



Office for
National Statistics
Swyddfa
Ystadegau Gwladol

Bringing Data Together at the UK Office for National Statistics

Mark Martin

Business Change Manager

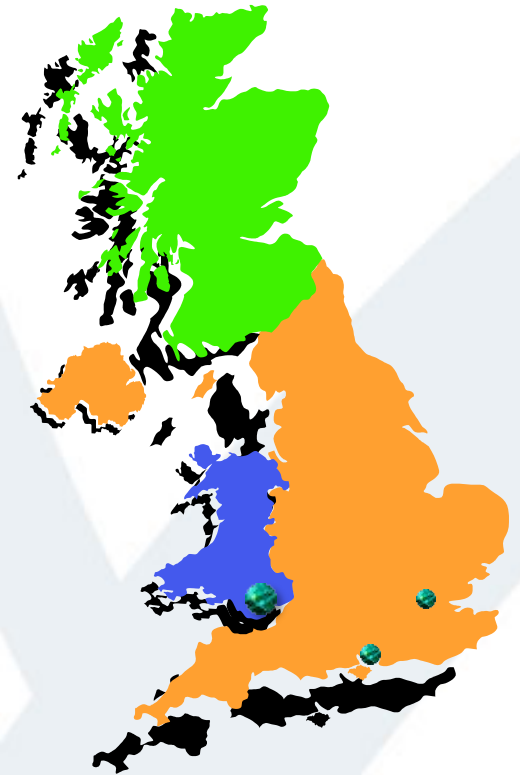
Electronic Data Collection Programme

Overview

- Who are ONS
- Where we are now
- Our Vision for data collection: where we want to be
- Challenges/Drivers
- Some experiences of Administrative Data Collection

The Office for National Statistics

- The UK's main provider of official statistics
- What's important to us:
 - Trusted
 - Respected
 - Independent
 - Efficient
 - A minimal burden



ONS's Mission Statement

Trusted Statistics - Understanding the UK

Where we are now – Our questionnaire tools



Field Force – Used for social surveys to seek qualitative information

£ £ £ £ £ £ £ £ £ £ £



Telephone Interviewing – mostly used for social surveys

£ £ £ £ £ £ £ £ £ £



Paper Questionnaires – The main collection method for business surveys and widely used for social surveys

£ £ £ £ £ £ £ £



Online – ONS has conducted a few experiments, but no large scale collection.

£ £ £ £ £



Secure Electronic File Transfer (SEFT) – allows businesses to send complex or large returns using spreadsheets

£ £ £ £



Telephone Data Entry (TDE) – Useful for short surveys asking for numerical information

