



Office for
National Statistics

Transforming ONS Household Financial Statistics

Sean White, Head of Household Finance and Inclusive
Growth Development

Matt Corey-Bale, Household Finance Survey Transformation
Lead, Social Survey Transformation Division

Matt Greenaway: Economic Activity, Income and Earnings
Lead, Statistical Design & Research , ONS

Outline

- Transforming the data:
 - HFS development – why/how – where have we got to so far – what's next?
 - Admin data research why/how – where have we got to so far – what's next?
 - ETB/HBAI mapping
- Developing our outputs:
 - Recent Developments to ETB
 - Distributional NA and micro/macro comparisons
 - Small area income estimates / ABIS
- LCF update

Why are we transforming?

New technology

Access to Admin data

Policy interest in income, consumption and wealth

Greater precision of estimates required

Reducing the respondent burden

Transformation programme

- 3 main strands
 - Household finance survey
 - Administrative data development
 - Research projects to develop our outputs

TRANSFORMING THE DATA

1. Household Finance Survey Transformation

What I'll Cover

- Our starting point
- Why transform?
- The HFS model
- Progress to date
- Current work and priorities

Where did we begin?

Living Costs and Food Survey
LCF

5,000 households

CAPI +
Paper diary

Survey of Living Conditions
SLC

12,000 households

CAPI

Wealth and Assets Survey
WAS

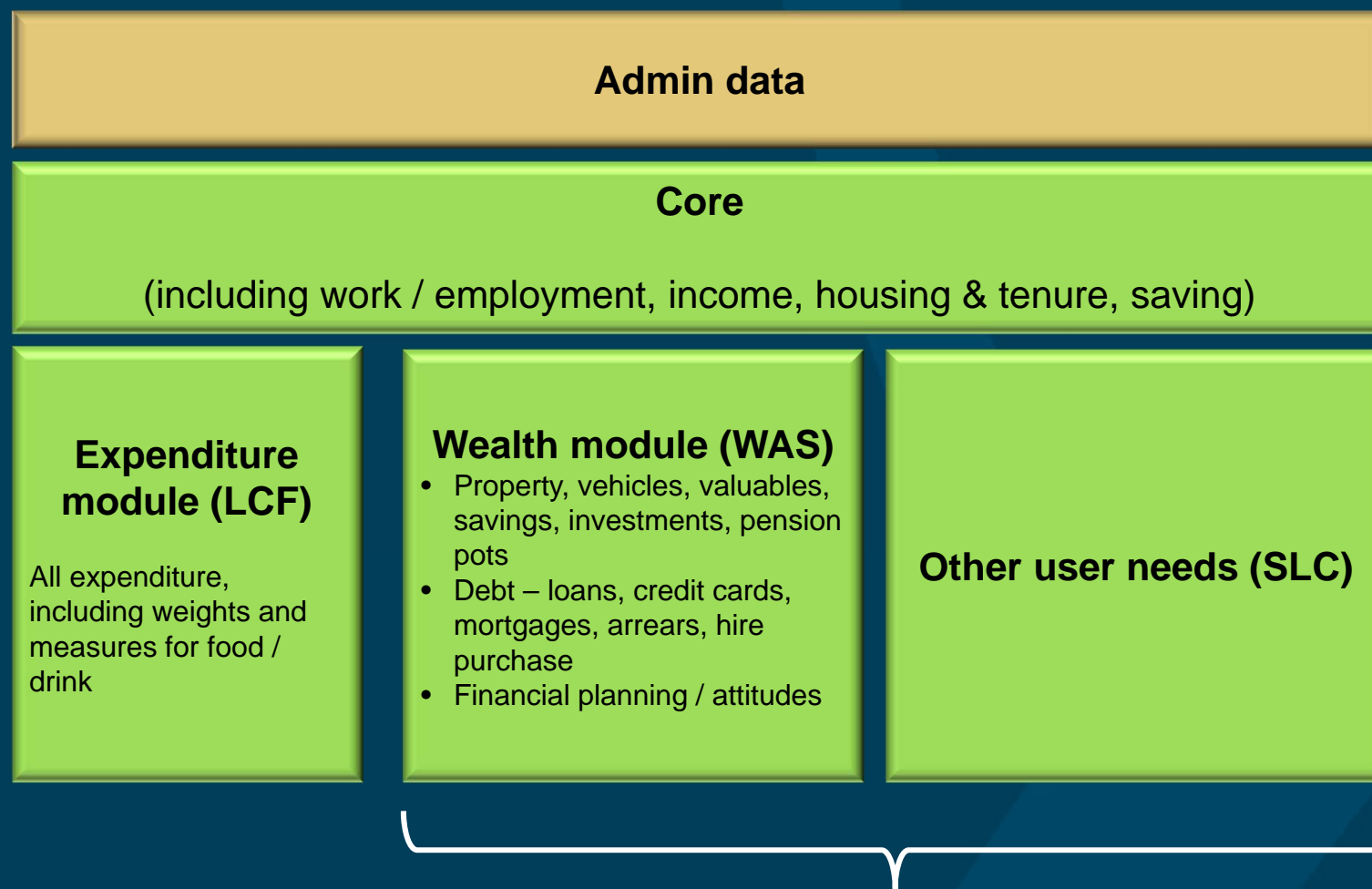
10,000 households

CAPI

Why transform?

- Better technology
- Variety of alternative data sources
- Coherence
- Cost
- Precision

The HFS Model



Data available longitudinally

Progress to date

Sampling

- Integrated sample for LCF/SLC
- LCF, SLC and WAS all now sampled based on a fiscal year

