

Integrated data for Population and Household statistics – ONS's plans for whole system change.

Becky Tinsley & Ian O' Sullivan

Integrated Data Division & Social Survey Transformation Division
Office for National Statistics

What are we doing?

ONS are exploring the potential of using linked multiple administrative data sources to produce outputs on the population that are traditionally based on a full population Census.

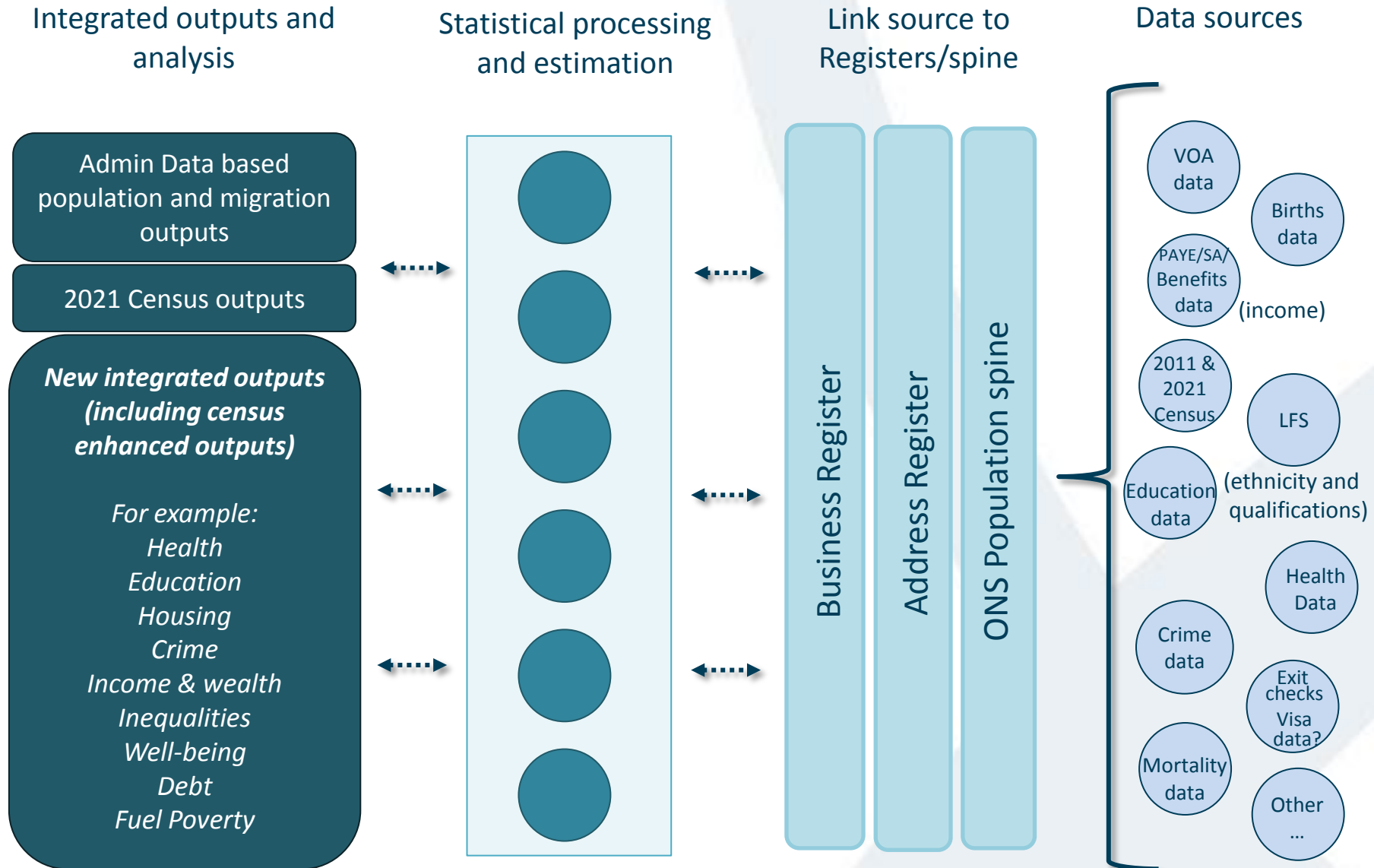
And we are also...

Taking this opportunity to transform how data traditionally collected via large scale household surveys are sourced and think about the future framework for household surveys.

Why are we pursuing integrated data?

- Decision-makers increasingly need better local data on the size and characteristics of their areas to understand the changing dynamics of the population
- We have a rapidly changing population and need up-to-date data to make decisions
- Census = key in meeting needs in 2021, granular, local data but quickly out of date
- Potential to help us answer questions we haven't been able to answer before eg what international students do after their studies

