

## Regional Workshop on Editing (Data Cleansing) of Statistics produced from Administrative Registers

### Using Survey Data to help Validate and Correct Administrative Based data

Nancy McBeth  
GCC-Stat



حاصل علي شهادة الأيزو 27001  
في أمن المعلومات



# Surveys

- Designed for statistical purposes – measure statistical concepts
- Information provided directly by respondents
- Can measure elements not in registers, eg Relationships and family information
- Can measure some concepts quickly – eg Main Activity/Labour force Status
- But
  - Sampling and Non-sampling Errors



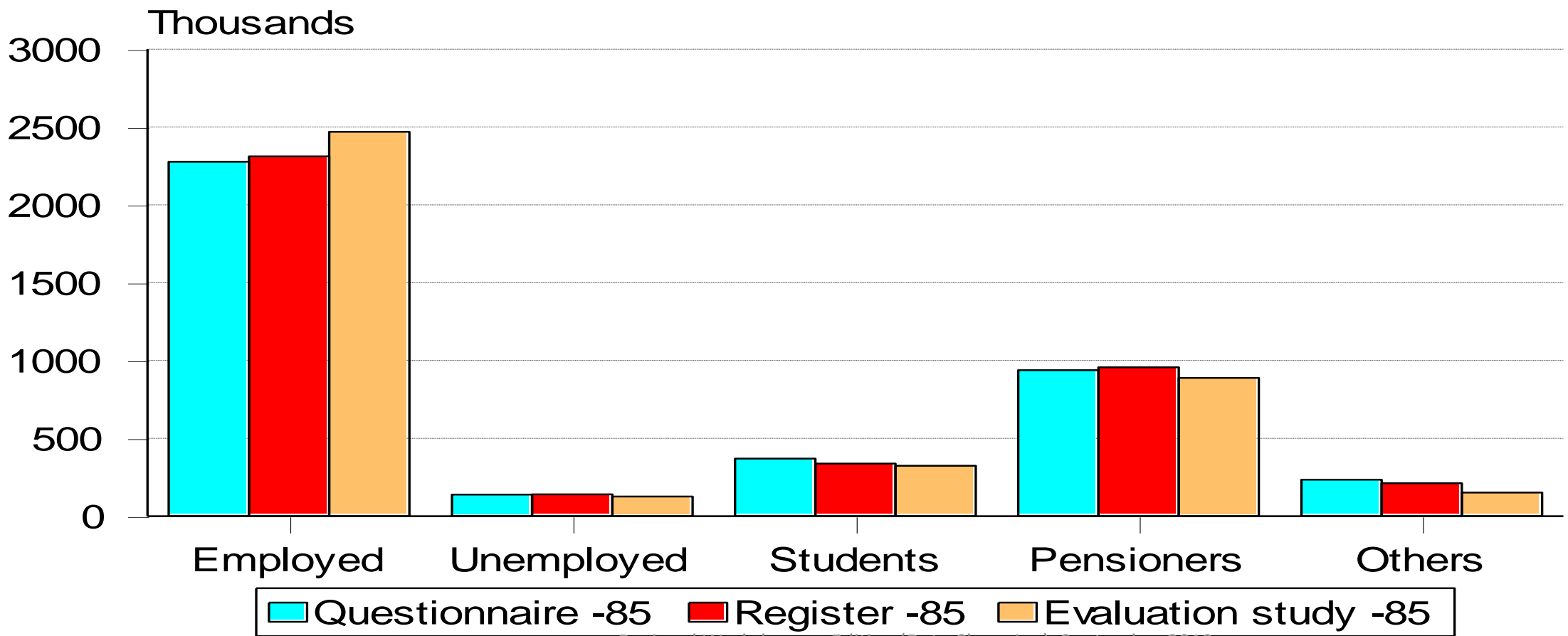
# Using Survey Data

1. Feasibility assessment of using /changing administrative sources
2. Check quality of registers
3. Compare figures at aggregate level
4. Compile data linked between surveys and statistical registers