Questionnaire

- Implemented harmonised core for LCF/SLC
- Partial harmonisation for WAS (employee pay, benefits)

Cleaning / processing

- SLC introduced monthly data processing
- Creation of experimental combined HFS dataset (SLC/LCF)

Current priorities

Online Uptake test

- Debate - will the public engage with an online financial survey?
- Launching 28th June
- This will inform our approach to online collection on HFS, with a wider research plan into mixed-mode collection being developed from this

Council Tax data from local authorities

- Aim is to use these admin data as the primary source of Council Tax data
- Developed a research plan to assess the suitability of the data for variable replacement

Current priorities cont...

Sampling frame

- Move to AddressBase
- Reduce the proportion of ineligible addresses sampled
- Again, dependent on research outcomes

Developing admin data research plans

- Can get a lot of benefit from admin data, in particular HMRC PAYE RTI and Benefits datasets from DWP
- Data not in the office yet, but research plans being developed so that we are prepared when it does come in



Office for
National Statistics

2. Admin data development for household finance statistics at ONS

Outline

- ONS ambition for use of admin data
- Update on ONS data acquisition
- Research -
 - Survey Question replacement & Quality Improvement
 - Admin based income statistics

ONS ambition for admin data

- Ambition is to put admin data at the heart of social statistics
- In household finance statistics context, this includes -
 - Developing our surveys - Survey question replacement, improving estimates, and improving sampling
 - Improving small-area income statistics
 - Producing 2021 Census income statistics