Integrated Data Framework



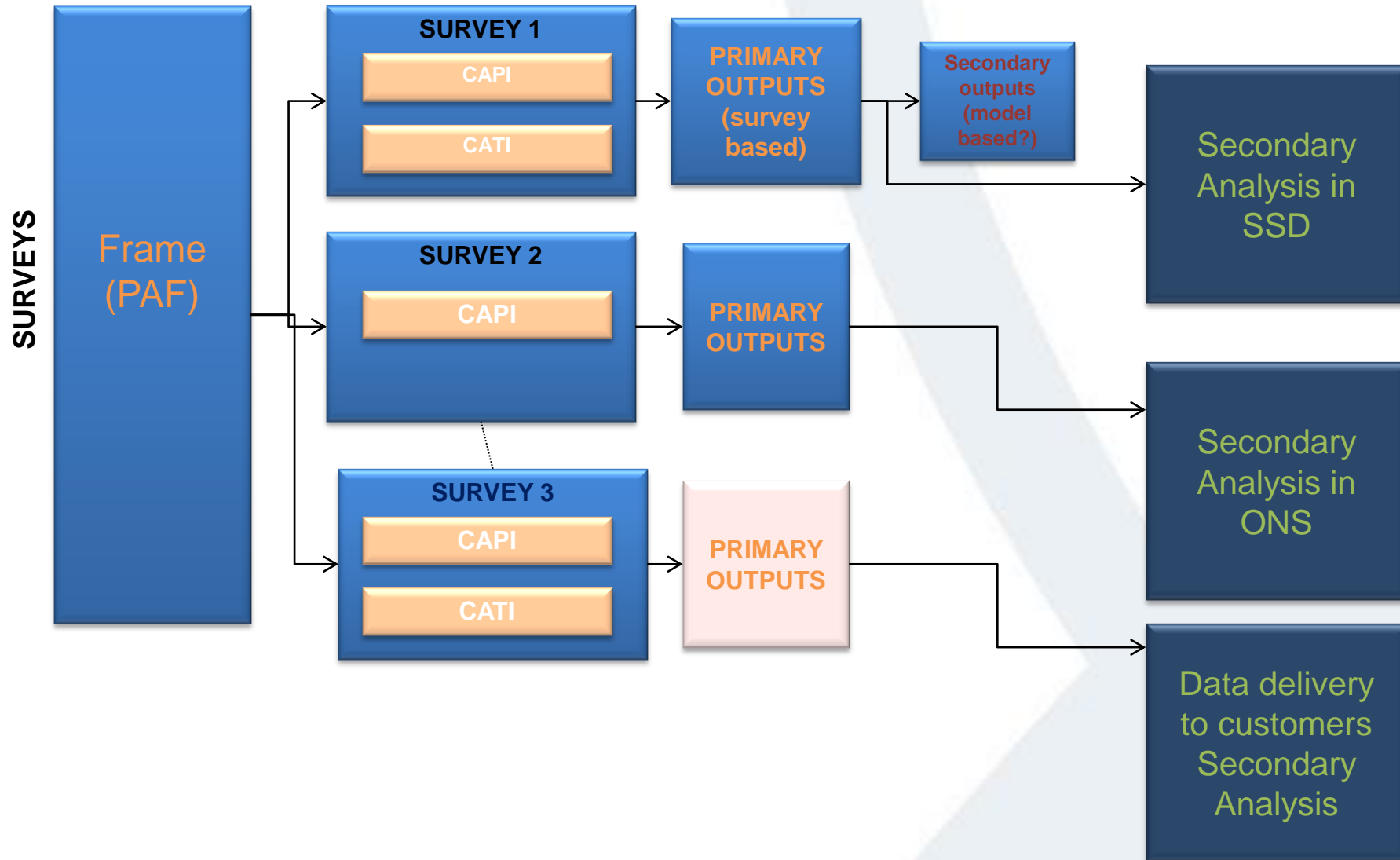
Transforming the population and migration statistics system...

