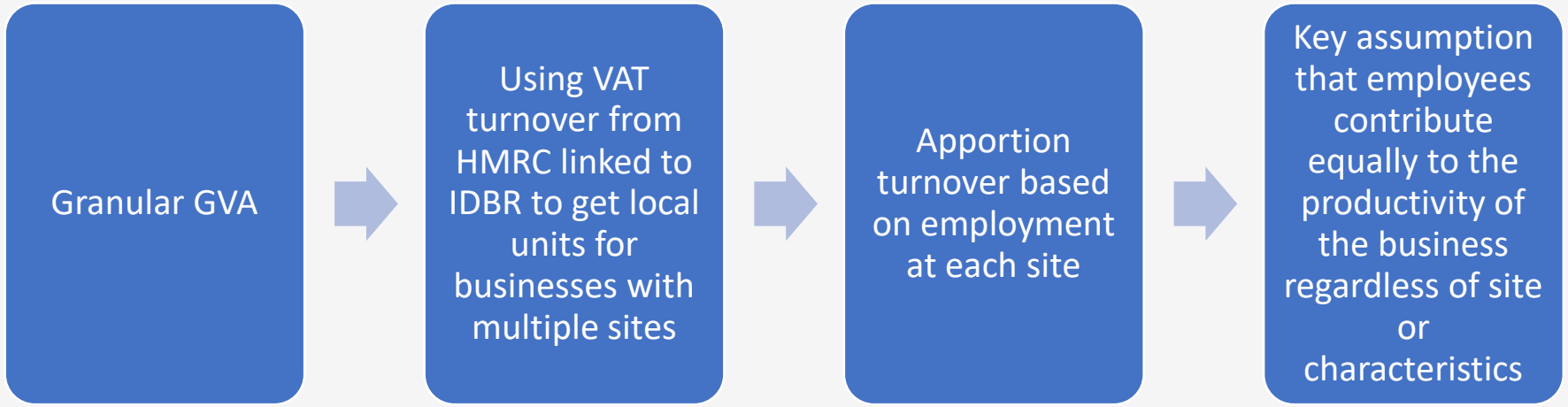
- There is need for a pre-existing national framework within which development is constrained e.g. National Accounts framework
- Secure access to administrative and other proxy data
- Existence of an up-to-date business register
- Devise an approach for dealing with complex business operations operating across multiple sites
- Select the appropriate geographical level
  - We target the lower-layer super output area (LSOA) | data zone (DZ)| super output area level (SOA)

Disaggregating data to lower levels of geography

- In developing the UK gross value statistics for small areas, we use a simple apportionment method.
  - We apportion using administrative and bespoke data sources
  - We apportion the VAT turnover of large and complex enterprises to local units using employee counts at each site
    - This approach assumes equal productivity, which may not be the case

Pre-apportionment data processing

- Data engineering:
  - Data linkage – VAT Turnover data + IDBR
    - We start with data engineering to allocate VAT turnover to enterprises' local units. This is achieved by matching VAT turnover records to the Inter-Departmental Business Register (IDBR) to create a new dataset that is used to allocate VAT turnover to business sites.



We apportion to LSOA, DZ and SOA, which we call 'the building blocks'

We apportion at Section level: reminder of the sections…

Section	Description
A	AGRICULTURE, FORESTRY AND FISHING
B	MINING AND QUARRYING
C	MANUFACTURING
D	ELECTRICITY, GAS, STEAM AND AIR CONDITIONING SUPPLY
E	WATER SUPPLY; SEWERAGE, WASTE MANAGEMENT AND REMEDIATION ACTIVITIES
F	CONSTRUCTION
G	WHOLESALE AND RETAIL TRADE; REPAIR OF MOTOR VEHICLES AND MOTORCYCLES
H	TRANSPORTATION AND STORAGE
J	INFORMATION AND COMMUNICATION
K	FINANCIAL AND INSURANCE ACTIVITIES
L	REAL ESTATE ACTIVITIES
M	PROFESSIONAL, SCIENTIFIC AND TECHNICAL ACTIVITIES
N	ADMINISTRATIVE AND SUPPORT SERVICE ACTIVITIES
O	PUBLIC ADMINISTRATION AND DEFENCE; COMPULSORY SOCIAL SECURITY
P	EDUCATION
Q	HUMAN HEALTH AND SOCIAL WORK ACTIVITIES
R	ARTS, ENTERTAINMENT AND RECREATION
S	OTHER SERVICE ACTIVITIES
T	ACTIVITIES OF HOUSEHOLDS AS EMPLOYERS; UNDIFFERENTIATED GOODS-AND SERVICES-PRODUCING ACTIVITIES OF HOUSEHOLDS FOR OWN USE