£ £ £

Non questionnaire based tools



Administrative Data from other Government bodies

£



Future Technologies



Direct Data Feeds – Businesses provide automated reports

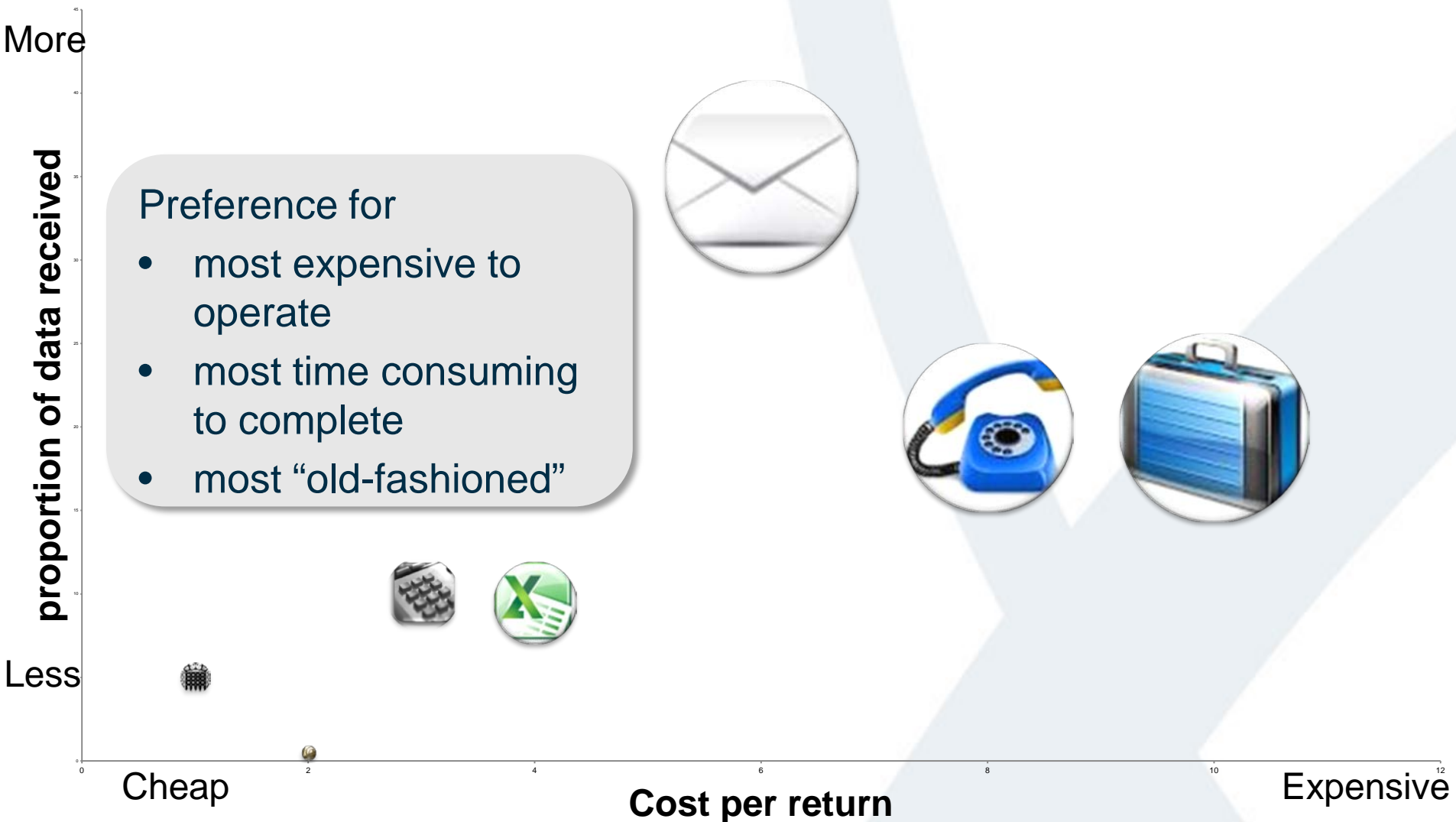
£



Bespoke Questionnaires

£ £ £ £

Where we are now – Which data collection modes do we use?



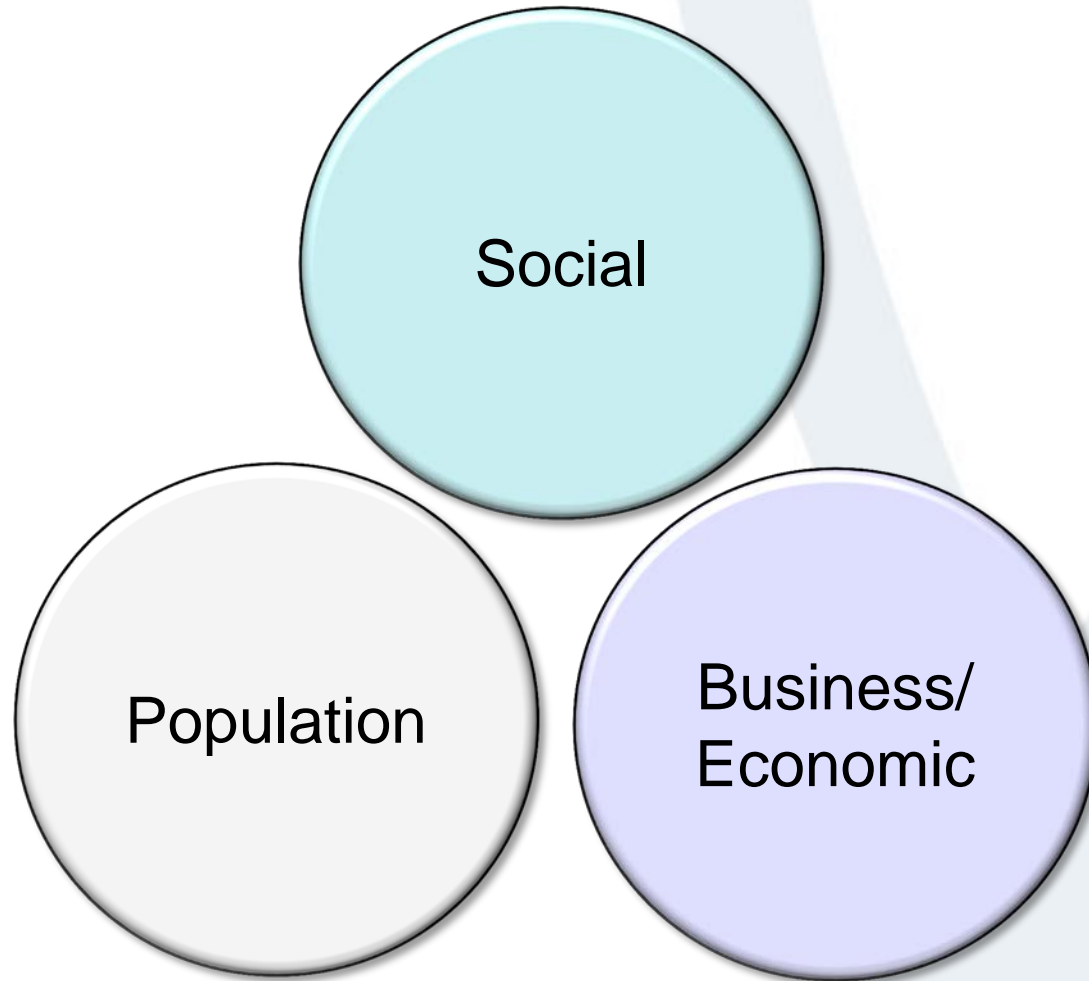
Where we are now - Structure

- Why talk about ONS's internal structure?