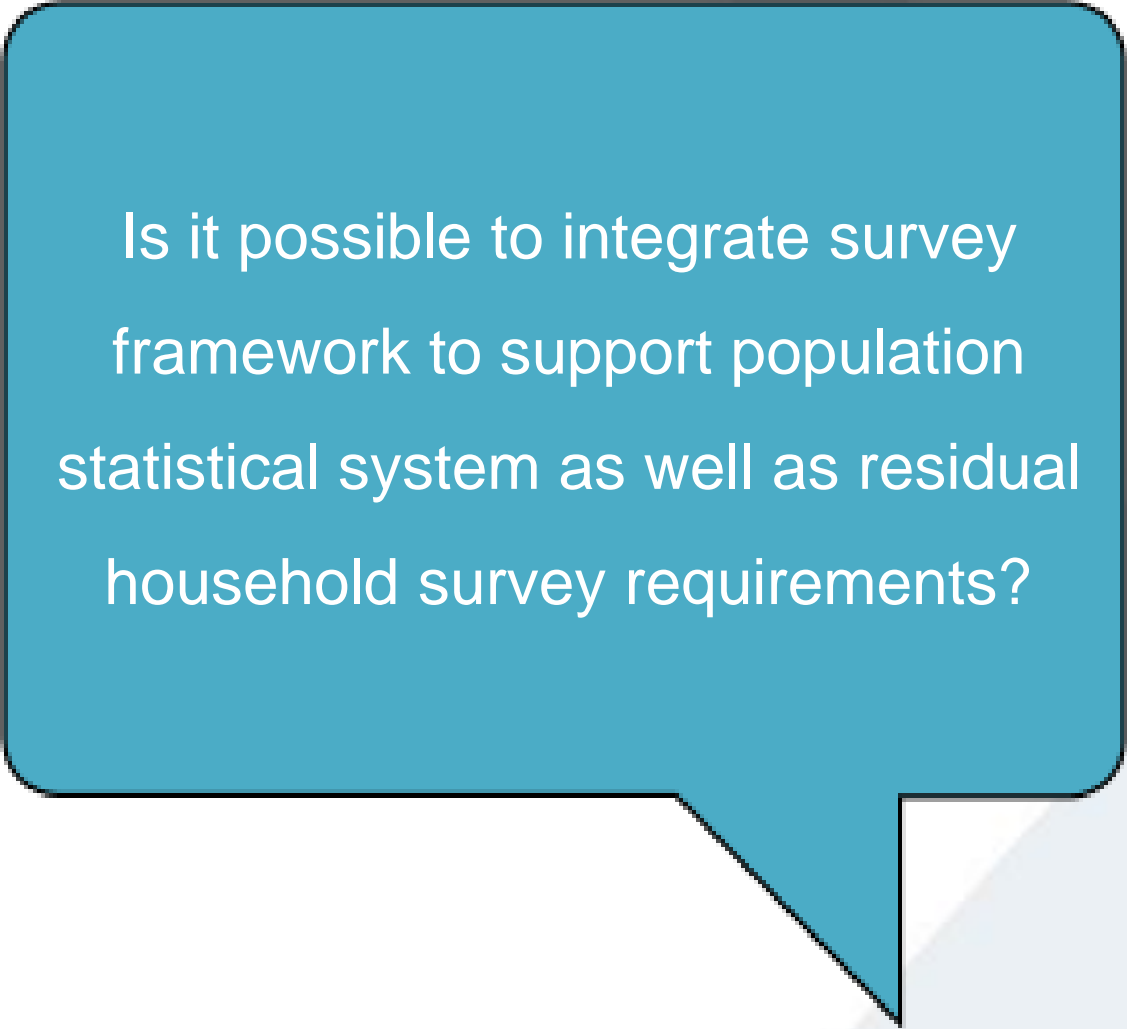
- Strategic - to make better use of all available data to produce better statistics, better decisions
- Current system - challenges to quality of population statistics:
 - 10 year before we can rebase population
 - International Migration statistics (survey based, emigration challenge)
- Un-met user needs - gaps between what users need and what we can currently produce (e.g. characteristics of migrants)
- Build flexibility to incorporate new data and methods.

What place for surveys within Integrated Data System...

- Need access to range of admin data and combine with surveys. **Likely to need new surveys:**
 - Coverage survey(s) to help measure size of population and households
 - Adjust for coverage biases in the SPD
 - Annual characteristics survey – size tbc
 - Collect target variables not available in admin data

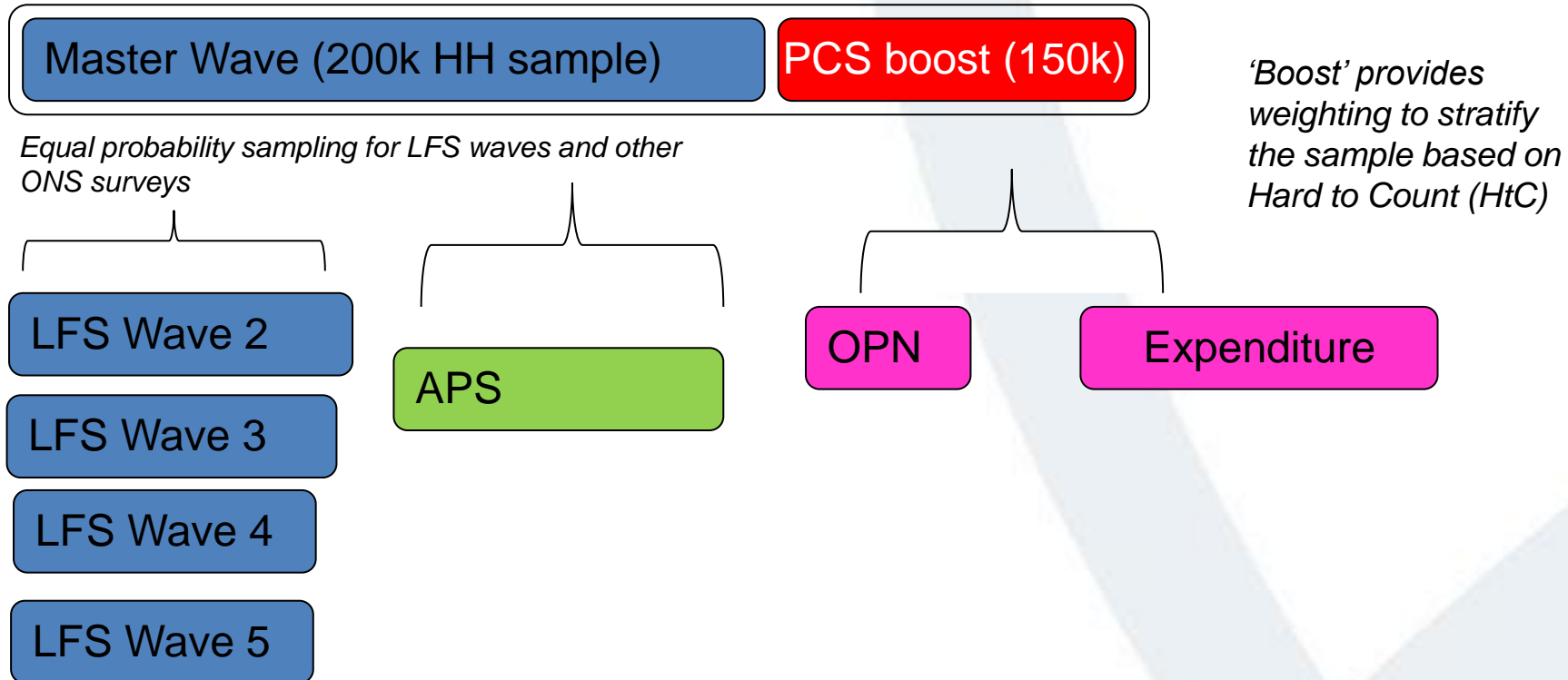
The Present Household Survey Data Framework