Apportioning local authority level GVA to building blocks level

- We start by apportioning GVA for all sections (except O, P, Q, T and part of L) using VAT turnover data:

a) $$\frac{\text{Building block Section } i \text{ VAT turnover}}{\text{LA SUM of Section } i \text{ VAT turnover}} * \text{LA Section } i \text{ GVA} = \text{Building block Section } i \text{ GVA}$$

*(where i = all other sections **except** O, P, Q, T, and part of L (68.2IMP))*

➔ This gives the building blocks GVA for all sections with VAT turnover data.

- Next, we apportion the GVA of sections O, P, Q, T and part of L.

b) **Sections O, P and Q**: $$\frac{\text{Building block Section } i \text{ employment}}{\text{LA SUM of Section } i \text{ employment}} * \text{LA Section } i \text{ GVA} = \text{Building block Section } i \text{ GVA}$$

(where i = Sections O, P and Q)

c) **Section T:** $$\frac{\text{Building block population}}{\text{LA total population}} * \text{LA Section T GVA} = \text{Building block Section T GVA}$$

d) **Section L (68.2IMP):** $$\frac{\text{Building bock dwelling stock}}{\text{LA total dwelling stock}} * \text{LA Section L 68.2IMP GVA} = \text{Building block Section L: 68.2IMP GVA}$$

Checking the apportioned GVA

- After apportioning all industries to building blocks level, we calculate the building block total GVA. That is, sum across industries to calculate total GVA for each building block:

$$\text{Total building block GVA} = \sum[\text{Building block Section } i \text{ GVA}] \quad \text{where } i = \text{Sections A to T.}$$

- Next, we perform a global check of the sum of all building blocks GVA, which must equal the local authority GVA we started with.

$$\sum \text{Total building block GVA in LA} = \text{Total LA GVA}$$

- This summation must equal the Local Authority GVA for a given year. We apply this method for the years 1998 to 2021 to produce GVA time series for each building block.

Points to note about LSOA, DZ and SOA data

- Customisable, flexible, bespoke geographies for analysis.
- Not constrained by published geographies.
- Focus on transport routes, hubs of industry, cutting across pre-defined geographical boundaries.
- Individual LSOA|DZ|SOA should not be compared with one another but aggregated to build larger areas for analysis – used as building blocks.
- We are not able to produce confidence intervals for out series – because of the nature of the datasets involved in apportionment.

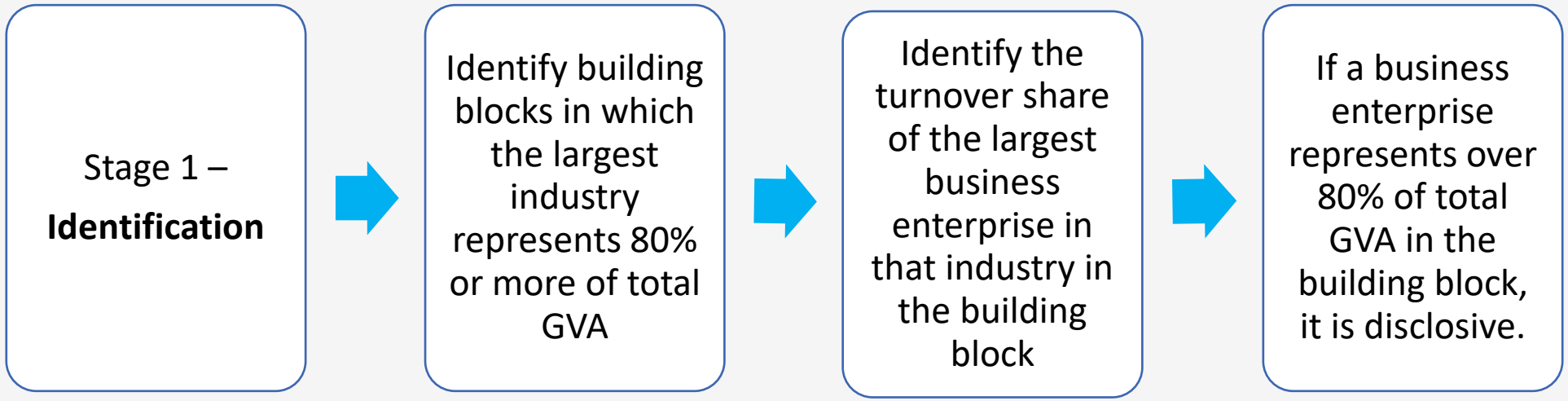
Statistical disclosure and treatment

In small geographical areas, there is a perceived risk of ‘disclosure’ of economic statistics:

Users with local knowledge may ‘guess’ the GVA of dominant local enterprises/businesses

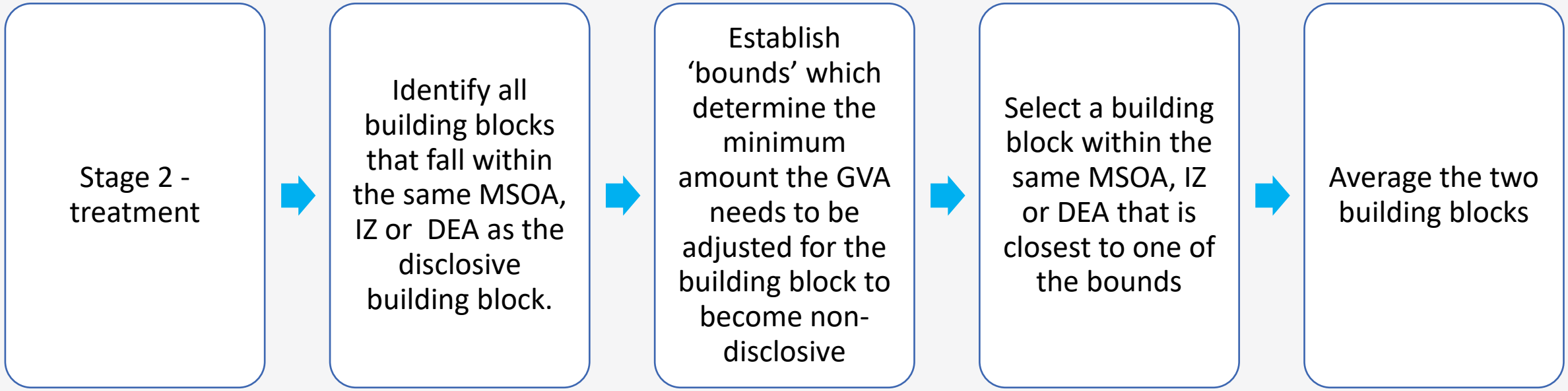
We must apply disclosure ‘treatment’ to adjust values

Testing for statistical disclosure



Treatment options: Suppression
Adding noise to data
Averaging / combining blocks

Treating statistical disclosure



We must choose building blocks to pair within the same MSOA, IZ or DEA so we maintain the correct local authority totals.

MSOA = middle-layer super output area (larger building block)
IZ = intermediate zone (larger building block)
DEA = district electoral area

Sensitivity analysis: Changing the threshold

Uniquely disclosive LSOA/DZ/SOA: 1998 to 2020 and 2012 to 2020

Cut off points/thresholds	LSOA/DZ/SOA count, 1998-2020	Percentage of disclosive LSOA/DZ/SOA	LSOA/DZ/SOA count, 2012-2020	Percentage of disclosive LSOA/DZ/SOA
75%+	3720	8.7	2611	6.1
79%+	2833	6.6	1962	4.6
80%+	2657	6.2	1831	4.3
81%+	2477	5.8	1706	4.0
85%+	1757	4.1	1205	2.8
90%+	1053	2.5	725	1.7

Upper limit: treated LSOA/DZ/SOA = 2657 x 2 = 5314

As a proportion of all LSOA/DZ/SOA: 12.5%

Outlier detection and treatment

- Not all building blocks series are smooth – some have outliers
  - Driven by VAT turnover data
  - Caused by the disclosure treatment.
- We have developed a tool to detect outliers, and to plot the series to aid adjustment.
  - We generate a list of adjustments for each local authority
  - Adjustments are re-apportioned in each local authority, to LSOA|DZ|SOA that were not adjusted.
    - Sum of all building blocks GVA in LA must equal LA total GVA

Key messages

- ✓ Building blocks GVA can allow us to create flexible geographies for analysis
- ✓ Subnational administrative institutions value subnational statistics because they informs local policy formulation decision making
- ✓ Subnational statistics methods continue to development and improve