Update on ONS data acquisition

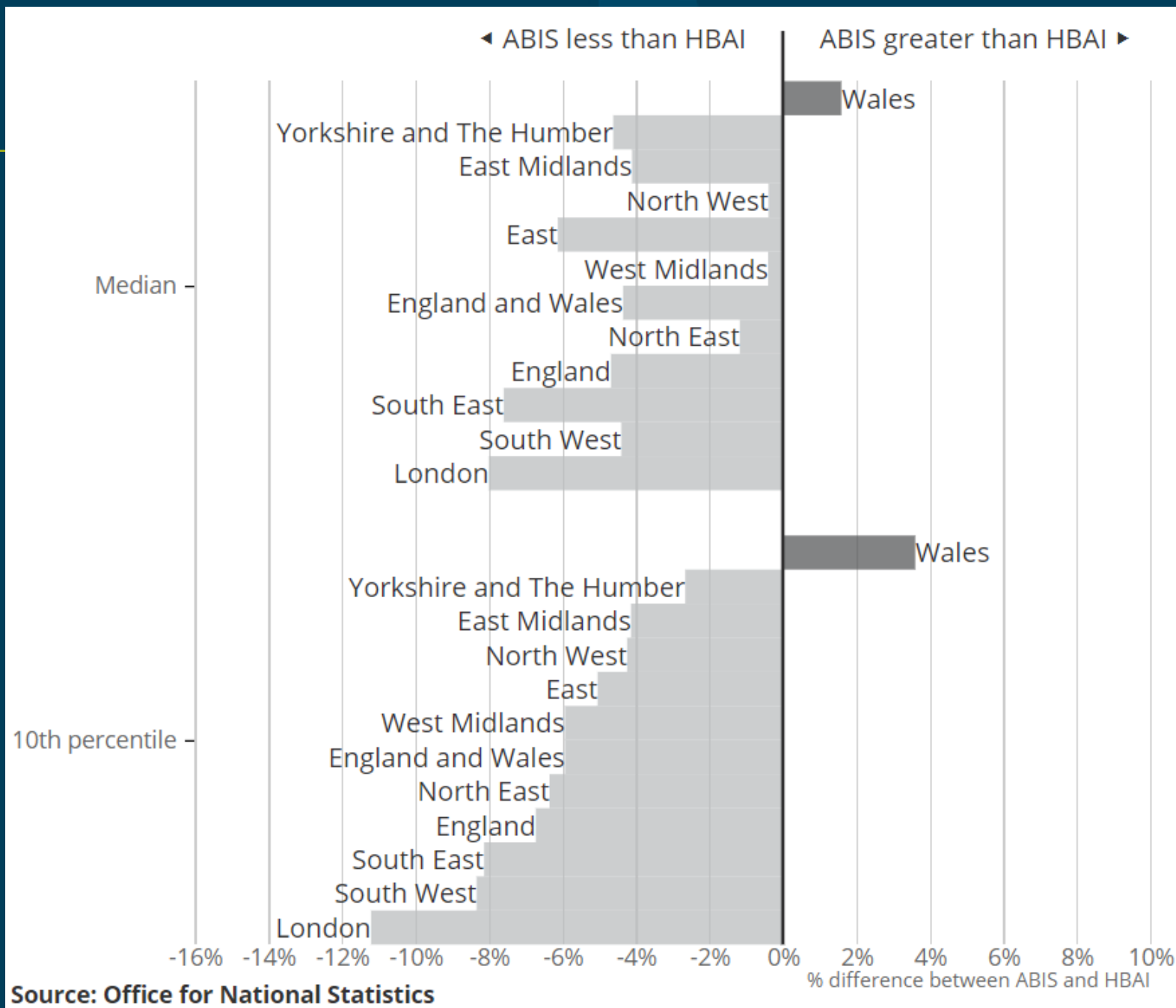
- Council tax data now arriving at ONS on monthly basis
- First full cut of Self-Assessment data from HMRC just arrived at ONS
- Currently negotiating on ONS use of other datasets, including in particular real-time PAYE data ('PAYE-RTI') from HMRC and benefits data from DWP

Research – Developing our surveys

- Desire to shorten household finance surveys, partly to facilitate move online
- Vision is that we will cease asking some questions, and instead using linked administrative data
- Currently conducting research on use of Council Tax data (circa 20 questions on some surveys – including council tax band/amount payed/discounts etc)
- Aim to implement question replacement for council tax questions on Household Financial Survey in 2020
- Admin data offers opportunity to improve the quality of our household financial statistics
- Planning additional research on use of self-assessment data, and other relevant datasets when they arrive
- Supported by concept mapping work