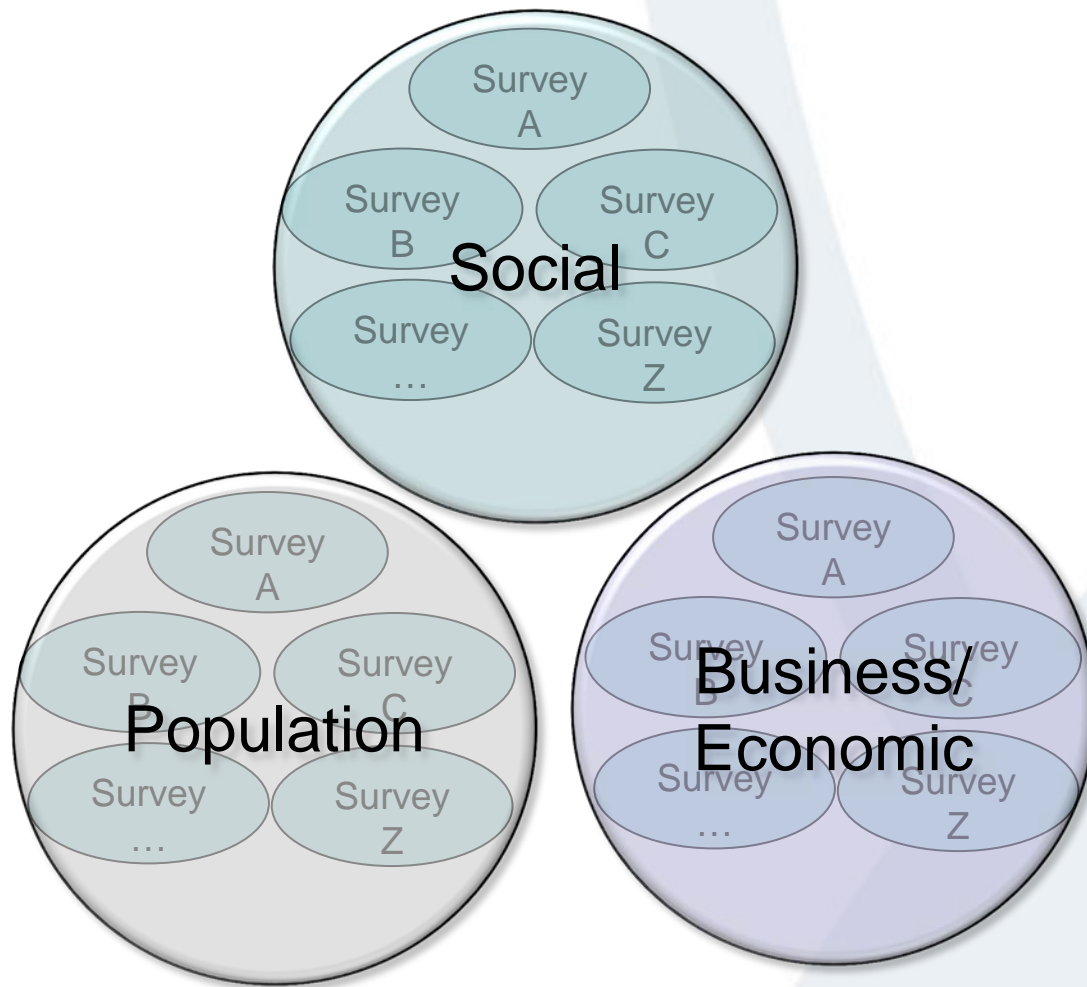
Many government services rely on digitised versions of pre-digital business processes, layered on top of legacy IT systems, some of which are over 30 years old. They were not designed with a digital service in mind, being built to replicate paper forms and processes rather than taking advantage of opportunities to pre-populate or respond to user's selections.

(UK Government Digital Strategy, 2012)