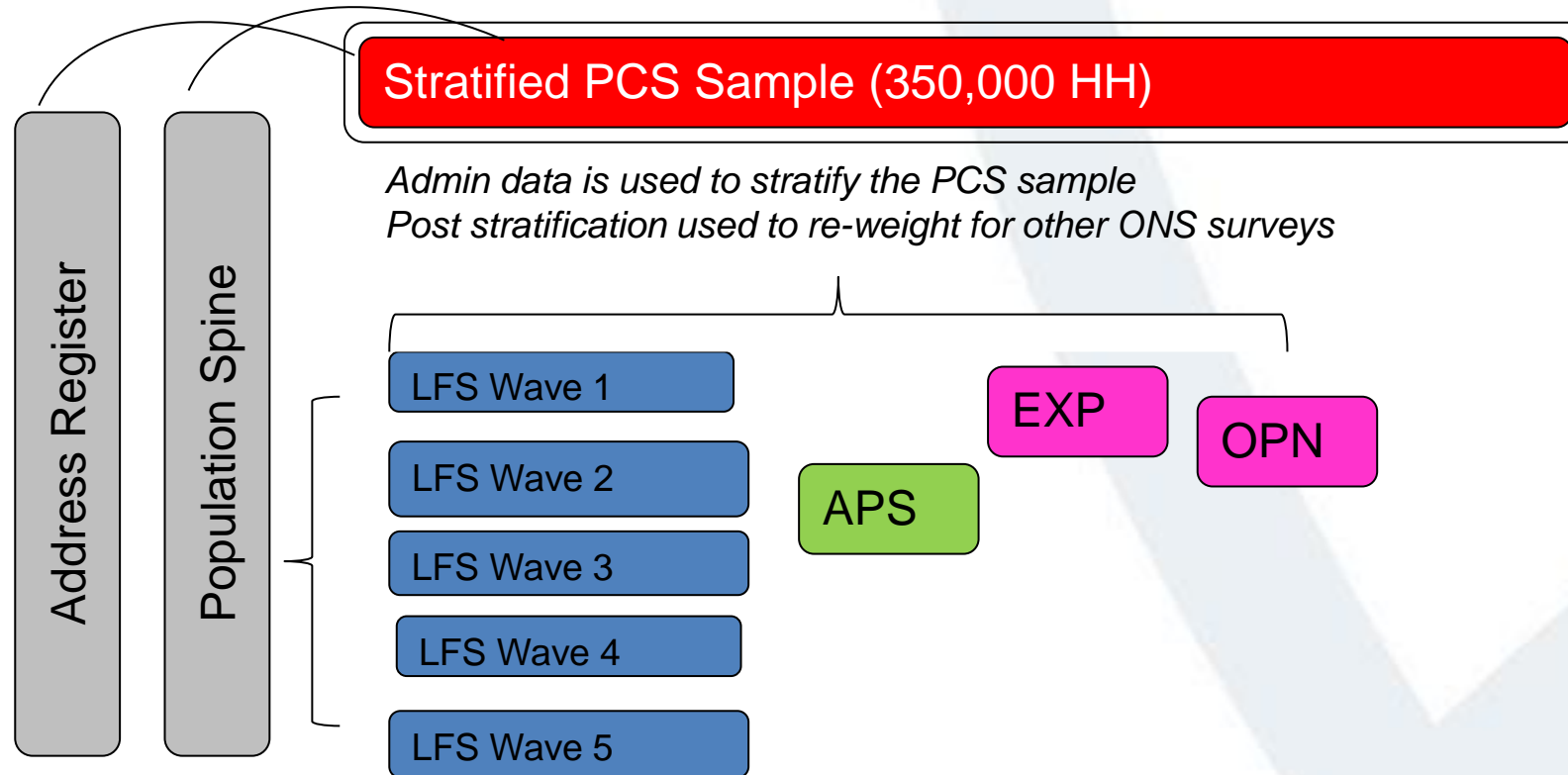
Is it possible to integrate survey framework to support population statistical system as well as residual household survey requirements?

