Maximising Response Strategy Overview

Orlaith Fraser – May 2019

# Introduction

This paper provides an overview of the maximising response strategy, focussing on each of the following maximising response work epics and how they interact to enable us to meet our quality targets and inform the statistical design for the collection operation of the 2021 Census:

* + Hard-to-Count index (digital and willingness domains)
	+ Self-response profiles
	+ Field Operation Simulation model
	+ Response Chasing Algorithm
	+ Field Prioritisation Algorithm
	+ Targeted Action Groups

Each of these epics are dealt with in detail separately elsewhere (with the exception of the targeted action groups), and so the purpose of this paper is not to go into detail about any of these individually but to demonstrate how these work together to form a coherent strategy and to demonstrate where each epic fits within the bigger picture.

# 2021 Census Statistical Quality Targets

Targets have been set to meet or exceed response rates achieved in the 2011 Census. We aim to maximise response overall, with a target of at least 94% of households responding to the 2021 Census. We aim to achieve at least an 80% response rate across all local authorities. We need to minimise variability in response rates within local authorities and within hard-to-count categories (no specific target set yet). Evidence from international census-taking and large-scale survey operations, such as the Labour Force Survey, indicate that it is likely to be harder to get people to respond to the 2021 Census than to previous censuses. It is thus essential that we have a coherent strategy in place to ensure that we reach our targets.

# Maximising Response in the 2011 Census

Abbott and Compton (2014) provide a summary of the design and operational activities undertaken as part of the 2011 Census. The overarching strategy was to maximise response, but also minimise the variability of response across areas and sub-populations. The minimisation of variability was critical for producing population estimates with both high and uniform quality, a key objective for the 2011 Census following lessons from the 2001 Census. This meant greater focus on areas and populations that were likely to have lower response rates so that their response levels could be increased through special operations or increased resources – the consequence being that areas or populations with a likely high response would have lower resource levels. The key enabler for this was a statistical model for non-response, which enabled local area estimates of likely response rates together with a national hard-to-count index which indicated the level of difficulty that might be experienced by the census in an area (covering all areas – both easy and hard). These drove the allocation of resources across England and Wales for the field operation, publicity and the Census Coverage Survey (CCS).

The operational activities were, in summary:

* The questionnaire was developed to ensure that it was easy to complete and maximised response from some of the key populations that might be missed
* The field strategy was designed to enable effective targeting to assist and persuade non-responders to complete the questionnaire. This was through initial allocation, additional recruitment, authorising extra hours and movement of field staff to areas with low response (although the movement was limited to fairly close areas). In addition, ONS tried to recruit collectors with skills relevant to the types of community they would be visiting, especially language skills.
* Special collectors were introduced to oversee the enumeration of specific Communal Establishments, many of which contained hard-to-count populations such as students, prisoners or the elderly.
* Hard-to-count populations were identified and prioritised so that specific operational activities could be put in place for each designated group as part of the census design.
* Communication activities, including community engagement, were designed to invite everyone to participate in the census. This used a broad mix of communication channels to reach both the general population and hard-to-count groups.

# What is different this time?

The 2021 Census will, unlike the 2011 Census, undertake the collection using an online questionnaire as the primary response mode. This change in the basic collection methodology means that the non-response patterns observed in the 2011 Census may no longer be valid for the 2021 Census. For instance, the public will now have to have access to and be able to use the internet in order to respond using the primary mode. This additional barrier to completion means that the digitally excluded population (and those who do not wish to take part online) will be harder-to-count than in the 2011 Census. On the other hand, an online census may make it easier and more accessible for young people which may mean they are more likely to respond.

The landscape across government is different, especially for digital services. As the 2021 Census will be primarily online, the service it offers needs to be developed in the context of policies and good practices developed by Government Digital Service. As is the case in every decennial census, any strategy for hard-to-count must be reviewed in the context of ongoing population and social change. Factors such as ageing, mobility, public opinion etc will have an impact on the extent and makeup of hard-to-count groups.