# 1. Feasibility assessment

Statistics Finland, Initial Assessment of Administrative data 1985  
- Main type of activity by survey and register



## 2. Quality Checks through targeted survey



### Statistics Finland - Checking data in registers

- Once a year, Labour Force Survey contains question on quality and accuracy of address data in Population Register
  - 96.8% in 1998
  - 98.4% in 2010

Other topics – eg occupation, tenure – checked less frequently



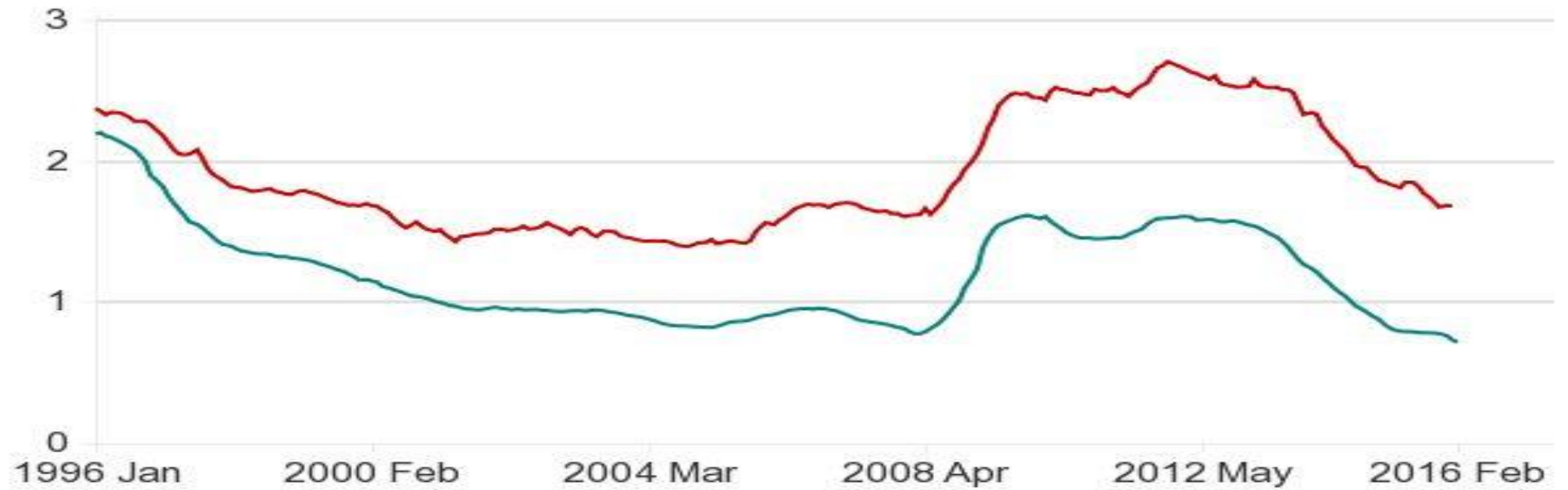
# 3. Comparing Distributions and Totals

## Unemployment in the UK

### Jobless total

Unemployment and claimant count in the UK, millions

Claimant count    Unemployment



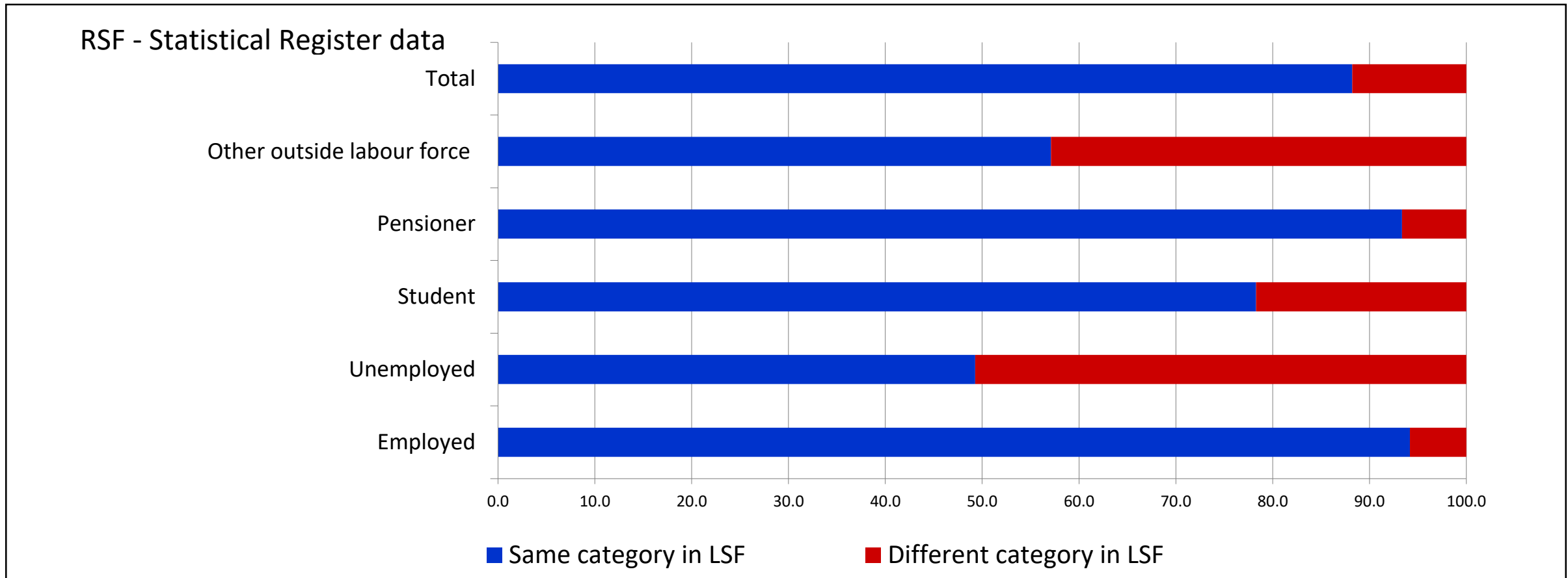
Source: Office of National Statistics



# 4. Analysing Linked Data

## Statistics Finland - Comparing Main Activity – between Survey and Statistical Registers

*Percent classified in the same and different categories as Labour Force Survey (LSF)*





# Using Labour Force Data to Assess Administrative Data - Example



**Theoretical Example**

	Administrative data source	Labour Force Survey
<b>Description</b>	Government recruitment register – from 2017	Quarterly survey – rotating panel of 20,000 citizens and non-citizens Available from 2013
<b>Coverage</b>		
Unemployed	Citizens registered with Job Search agency	Citizens and Non-citizens who meet ILO definition (Available and looking for work)  Job Search Methods – include Registered with Job Search Agency
<b>Punctuality</b>	3 months after end of year	Available 2 months after quarter