Option 1 – LFS design with PCS ‘boost’



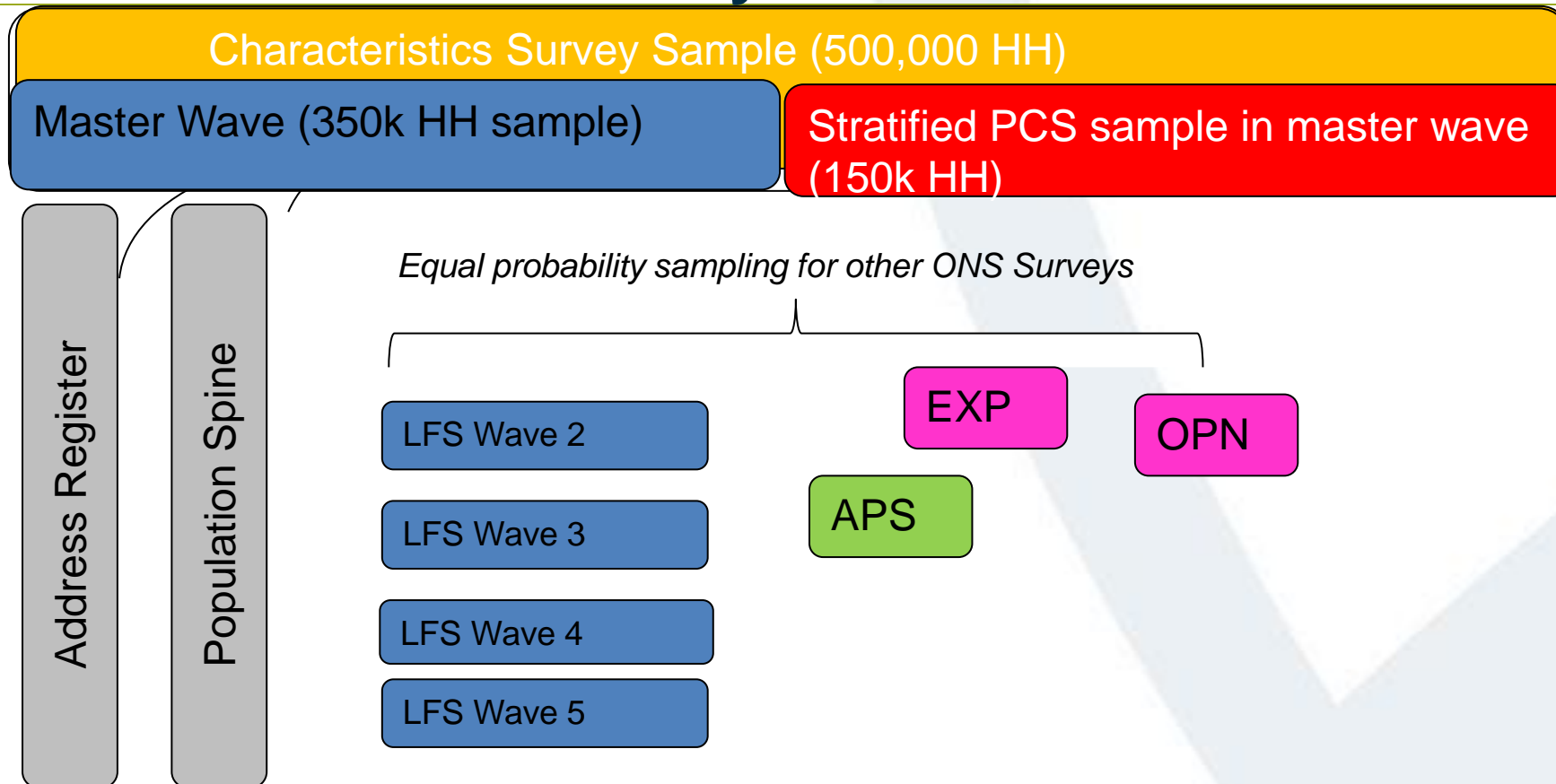
- Master wave randomly sampled - provides a frame for other ONS surveys
- Boost needs to be large enough to meet all PCS requirements (HtC weighting, dependant sampling etc)

Option 2 – Master Wave designed for PCS



- Use administrative records to stratify the PCS sample
- Admin data needs to be of high quality to re-weight PCS sample for LFS and other surveys

Option 3 – Master Wave based on ADC Characteristics Survey



- Master wave based on larger random sample of households
- Use admin data to check that PCS requirements met by master wave

Challenges

- Data access – availability and quality of data – is it fit for purpose?
- Admin data alone is unlikely to deliver population statistics of sufficient quality, likely to need surveys to support the transformation.
- Methods – new and innovative approach required, developing methods alongside other NSIs.
- Will it be possible to integrate?

Vision for Household Statistics

To use administrative data to produce and improve full range of household statistics

Make effective and efficient use of administrative data held across government to reduce respondent burden

Residual surveys to support the population and household statistical system

Where surveys are needed they will be 'digital by default'!

Online first – but not online only – mobile first design BUT optimize for mode

Start Wave 1 asking all (or most) to complete online, and follow up non-response with f2f or phone contact

A lot of learning already in 2017 about the best way to maximise take up online, and best incentive to use.