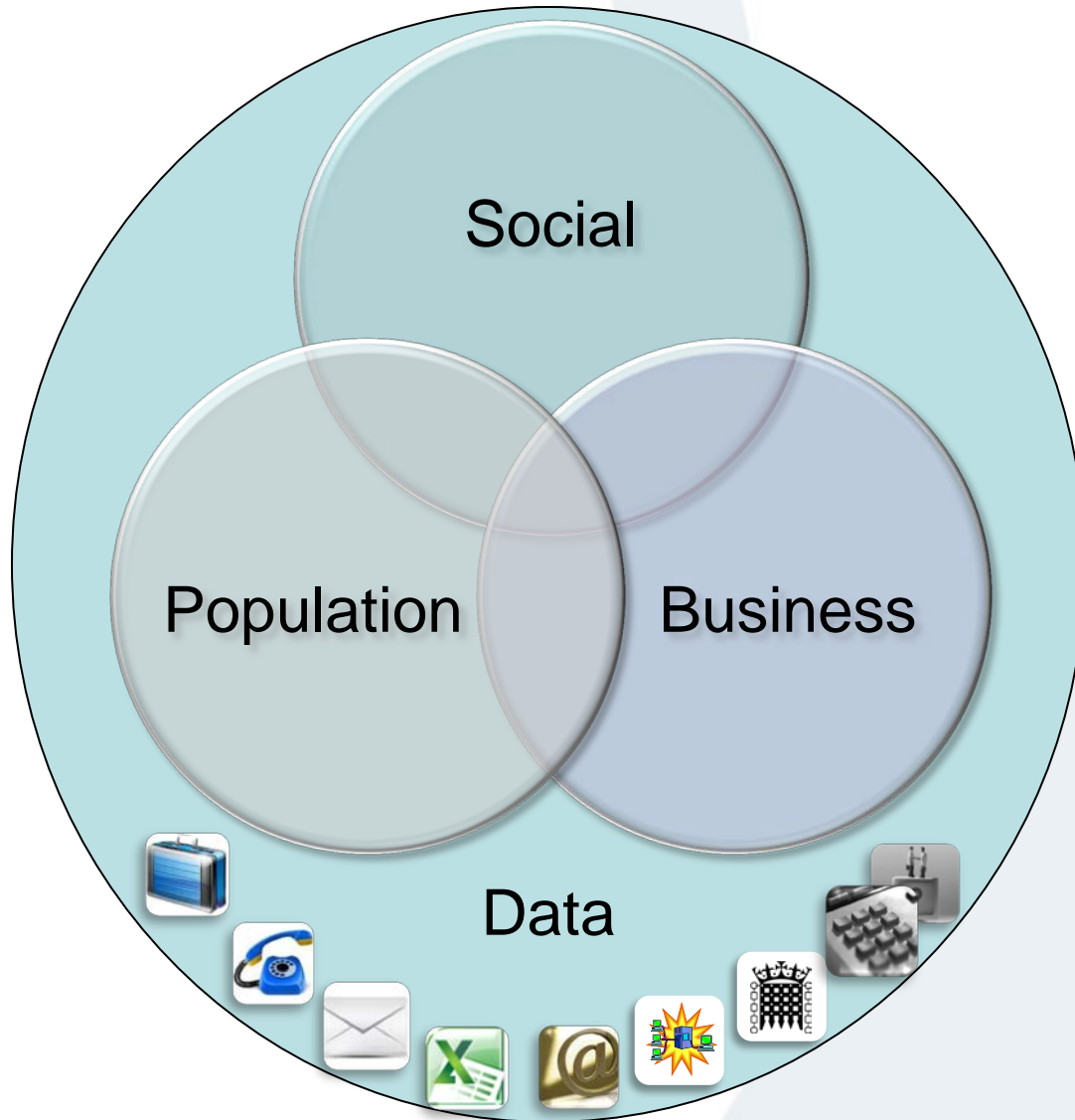
Where we are now - Structure



Where we are now - Structure



Where we are now - Structure

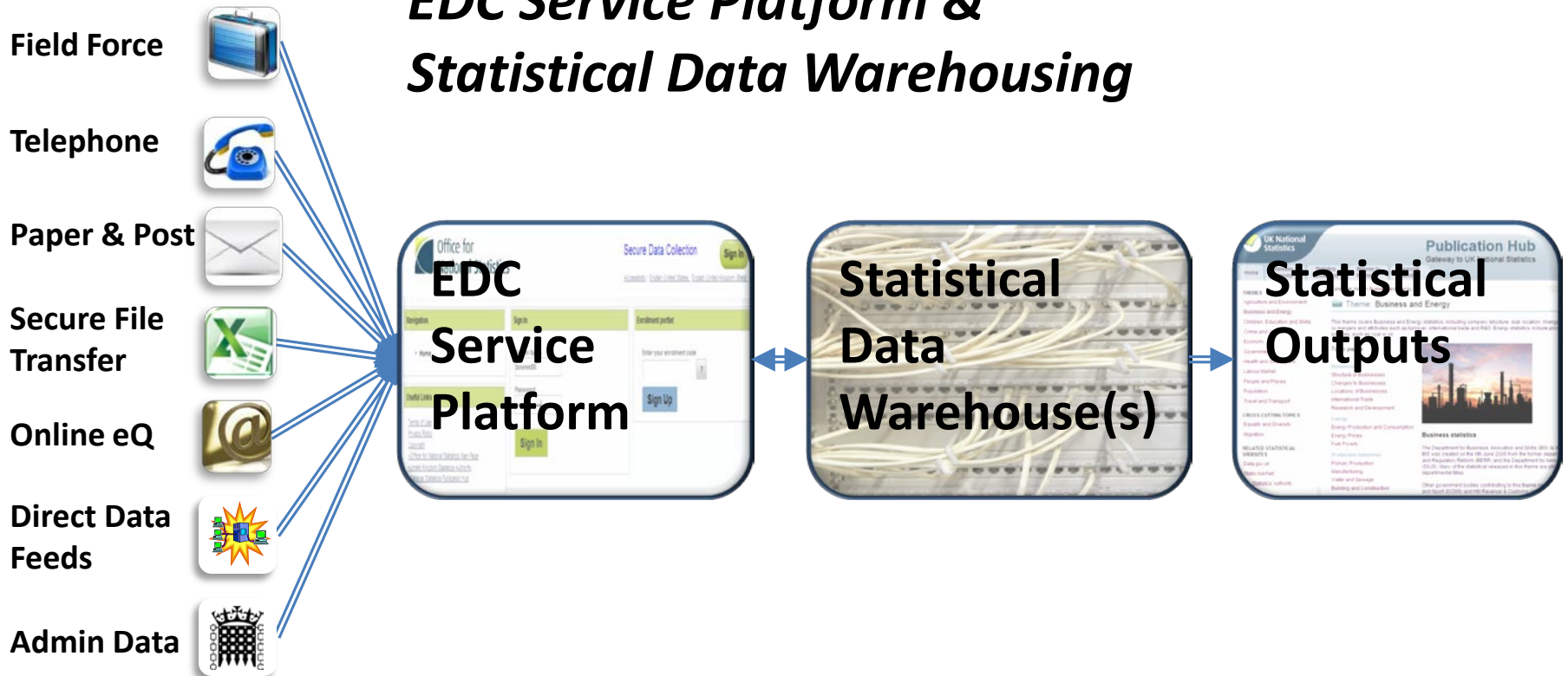


Structure Summary

- Internal Barriers
- Lots of surveys managed separately
- Based on questionnaire/survey paradigm
- Not really compatible with future forms of data collection

EDC Programme Vision

EDC Service Platform & Statistical Data Warehousing



- Multiple, prioritised electronic collection modes and channels
 - Hierarchy based on efficiency
 - Provide choice and flexibility to respondents
- Integrated & coordinated via EDC Service Platform + Statistical Data Warehouse(s)