# How will we maximise response and minimise variability in 2021?

Despite the differences to the 2011 Census, the vast majority of the operational activities applied in 2011 to maximise response are still valid for the 2021 Census, including all examples listed above, and form part of the strategy for 2021. However, the much improved access to relevant data sources, including better data from the previous census, administrative and survey data has enabled us to build more sophisticated tools to be able to more accurately predict patterns of response and capitalise on the rapid influx of live data during the collection operation to respond intelligently and make the best use of available resources.

# How will we ensure that those who are unable or unwilling to respond online still complete a census questionnaire?

We are very much aware that not everyone living in England and Wales will have the digital skills or capability to respond to the 2021 Census online. Whilst our position is to encourage online completion as the primary mode of response, we will also be offering paper questionnaires to those who would prefer to complete their forms on paper.

Paper questionnaires will be available via a number of routes to ensure that, although we will encourage anyone who is able to respond online to do so, there are no barriers to completion for those who cannot, or choose not to, complete their forms online.

There are four different options for receiving paper questionnaires:

## Paper as first contact

While the primary mode of response for the 2021 Census will be online, where we believe that people will struggle to be able to complete the online questionnaire, the most effective and efficient action is to send paper questionnaires as an initial contact. This will reduce the burden on respondents, be more successful in achieving a response, and be the most cost-effective way of providing paper questionnaires. These paper questionnaires will also include a unique access code to enable those who prefer to complete the census online to do so.

Paper questionnaires will be delivered to households in areas where the respondents are most likely to self-respond but where the level of broadband uptake is low, indicating that people in this area may be happy to respond, but not have the knowledge or ability to respond online. The Hard-to-Count willingness and digital domains will be used to determine where these areas are. Additionally, addresses recorded as sheltered housing will receive paper questionnaires as initial contact, as may a small number of areas with particular population groups where engagement with these communities suggests that they might be unlikely to respond online.

 It is expected that about ten percent of all households will receive paper questionnaires as initial contact.

## Paper questionnaire as a reminder

For some areas, reminder letters will include paper questionnaires. Non-responding households in these areas will receive a paper questionnaire in the post, due to land on the third Friday after Census Day. These areas will be LSOAs with poor digital uptake and a low willingness to respond, as these households will not have received paper as an initial contact but are more likely to have a need or desire to respond on paper. Paper questionnaires may also have a greater effect in these areas where the households are less willing to respond than reminder letters, simply because a paper questionnaire is less easily missed than a letter.

## Via contact centre or online

Anyone who wants to complete a paper questionnaire will also be able to request one at any time either via the website, or by calling our helpline, and information about how to do this will be included in the initial contact letter inviting people to respond.

## Handed out by field officers

Census field officers will visit all addresses where no response has been received to remind people to respond. At this point, field officers will also be able to supply paper questionnaires on the doorstep if requested.

## Other forms of assistance

For those who do wish to complete the census questionnaire online, but may require help to do so, face-to-face assistance will be offered through an assisted digital service at a variety of locations. ONS will also be providing help with completion at organised completion events and help will also be available via our contact centre.

# Targeting specific hard-to-count groups

Some population groups require additional specific support or targeted action to ensure that they are aware, willing and able to respond to the census. The **Targeted Action Groups** (TAGs) consist of a list of groups of people, defined by some characteristic, who are thought to be harder to count in a census than the general population. Groups are included in the list because not counting them will have an impact, either national, local or reputational, on the census and specific action is required in order to ensure that they respond. This can range from ensuring that specific information is available to respondents on the online guidance, to providing specific material (such as translation booklets or questionnaires in Braille or large print) to having a dedicated community advisor working in the field to engage with that particular community for up to a year before Census day. The TAGs form an important part of our strategy to ensure that everyone has the opportunity to respond to the census, and to ensure that we maximise response from all population groups, ensuring high quality census estimates can be produced at low geographic area level. See the appendix for a list of targeted action groups and the methodology for selecting those groups. The groups included on this list are reviewed regularly and may be updated at any time if further information becomes available to indicate that changes should be made.