Research: Admin data based small-area income statistics

- Admin-based income statistics – Lower Level Super Output Area level income estimates calculated purely from administrative data.
- Response to user demand for small-area income statistics
- Published update in December 2018
- Various developments – including to equivalisation, publication of percentiles, publication of net (in addition to gross) income, comparison to survey-based estimates
- Still missing some components of income, including in particular self-employed income





Office for
National Statistics

3. HBAI and ETB mapping project

Background

- ONS produce income estimates in their annual 'Effects of Taxes and Benefits' publication
- DWP also publish income estimates in their 'Households below average income' publication
- the Resolution Foundation, in their 'Unequal Results' publication recommended:
- *'that statistical producers do more to explain and quantify how methodological differences between related statistics contribute to the different results; and present advice on whether one measure is preferable to another for a particular use'*

Background: What

- Both publications produce income estimates

- Both publications have national statistics status

- However, each have different uses

- HBAI is a good source for measuring the number of households with low income

- ETB provides data on the full redistributive role of taxes and benefits in kind

Background: Why

New technology

Access to Admin data

Policy interest in income, consumption and wealth

Greater precision of estimates required

Reducing the respondent burden

Aims

- **The project aims to work with DWP to map processes and understand components of the income calculations. We want to highlight the differences and similarities**
- **This analysis aims to contribute towards the co-ordination of sources and outputs so that key concepts are harmonised, and associated publications have a clear purpose and use, minimising duplication and confusion**

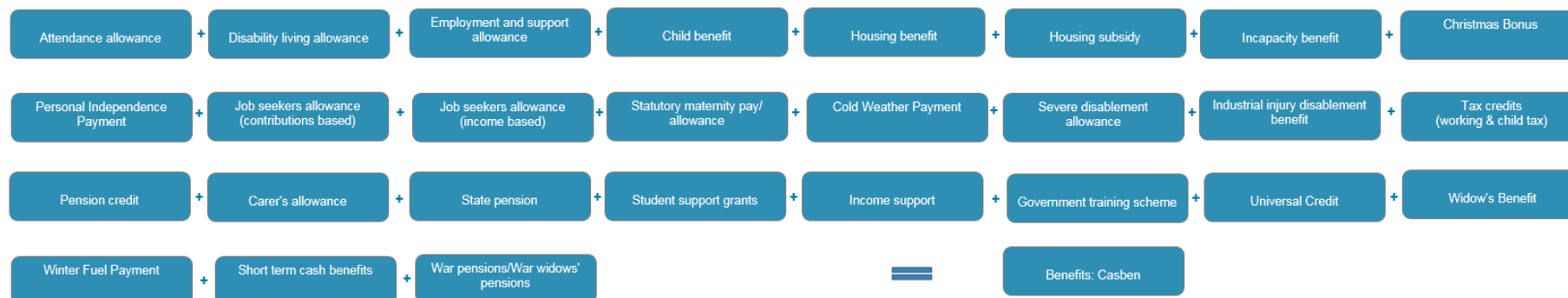
Progress

- Questionnaire:
 - Benefit mapping
 - Income mapping
- Collection details (definitions, coverage etc)
- Editing and imputation
- Weighting
- Analysis procedures

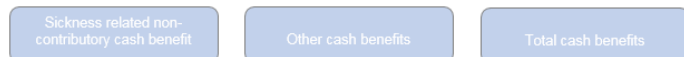
Progress and findings so far: Benefits Mapping