- Underpinned by a Common Information Model i.e. common concepts, definitions, data models that draw on existing standards

A few specific drivers for ONS

- **80%** of Business Users would prefer to provide data electronically
- Over a quarter of Internet users (**27%**) have submitted official forms electronically. (ONS data – 2011).
- HMRC & Companies House moving to mandating of online data collection.
- Government Digital Strategy: “**Digital by Default**”
- Customers for social surveys are increasingly demanding an online capability – and ONS is losing business

Our Mission (again)

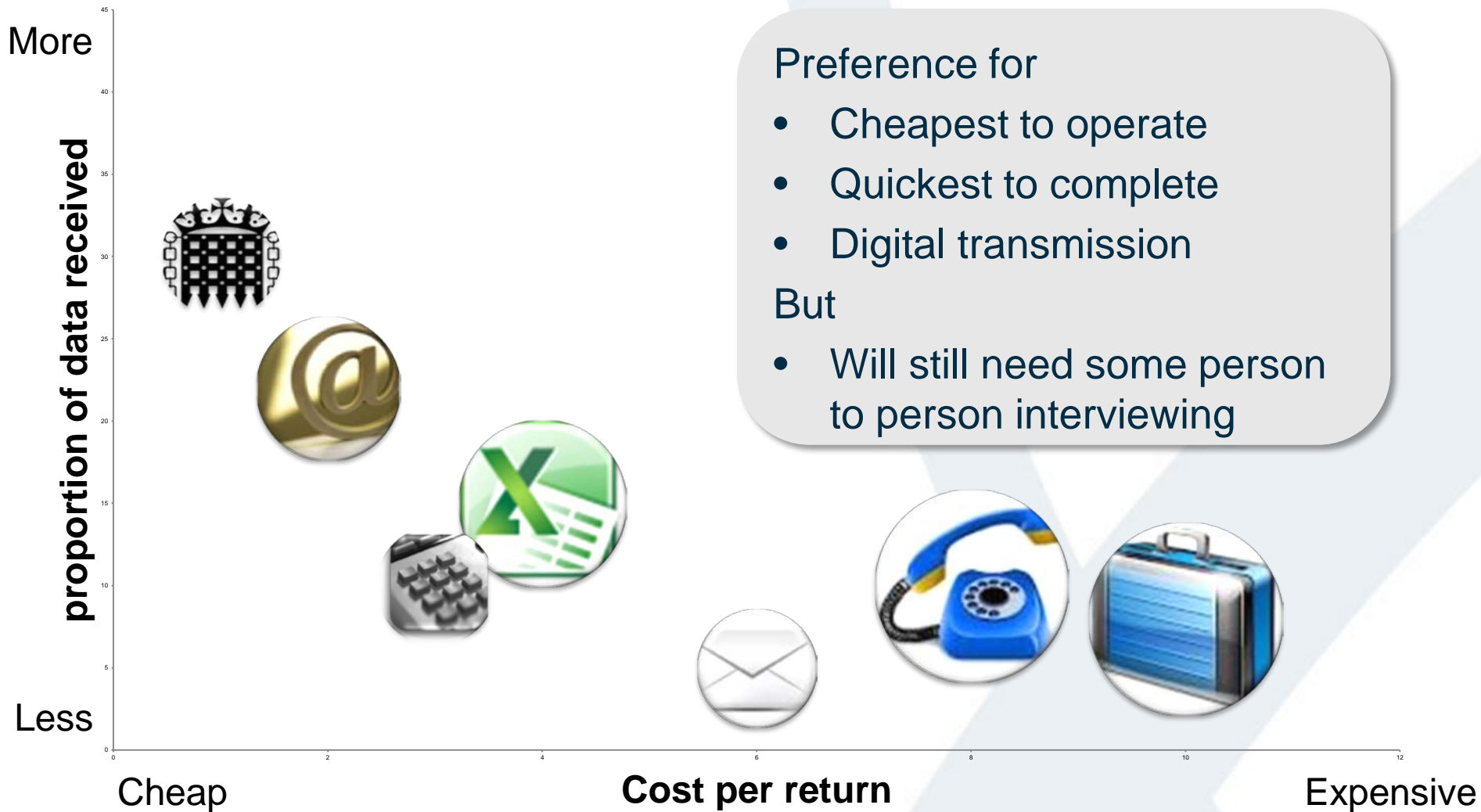
Electronic channels have the potential to improve the quality and especially the timeliness of our data