# Predicting patterns of census non-response

Reminder letters, paper questionnaires and field visits will be used to follow-up non-responding households. The communications and media strategy will also be tailored to the point in time of the collection operation and may be targeted to particular areas and/or population groups where non-response is high. In order to maximise response, we need to be able to predict likely patterns of census response to ensure the most efficient and effective allocation of resources during the planning stages and to identify where areas and or population types are not performing as expected during live operations to target interventions where they are needed most and provide the best chances of meeting our quality targets and ultimately producing high quality census estimates.

The **Hard-to-Count index** identifies areas at risk of census non-response, and comprises two domains, the willingness to self-respond and the digital domain. The willingness domain is constructed using an area level (Lower Super Output) model that predicts non-response by day 10 after census day. The covariates used to build the model parameters are from previous census and updateable administrative data sources. The digital domain is built as an indicator to measure willingness and ability to respond online to a census. Both domains rank (separately) all Lower Super Output areas (LSOAs) and classify them on a scale of 1-5, with hard-to-count group 1 representing the top 40%, group 2 representing the next 40%, group 3 representing the next 10%, group 4 representing the next 8% and group 5 representing the bottom 2% of all LSOAs.

The willingness and digital domains of the hard-to-count index are used to identify areas where we need to send paper questionnaires as initial contact and to feed into the shape of the self-response profiles and the volumes of field follow-up visits.

The **self-response profiles** are a best estimate of the volume of responses we expect to receive during the census period. They will model when, where and how people are likely to respond to the census and are used to inform operational decisions as well as feeding into the field operations simulation model to determine the number of field visits and reminder letters required.

# Allocation of resources for follow-up activities

The **field operations simulation** (FOS) model simulates the 2021 Census collection operation in terms of the resources required over time and geography (at Lower Super Output Area level). The FOS model uses the predicted patterns of response provided by the self-response profiles and the hard-to-count index (as well as other data sources) to produce estimates for the volumes of field visits required, the expected volumes of reminder letters and expected response profiles, amongst other things. This model provides significantly more detail on the basis of much more data than the basic volumetric model used in 2011.

The **Response Chasing Algorithm** (RCA) assesses live response (and return) rates against expected patterns and identifies and prioritises any shortfalls in responses during the census collection operational period. The RCA then recommends appropriate interventions to address those short-falls. The RCA forms an important part of our strategy to ensure that we are not overly reliant on our predicted patterns of response. This tool makes the best use of the intelligence we receive during live operations to target our resources towards those areas that need it most, to correct for any errors in our assumptions and still enable us to meet our quality targets. This represents a significant improvement from 2011, when fewer, less timely, data were available during live operations and there was no systematic evidence-based decision-making process.

Figure 1 shows how the Hard-to-count index, response profiles, field operations simulation and response chasing algorithm interact. For a particular LSOA, the hard-to-count index indicates the likely willingness to self-respond. This then informs the self-response profiles, which along with the HtC categorisation, feeds into the FOS to provide the number of field staff hours required for that area and produce the expected overall response profile for that area. This is then used by the RCA to assess performance during live operations and recommend interventions where there is a risk of not achieving final response rate targets.