Effects of Taxes and Benefits

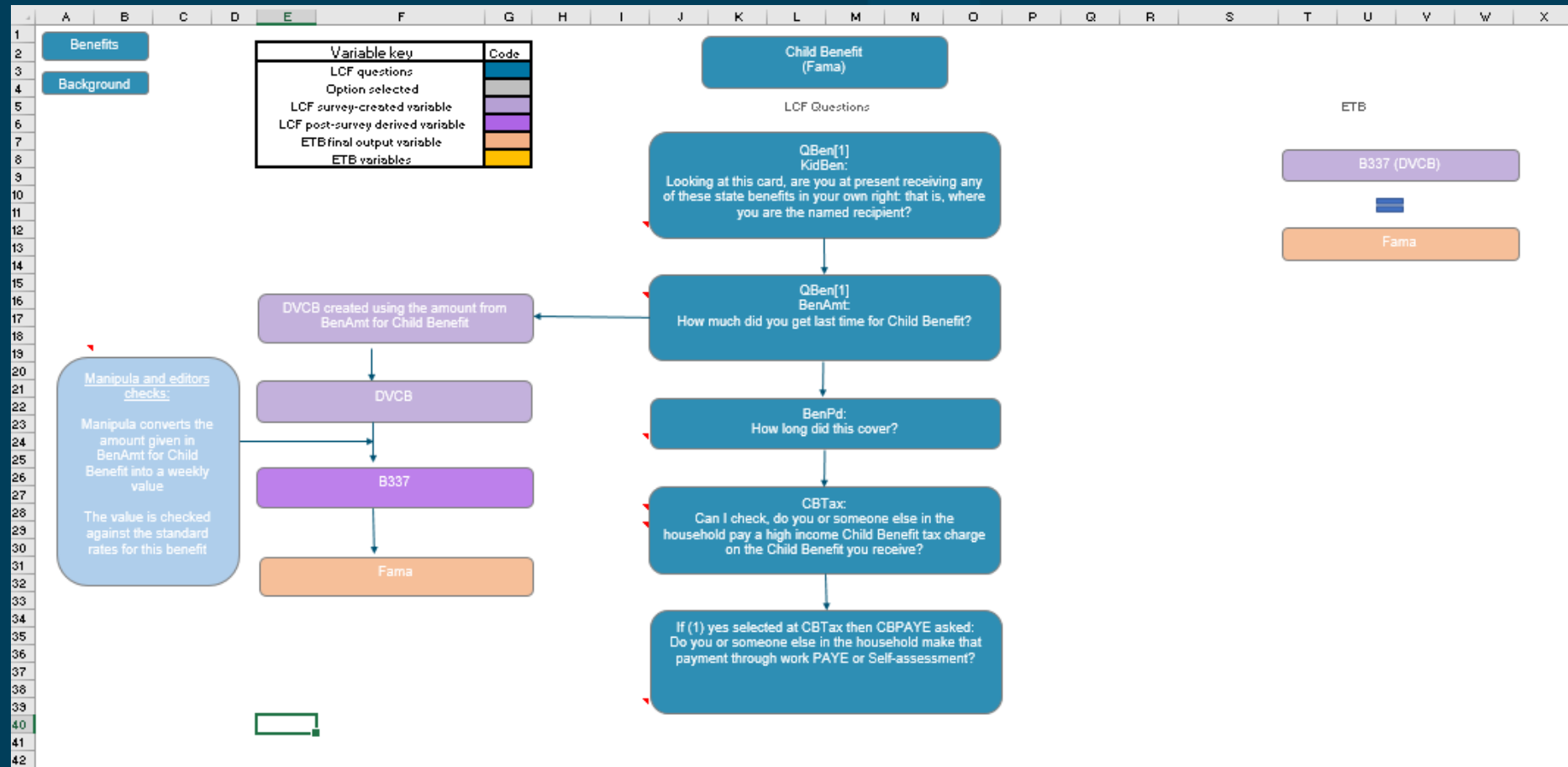
Benefits Mapping



Component variables within ETB, select the boxes below to see the mapping of each:



Progress and findings so far: Benefits Mapping



DEVELOPING OUR OUTPUTS

1. Recent Developments to ETB

ETB Developments

- Using tax data to better capture top earners in household income inequality statistics (presentation this afternoon)
- Switch from measuring household income at a household level, to an individual level
- new measures for the distribution of benefits-in-kind arising from publicly provided adult social care
- Shorter, simpler statistical bulletins