Trusted Statistics - Understanding the UK

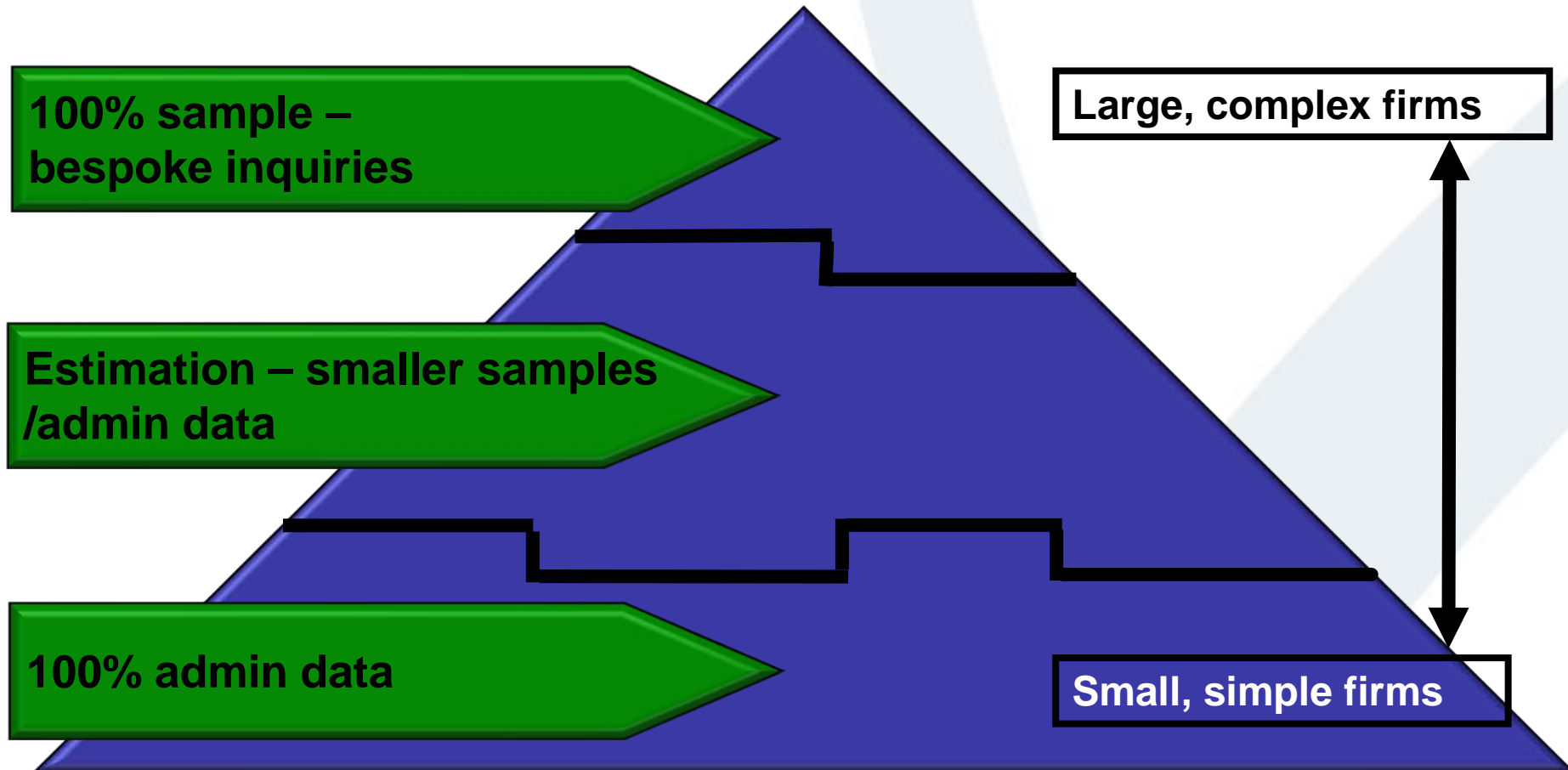
Appearing old fashioned jeopardises this trust

Can't claim to understand a complex thing like the UK if we're not engaging appropriately

Where we want to be



Integration of Admin Data (Business Surveys)



Challenges

- **Technology:**

Data Warehouse – big and complex. Working with other European NSI's to understand implications. Not ready to build it yet.

Require complex identity model which goes beyond standard user name/password concepts

- **Methodology:**

Admin Data – not collected for statistical purposes, so need to understand impact fully.

Matching datasets

Modal impacts not fully understood. Running some pilots to better understand this

- **Legal issues** – often need legislation to share data with other Government bodies.