**Figure 1.** Interaction between the hard-to-count index, response profiles, field operations simulation and response chasing algorithm



While the RCA focusses on the maximising response at LSOA level, targeting interventions at problem areas, it is also important to ensure that we are actively and continuously minimising variability in response rates at lower levels of geography. The **Field Prioritisation Algorithm** (FPA) does this by analysing return rates at output area level within Census Coordinator Areas to prioritise field visits in areas where return levels are lowest, reallocating census officers from the best performing areas to the worst performing areas to reduce variability in response rates. Whereas the RCA may identify specific problem areas where additional resources are needed (implemented by increasing working hours of existing field officers or moving officers in the area from elsewhere), the FPA only reprioritises existing resources within the Census Coordinator Area. Up to 12 census field officers may work within one Census Coordinator Area and all may be deployed to work in any part of that area. The FPA takes a business-as-usual, automated approach to ensure that visits are always prioritised where response rates are lowest without any additional resources or interventions.

# Design of the Collection Operation

The collection operation has been designed in order to make a number of aspects of the maximising response strategy feasible. In particular, the contract with the recruitment supplier for census field staff includes a requirement for one percent of all staff to be able to be moved to anywhere in the country, at least two percent to be moved more locally (to a neighbouring local authority, for example) and for almost all field staff to be able to move anywhere within their Census Coordinator Area (the exceptions being those who are only able to walk, without access to public transport who may be limited in how far they can travel). Furthermore, the contracts offered include a range of part-time hours, providing the potential for greater flexibility of resources and for increasing field staff hours in a particular area without the need to move officers away from another area.

All non-responding households are expected to receive one reminder letter. Reminder letters will be sent out in three batches, sent one, two or three weeks after Census day. The areas selected to receive each batch are pre-determined according to the likely timing of responses in that area, with those expected to be ‘early responders’ getting the first batch of reminder letters, those expected to respond around Census day getting the second batch and ‘later responders’ getting the final batch of reminder letters. These are the ‘planned’ reminder letters and every household should receive only one of these letters if they haven’t responded by the time the cut of non-responding addresses is taken for the reminder batch allocated to their area.

In addition to these ‘planned’ reminder letters, additional ‘response-driven’ reminder letters may be sent in order to target the areas responding worse than expected, where a reminder letter is considered the most appropriate intervention. Once a week for the four weeks following Census day, we will have a certain number of reminder letters that may be targeted at non-responding households in LSOAs recommended by the RCA to receive these additional ‘response-driven’ reminder letters. Selection of areas to receive response-driven reminders will take into account the planned wave of contact to ensure that no household receives a response-driven and a ‘planned’ reminder letter within the same week, for example.

The statistical design underpins the wave of contact (the timeline of all census touchpoints with the public during live operations), dictating how and when we interact with respondents in order to maximise responses and make the most effective use of resources. The FOS model simulates the events as stipulated by the wave of contact, but also functions as a feedback loop into the statistical and operational design and the wave of contact so that these can be adjusted according to their impact on the response model to ensure that these are all consistent with the quality targets.



# Challenges and limitations

There are specific limitations associated with each of the maximising response epics, but a common theme is that each relies on the accuracy of the underlying assumptions and our ability to accurately predict how people will respond. While we have made the best use of the data we currently have available, these are often far from perfect and we will not know the accuracy of our models until during, or even after the 2021 Census. Although the 2019 Census Rehearsal will play an important role in validating and informing our assumptions, a particular challenge is in identifying ‘worst-case’ scenarios and the difference between these and the most likely or expected scenarios for operational planning. Our best mitigation for this is the use of the response chasing algorithm to redirect resources where needs differ from our anticipated requirements, but even the response chasing algorithm uses predicted response patterns to identify problem areas and is limited by operational constraints in how many adjustments we can make to planned volumes and distribution of available resources.

The strategy presented here, with the exception of some of the TAG work, considers only individuals living in households and not those living in communal establishments. Further work is needed to establish how some of the response modelling work conducted for household respondents and the use of interventions during live operations can be used to ensure that we also maximise responses from communal establishment residents.

# References

Abbott, O. and Compton, G., 2014. Counting and estimating hard-to-survey populations in the 2011 Census. In: R. Tourangeau, B. Edwards, T. Johnson, K. Wolter and N. Bates, eds. 2014. *Hard-to-survey populations*. Cambridge: Cambridge University Press. Ch.4.