2. Distributional National Accounts and Micro/Macro comparisons

Aims

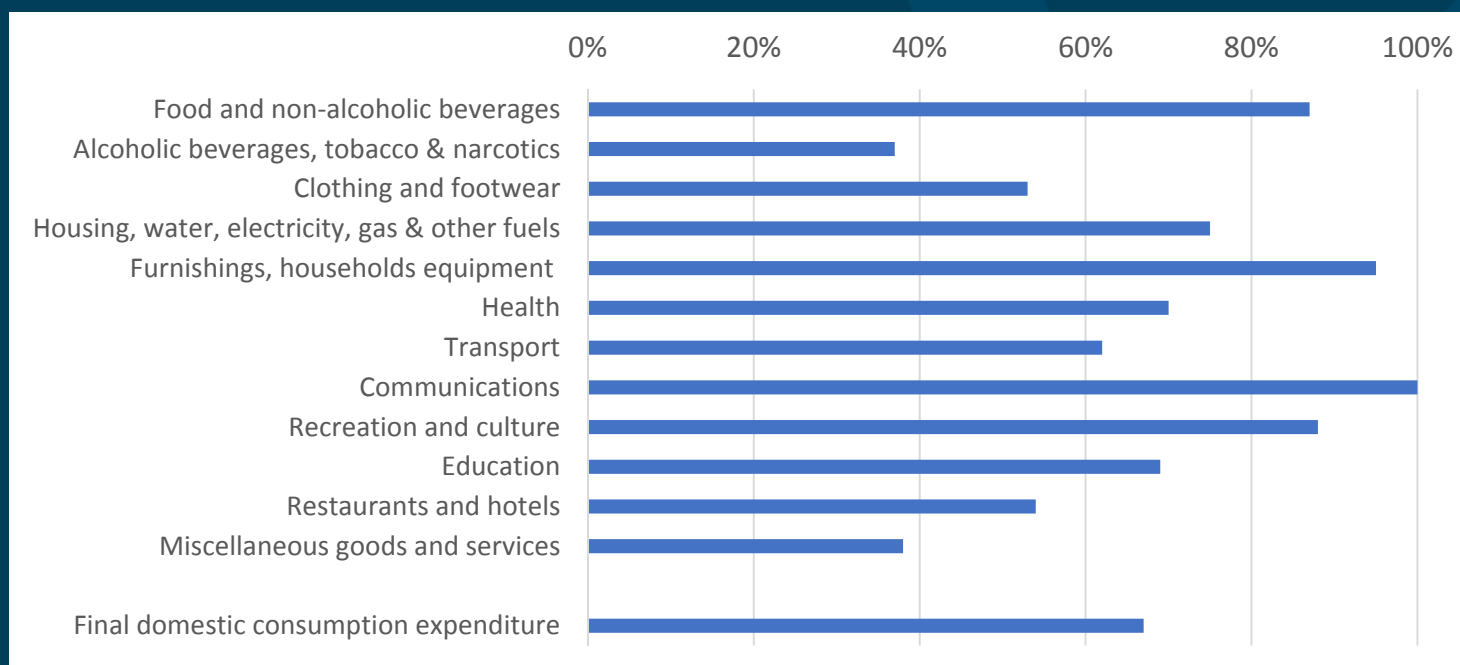
- There are discrepancies in estimates of household income captured in Micro and Macro based estimates from ONS
- Macro – National accounts based estimates
- Micro – estimates based on survey outputs (e.g. LCF)
- Reconciliation of micro/macro allows for production of distributional analysis consistent with national accounts

Distributional analysis within a NA framework

- Why?
 - Provides view of distribution of income and wealth consistent with economy-wide totals.
 - **Potential** for timely monitoring of economic well-being.
 - **It will help assess inclusive growth policies – e.g. how the benefits of GDP growth are shared in society. It will improve coherence for users of our data. Aligning survey data to SNA will also aid international comparison**

Progress

- On the consumption side, we are mapping differences between micro and data at a detailed COICOP level.
- aim is to understand where in the process differences arise.
- The chart presents the disparity in consumption between NA and survey data



Why might micro and macro figures differ?