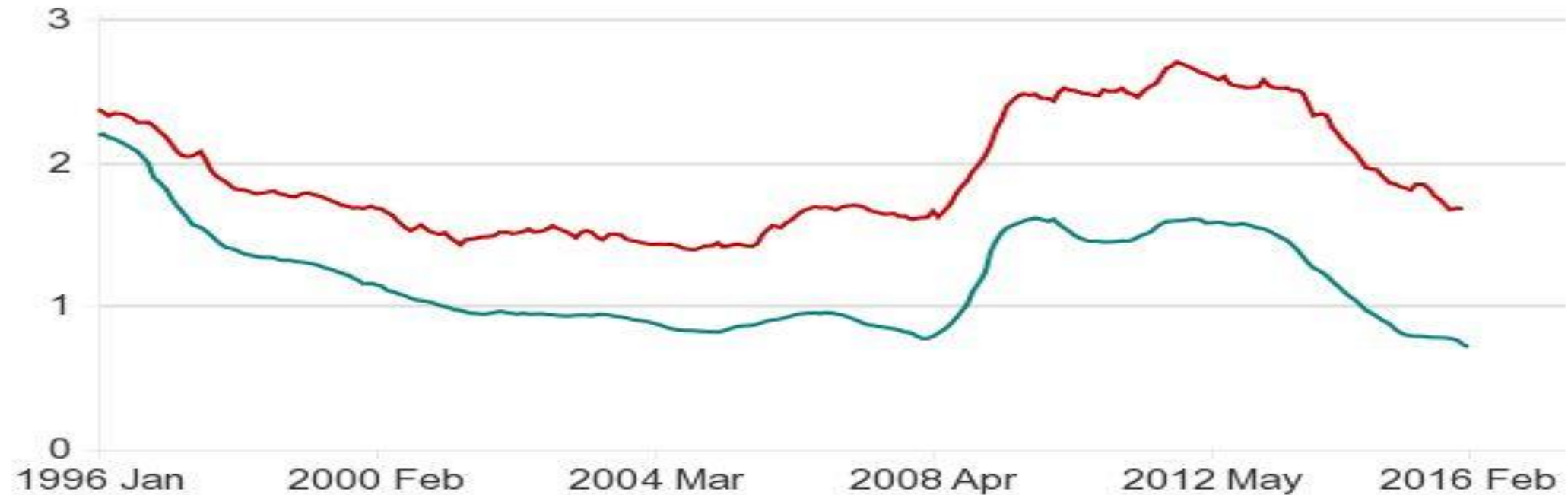


# Plot data over time if possible

## Jobless total

Unemployment and claimant count in the UK, millions

— Claimant count — Unemployment



Source: Office of National Statistics



# Possible areas for differences

- Time Differences
  - LFS – Average for Quarter
  - Register – continuously updated
- Conceptual differences
  - LFS – Not in work, Actively Looking and Available
  - Register – Only people eligible for support (benefits), etc
- Accuracy issues
  - LFS – People forget they have registered - or think they have registered when they have not
  - Proxy effect
  - Register – Can include people who are no longer actively looking for work

## Sampling Error



- **Importance of Understanding Sources – Metadata**
- **Statistical perspective**



# Can survey data be used to correct individual administrative data records?

- Statistical Register - Yes – but
- Administrative Register - No

### Be statistically confident you have the correct unit

- Same ID numbers on both the survey and administrative record, or
- Using agreed statistical criteria, confident that likely that records are the same

### If you plan to use unit record Survey Data with Administrative Data

- Include the ID number on the survey
  - eg Personal ID Number on Household Surveys
  - Business ID numbers on Business Surveys
- Or develop statistical rules/criteria for linking/matching records

**and**      Check the correction is valid –don't introduce errors



## Correcting Unit records on Administrative Registers

### Correction to Administrative Register – **No**

Adjustments for statistical purposes can have unexpected administrative consequences





# Summary

- Survey data – important role in helping to validate and correct administrative
  - Feasibility assessment
  - Check quality of registers
  - Compare figures at aggregate level
  - Compile data linked between surveys and statistical registers
- Preparation is key
  - Including identifiers in survey
  - Understanding the differences between survey and administrative data



## Sources

- Use of Survey Data (LFS) to Evaluate the Quality of Register-based Census in Finland – UNECE-Eurostat Expert Group Meeting on Censuses Using Registers, 2012 -  
[https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2012/use\\_of\\_register/WP\\_11\\_Finland.pdf](https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2012/use_of_register/WP_11_Finland.pdf)
- Unemployment and Claimants Count  
<https://www.ons.gov.uk/file?uri=/employmentandlabourmarket/peoplenotinwork/outofworkbenefits/methodologies/claimantcount/unemploymentandtheclaimantcounttcm77327614tcm77387938.pdf>.



# Questions

# Thank you