# Linked papers:

Dini, E. Hard-to-count index for the 2021 Census – presented to the External review assurance panel in October 2018.

Meirinhos,V. Self-response profiles for the 2021 Census – to be presented at the External review assurance panel in June 2019

Meirinhos,V. Response Chasing Algorithm – to be presented at the External review assurance panel in June 2019

Meirinhos,V. Field Prioritisation Algorithm – to be presented at the External review assurance panel in June 2019

Ward, K., Priestley, M., Fraser, O. Simulating Census Operations to Inform Resourcing Decisions – to be presented at the External review assurance panel in June 2019

Appendix - Targeted Action Groups

Selection Methodology
**Reasons for including groups includes:**
• Anecdotal evidence that this group is hard-to-count
• High person imputation rates (we expect a 94% response rate so a person imputation rate of more than 6% indicates a problem with response)
• A non-response bias in the item imputation rates (this indicates that proportionally fewer people have responded that have a particular characteristic than we believe exists in the population)
• Specific action was taken last time to encourage a group to respond, without which there is reason to believe the group may not have responded as well as they did.

A red/amber/green (RAG) status is assigned to each group. The RAG status is based upon the size of the group and the impact that not counting them will have. So, a large group which will have a big impact if not counted will be awarded a red status. A small group which might have a big impact if not counted may be awarded a lower status. If actions are put into place that mean that we think we will be able to count the group then the RAG status may be lowered. Note that lower imputation rates in 2011 compared to 2001 may be as a result of such mitigating actions in 2011 and so should not be a reason on their own to lower the RAG status. The following figure shows how the likelihood of the group to respond, and the impact of not counting it, will contribute towards the groups RAG status.



## List of Targeted Action Groups

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Targeted Action Groups | RAG Status  | Targeted Action Groups | RAG Status  | Targeted Action Groups | RAG Status  |
| **HEI Students not in halls aged 16-24** | Red | **Cognitive health issues** | Amber | **One-person households** | Green |
| **Students in halls of residence** | Red | **Households with no internet access** | Amber | **People with more than one residence** | Green |
| **Gypsies/Travellers/Roma** | Red | **Pre-schoolers** | Amber | **Boarders/lodgers** | Green |
| **Bangladeshi Ethnic Group** | Red | **Adults 80+** | Amber | **School boarders** | Green |
| **Black African Ethnic Group** | Red | **Low income households** | Amber | **Polish/Eastern European migrants** | Green |
| **Black Caribbean Ethnic Group** | Red | **Beds in sheds** | Amber | **People in Asylum process** | Green |
| **Somali Ethnic Group** |   | **Deaf/Hearing Difficulties** | **Amber** | **Gated communities** | Green |
| **Chinese Ethnic Group** | Red | **Indian Ethnic Group** | **Amber** | **Emerging groups who don't use online gov services / IT security phobia**  | Green |
| **Young Adults**  | Red | **Pakistani Ethnic Group** | **Amber** | **LGBT+ Community** | Green |
| **Short-term Migrants (3-12 months)** | Red | **Arab ethnic group** | **Amber** | **Nepali Community**  | Green |
| **People living in multiple occupancy dwellings**  | Red | **Sikhs** | **Amber** | **Household visitors** | Green |
| **Illegal Immigrants** | Red | **Orthodox Jews** |  Amber | **Armed forces** | Green |
| **Homeless (not rough sleepers)** |   Red | **Cornish speakers** | Amber | **People living in prisons** | Green |
| **Rough sleepers** | Red |  |  | **People lacking digital skills** | Green |
|  |  |  |  | **People living in private rented accommodation** | Green |

Targeted action groups and RAG status as of June 2019. These groups are reviewed and updated on a regular basis as we obtain more data and our plans progress to mitigate for any likely challenges f

/ acing these groups.