- Population coverage
 - Micro-statistics based on surveys typically only cover population living in private households.
- Conceptual differences
 - Income and consumption concepts differ in macro (UN, 2008) and micro (UNECE, 2011; OECD, 2013) standards / guidelines.
- Measurement error
 - Evidence of under-reporting in household surveys in different parts of distribution, with some evidence of decreasing coverage (e.g. Crossley & O'Dea, 2010; Brewer & O'Dea, 2012).

Next steps

- Improve imputation of key income variables
- Continue micro/macro mapping exercise
- Produce coherent time series of household (excluding NPISHs) income / consumption/ saving for 2008-2017.
- Develop first NA-consistent distributional wealth estimates for UK.
- Improve coherence of micro / macro data through direct use of administrative data
- User need for distributional NA which are more timely / frequent

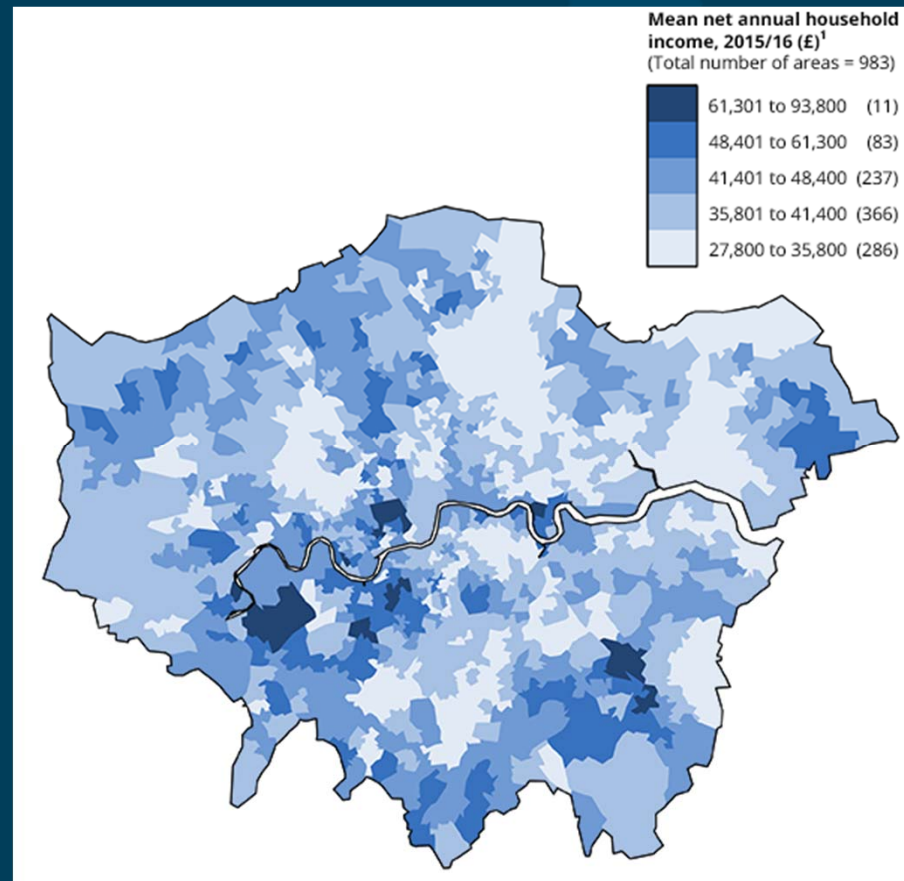
3. Small area income estimates

2. Survey data based small area income estimates

- SAIE are the official estimates of mean annual household income at a local area level and are designated National Statistics.
- SAIE are calculated using a combination of survey data from the Family Resources Survey (FRS), and published data from the Census and various aggregate admin data sources. As the FRS is a national survey, for most local areas there is a small, or even non-existent sample, so a model-based method is used. The models are used to discover which variables in the data relate to income on the FRS at a national level, these variables are then used to predict income at a local level