Transformation of the LFS

Now: Transformation, not translation!

Radical redesign: strip back to the output requirements
and design the online mode first!

What?

- Changes to survey content (cut down, rationalise)
- Changes to questionnaire flow (more relevant)
- Changes to questions (simplify)

How?

- Design principles
 - Question design mobile focussed
 - Questionnaire design respondent focussed
- Agile and iterative development
- Focus group with interviewers
- Designing from data
- Non-survey data

Design for mobiles!

Graceful Degradation



Progressive Enhancement



Image credit: [Beacon](#)

GOV.UK – GDS

design principles



5

ITERATE. THEN ITERATE AGAIN



Design
principles
Government Digital Service

1 Start with needs (user needs, not government needs) 2 Do less 3 Design with data 4 Do the hard work to make it simple 5 Iterate. Then iterate again 6 This is for everyone 7 Understand context 8 Build digital services, not websites 9 Be consistent, not uniform 10 Make things open 11 Make things better
www.gds.uk/designprinciples

+

G.O.O.B

(Get Out Of the Building)

Proto-type

Usability testing
&
Cognitive testing

Focus groups

Pop-up testing



Were you working / employed in ref week?

Original

- Did you do **any paid work** in the week ending Sunday the [date], either as an employee or self-employed?
- ‘Paper work?’ ; “No, just my normal job”

v.1

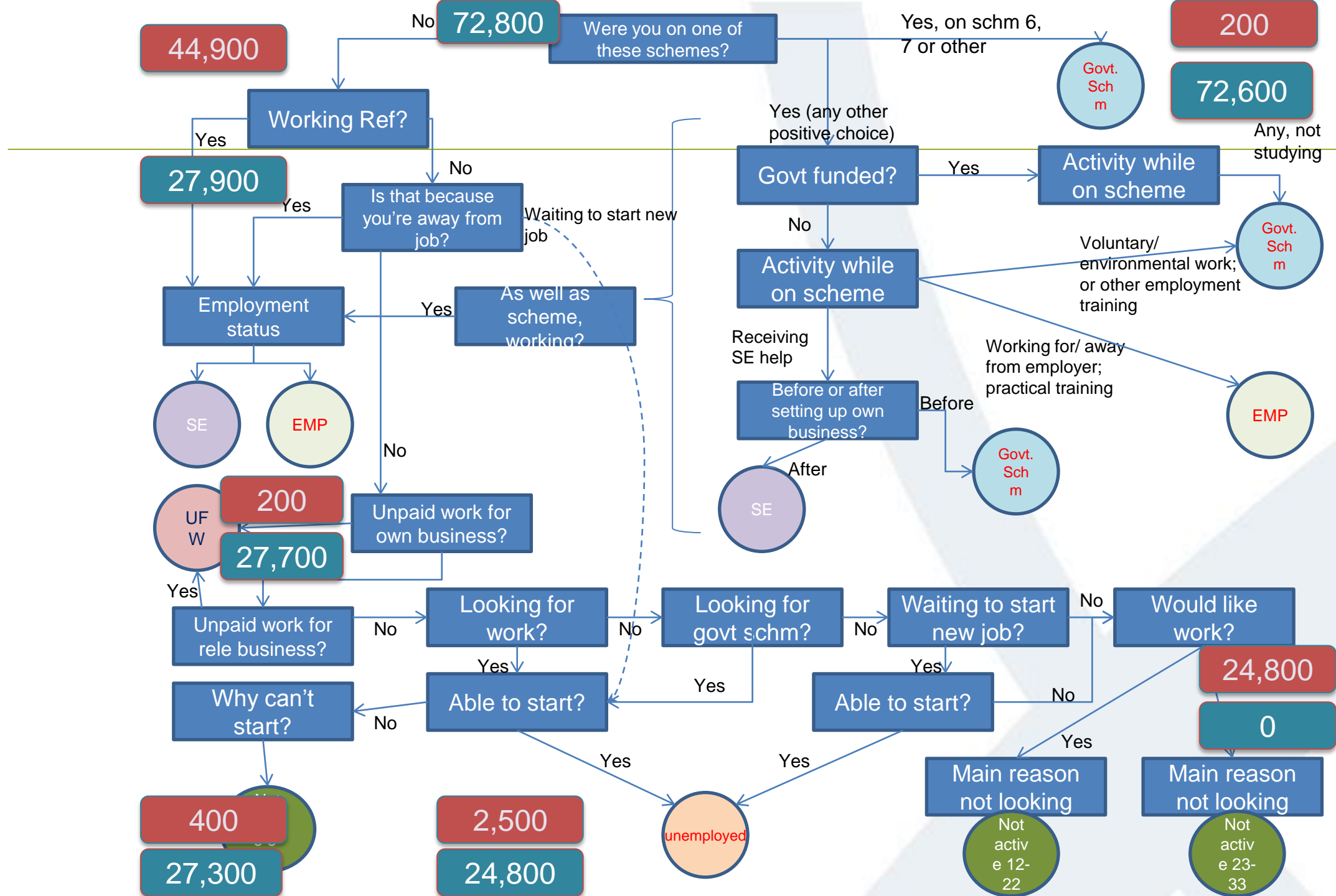
- Did you do any work **for payment or profit** (including in self-employment) in the week ending Sunday [date], even for as little as one hour?
- National account definition;
- “I would have said yes because they still gets holiday pay”
- “Well I was on holiday so I was paid”