Drivers - External

- **Social Changes**

People move about more

“death of deference”

More atomised – more households, more self employed etc

Concepts of identity

Blurring of boundary between work and leisure

Fears of “big brother”

Expectations of data sharing

- **Technological Changes**

Mobility

Internet ubiquity

Enterprise Resource Planning

Speed of transactions– eg: Vocalink

Ease of transmission and storage of data

Administrative Data – Some Experiences

- Population Estimates
- Short Term Enterprise Statistics

What ONS has been able to access

Since ONS became independent of the UK government in 2007 , new legislation has been passed to allow access to 5 new data sources:

- HESA Student data
- English School Census data
- Welsh School Census data
- Customer Information System (social security payments made and benefits received)
- VAT purchases data

Legal obstacles, and matching difficulties

- In the UK there is no legislation that gives ONS access to all data owned by other government departments.
- While there are some common identifiers used within government departments, there is no single identifier used by all departments, either for enterprises or for people.

What have we learned?

- No single administrative source will provide all the required information.
- When comparing sources it is important to take coverage and precise definitions into account

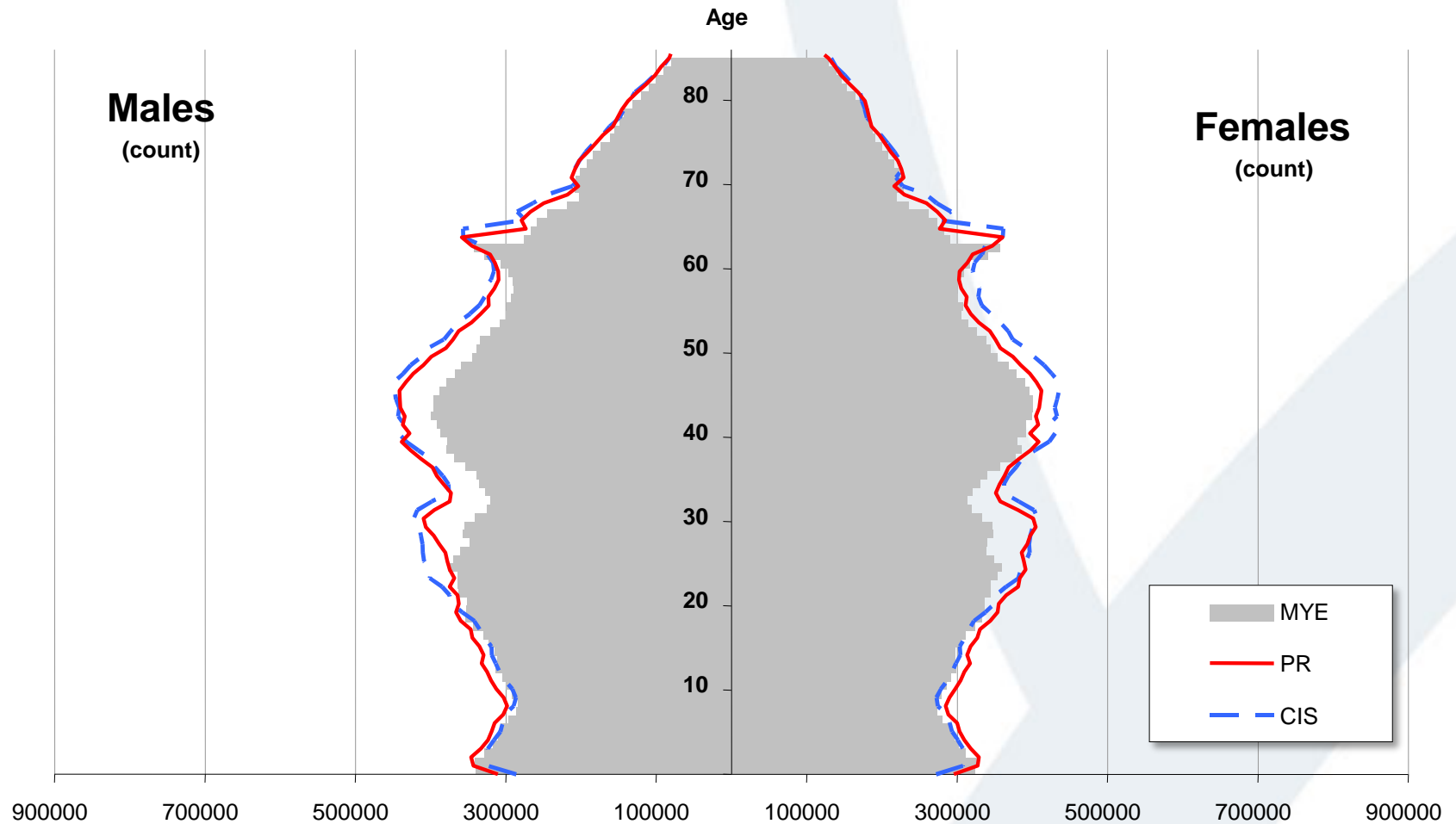
Population Estimates

The following slides compare population estimates from:

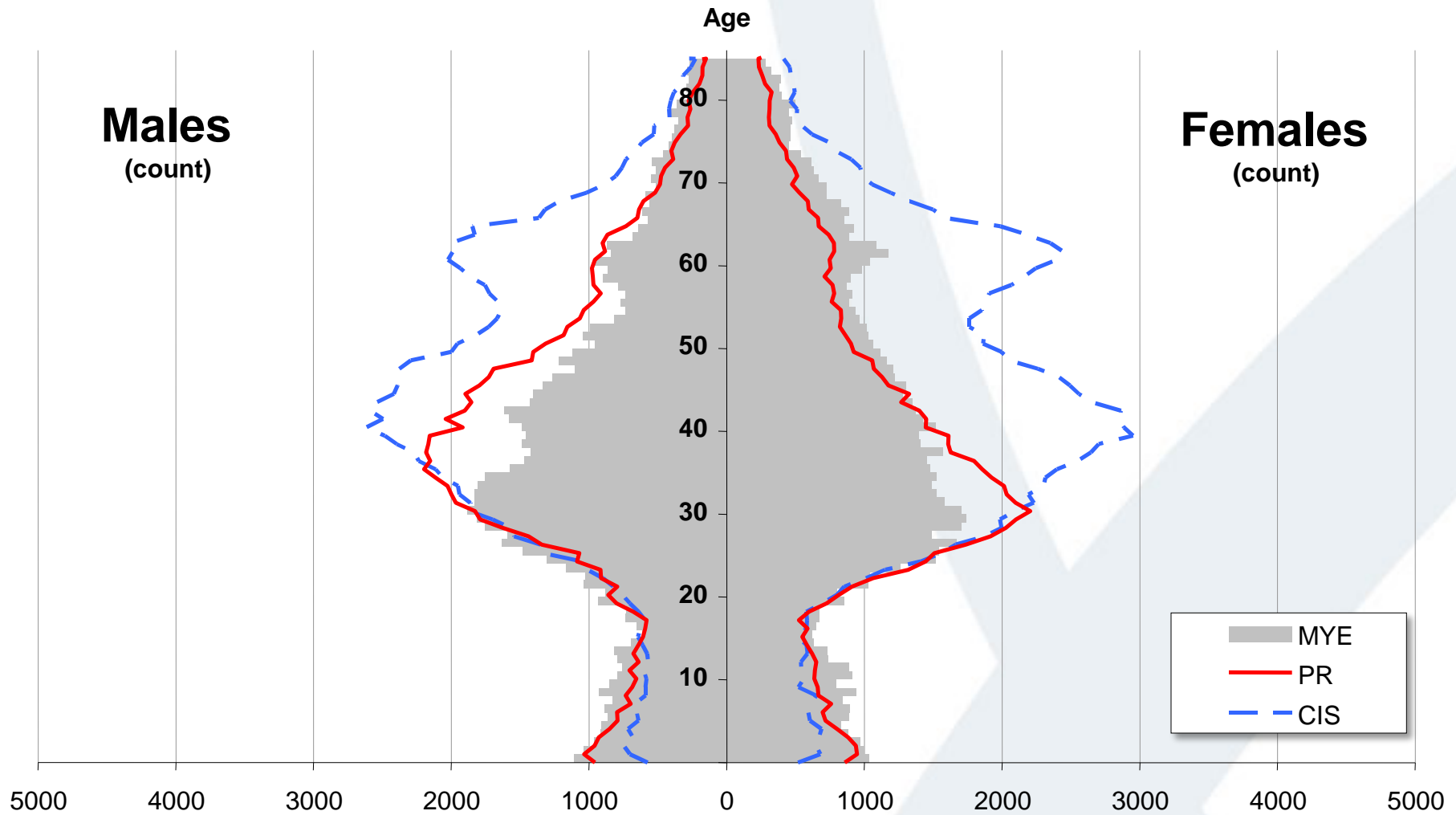
- mid year population estimates (survey-based)
follow international residency definitions
- GP patient register
list those who are registered with a doctor and
include students at their place of study
- Customer Information System (social security)
includes everyone who has ever interacted with the
benefits or tax system even if they are no longer
resident

Some areas show a good correlation...

England



...some areas are not so good



What next?

- looking at how administrative data can contribute to the production of demographic and social statistics.
- linking data from a range of sources.
 - No common identifier
 - Using name, date of birth and address information to establish the link
 - but even these few variables are not held consistently...

Matching

- Date of birth
 - DD/MM/YYYY
 - YYYY/MM/DD
- Names
 - middle names
 - double-barrelled names

Short term enterprise statistics (STS)

Two main problems:

1. Administrative data are due later (*timeliness*)
2. The reference period for the admin data is different - for example, quarterly VAT returns but monthly STS delivery (*periodicity*)

Need to estimate for the missing data, to determine growth rates of turnover (VAT data) and employment (social security data).

Short term turnover statistics

- The UK has a unique complication:
- VAT returns cover a 3 month period
- But not always calendar quarters – enterprise can choose:

	April	May	June	July	August	September
ONS Requirement	[Light blue bar]			[Light blue bar]		
Enterprise A		[Light blue bar]			[Light blue bar]	
Enterprise B			[Light blue bar]			[Light blue bar]

Next Steps

- Monthly survey calibrated by VAT quarterly data;
- Time-series analyses for separate estimation for small enterprises;
- Model-based: by comparing the growth rate for large enterprises with that for the entire population.

Any Questions?