Example

Model-based mean net annual household income by middle layer super output area, London, financial year ending 2016



LA estimates

- **SAIE are produced at middle layer super output area (MSOA) level. The estimates can be aggregated to LA level but the associated confidence intervals can't be produced, due to the modelling methods used.**
- **Transformation work is addressing this issue when improving the SAIE**

Next steps

- Exploring the use of survey data from the Household Finances Survey instead of the
- Looking at potential for greater use of admin data to help improve the SAIE models
- Examine potential for comparing areas (currently difficult due to Cis)
- Look to incorporate more contextual information e.g. male/female information

LCF developments (Jo Bulman LCF survey manager)

- Questionnaire development
- UN COICOP 2018
- Redesign of paper diary
- Review of incentives
- Receipt scanning; optical character recognition; automation of coding

Questionnaire development

- **Changes for April 2019 focused on ensuring improvements made in 2018 were optimised. Identified opportunities to further reduce respondent burden and improve data quality:**
 - Combined telecoms (TV, Telephone, Internet packages), improvements made to question wording
 - Insurances- improvements to flow
 - Re-ordering of period and amount questions for all topics
 - Review of consumer durables harmonised standard
 - Consolidation of questions within LCF to reduce questionnaire length
 - Proxy measures will be available for users
- **Changes for April 2020 will focus on preparing for implementation of COICOP 2018**

UN COICOP 2018

- Classification of Individual Consumption According to Purpose. Current version implemented in early 2000s
- COICOP 2018 has been endorsed by the UN Statistical Commission and is available for implementation
- Coordinating within ONS and with Eurostat to agree an implementation date. Current working assumption- April 2023
- Changes include:
 - Re-organisation at division level, e.g. 'Information and communication'; 'recreation, sport and culture'
 - Lower level of detail, 5th digit subclass level
 - Separation of goods and services, e.g. delivery charges/fitting to be recorded separately from the good

Redesign of paper diary

- Project timeline: April 2019 to March 2020
- Aims:
 - Update current diary (last revised in 1999); reduce respondent burden & under-reporting; prepare for implementation of COICOP 2018
- Scope:
 - Specification for a new design based on findings from the NatCen Research project completed in December 2017 and requirements for COICOP 2018
 - Qualitative testing of new design with respondents
- Timeline for implementation in the field to be confirmed. Not before April 2021. Seeking funding for quantitative test

Incentive tests

- LCF trial - started June 2018
- Each quota split into 3 groups:
 - £20; £30; £40 (for each responding individual in the household)
- Test which level of incentive offers best value for money and inform longer-term incentive strategy for LCF
- Findings/recommendations prepared by autumn 2019

Digital technology – receipt scanning

- Project ongoing following completion of feasibility study in April 2018. Key findings from the feasibility study ([Survey Methodology Bulletin, May 2019](#)):
 - Concept does have potential to improve the timeliness of LCF diary processing
 - Automatic classification from LCF item descriptions showed promise in improving timeliness.
- Current project timeline:
 - Discovery: Sept – Dec 2018
 - Delivery: Mar – Dec 2019
- International collaboration
 - Feeding into a Eurostat funded project in collaboration with Statistics Netherlands focusing on development of a SmartPhone App.
- ONS focus- improving in-house processing of the paper diary
 - Development of machine learning techniques to improve efficiency of current processes
- Application of methods in an on-line context to be considered in the future

Contacts

- Sean.white@ons.gov.uk
- Matt.bale@ons.gov.uk
- Matthew.greenaway@ons.gov.uk
- Jo.bulman@ons.gov.uk