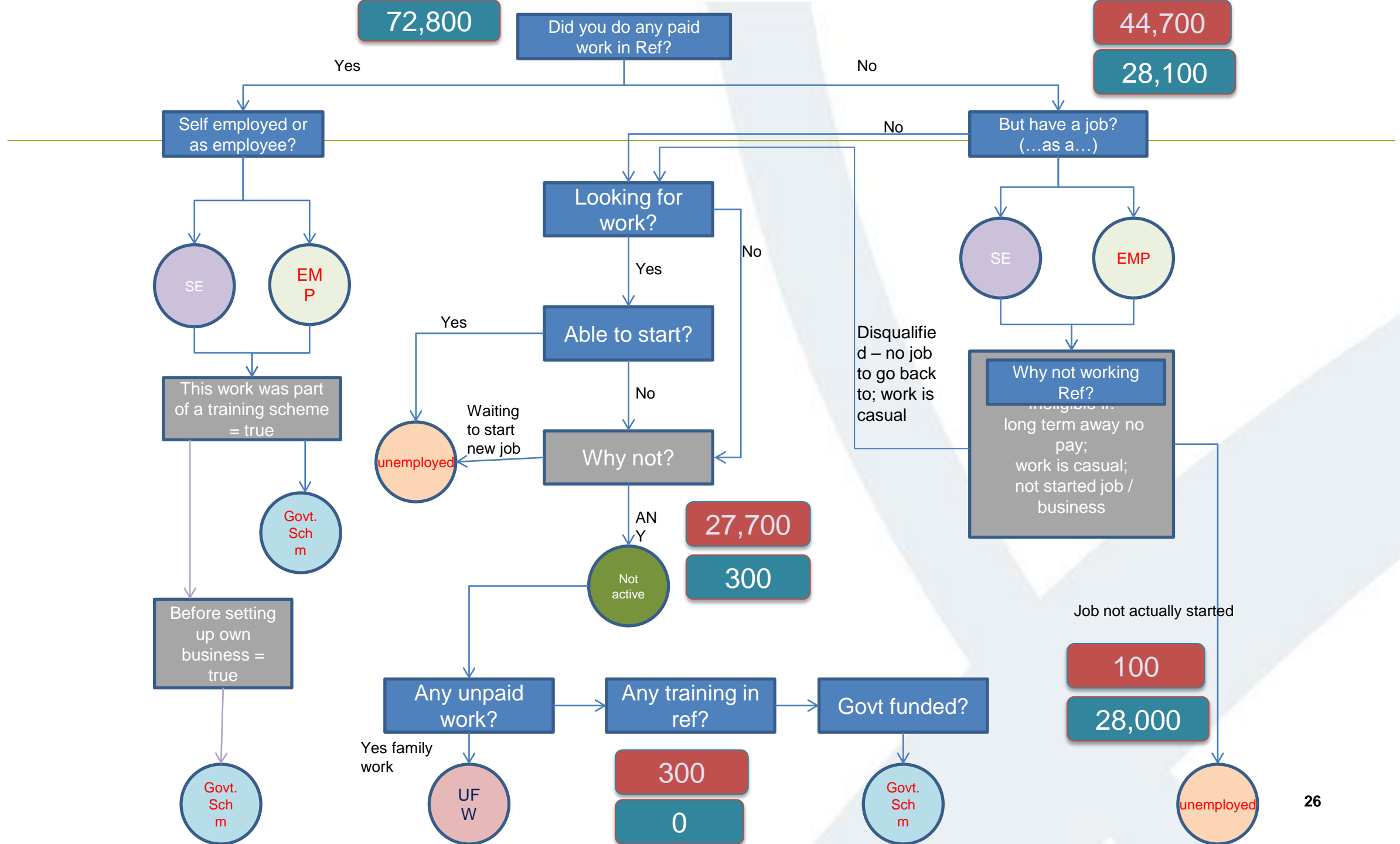
v.2

- Did you have a paid job or **business** in the week Monday [date] to Sunday...?
- Job easy concept; do not have to be present in Ref, or paid in Ref;
- “I was paying myself a wage but nobody was paying the company... the question is are you talking about it as a business or not?” – Sole traders

v.3

- Did you have a paid job, **either as an employee or as self-employed**, in the week Monday [date] to Sunday [date], [year]?
- Better captures different self-employed types;
- “but if it had been ‘self-employed’ that would have been a yes”





Key aims and Research Areas

1. Online take-up?

2. Mixed-mode response?

3. Online retention?

4. Mode switching?

5. Bias?

Research to date...on-line uptake Test1

Purpose

- Early indication of uptake for an online Labour Market Survey
- Provide evidence of the most effective comms strategy

Basic design

- Mail out to c.37,000 households across Great Britain
- Tested:
 - **Advanced mailing strategies**
 - **Envelope colour**
 - **Day of week the invite letter is received**
 - **National branding on envelopes**

Results – Test 1

19.9%
of households entered access code + started survey

Communication Strategy

- Invite + Reminder - 17.9%
- Prenote + Invite + reminder- 20.2%
- Invite + 2 reminders- 21.4%

Envelope colour

- Brown – 20.2%
- White – 19.5%

Mailing Day

- Wednesday – 20.7%
- Friday – 19.1%

Nation Specific Branding

- Scotland
 - With – 18.5%
 - Without – 18.9%
- Wales
 - With – 19.8%
 - Without -19.1%

On-line uptake Test 2

Purpose

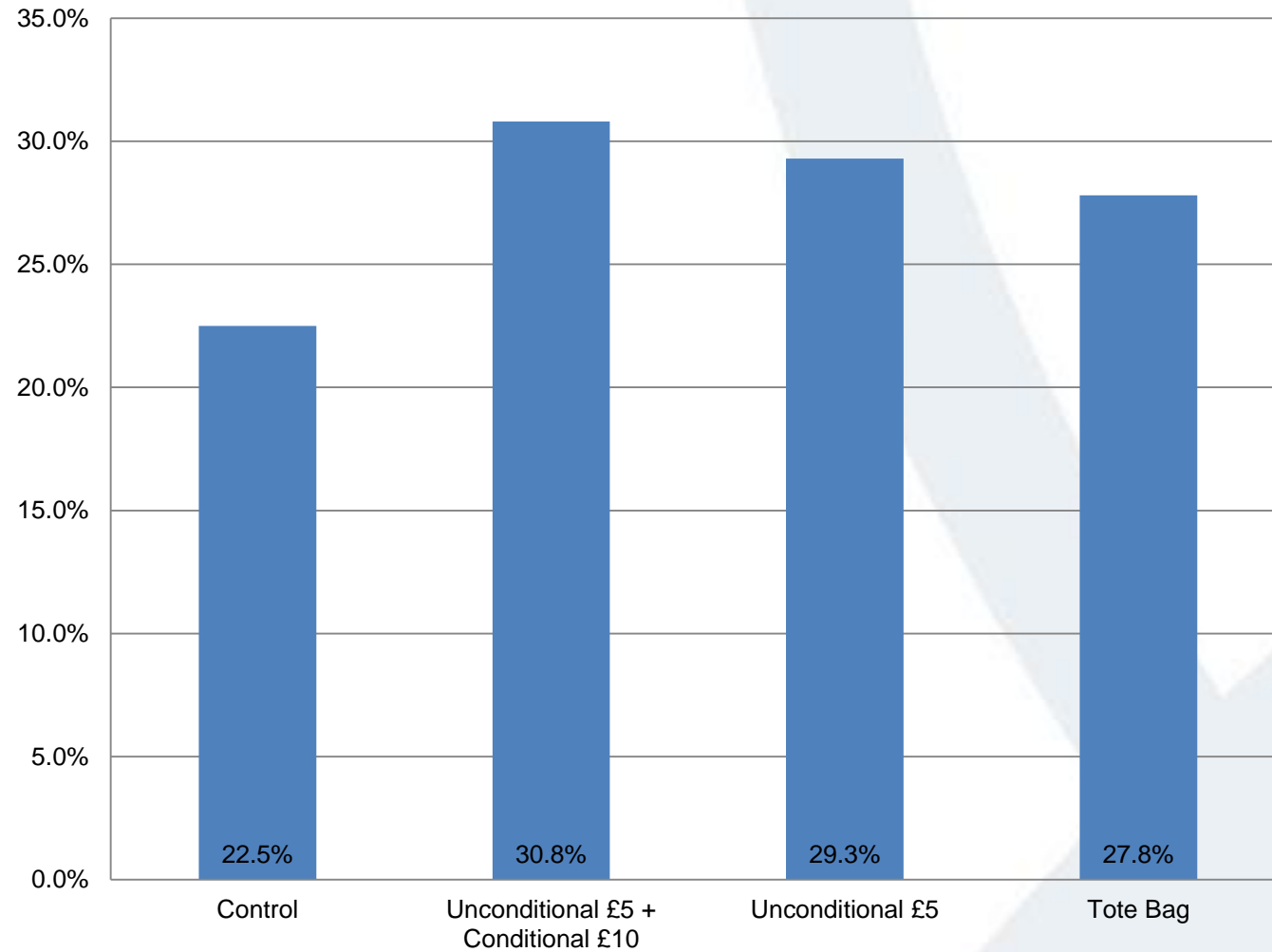
- Establish the likely most cost effective incentive strategy

Basic design

- Mail out to c.40,000 households
- Test impact of:
 - **No incentive**
 - **£5 unconditional + £10 conditional**
 - **£5 unconditional**
 - **Non-monetary incentive (tote bag)**



Results – Test 2



Household response rates

Response as a proportion of all accessing the survey	Test 1	Test 2
	(%)	(%)
Complete whole households	86.4	85.2
Partially complete households	6.2	7.0
Complete and partially complete households	92.6	92.2
Partial response	3.2	4.9
Accessed but did not complete household grid	3.8	2.9
Accessed but did not answer any questions	0.3	0.2
All accessing survey	100.0	100.0

Many thanks

Questions?