Detailed methods for assessing the benefits of Census 2021, England and Wales

1.0 Specific questions/discussion points

We would appreciate the panel's consideration of the following points:

- Is the proposed method for the assessment of the benefits of Census 2021 appropriate? Does it align with best practice; are there changes we should consider making?
- ONS are developing new methods using a range of data sources with the ambition to produce more regular and timely population and migration statistics at both national and local levels. As a result, although the data that stakeholders use will be reliant on census data, the census data may increasingly be supplemented with other data sources. How might this best be reflected in the benefits calculations, particularly as this will be a changing picture through the assessment period and we cannot say at this point exactly how it will progress?
- There are many smaller scale uses of census data that would be impractical to directly quantify, but their aggregated value may be considerable. In an assessment of the value of census data in Australia, researchers assumed (based on the 80-20 rule) that the value of this 'long tail' was 25% of the total value of those uses that were quantified. Does the panel consider that this would be an appropriate way to monetise the benefits that cannot easily be directly quantified? Are there alternative approaches we might consider?

2.0 Main Points

The Office for National Statistics (ONS) will be assessing the benefits of Census 2021 for users from central and local government, the private sector and the voluntary and community sector. We previously carried out an exercise to calculate the forecast benefits of Census 2021. We intend to use the same methods to assess the realised benefits, so that the results can be compared.

We will engage with stakeholders who contributed to developing the forecast benefits, as well as a widened stakeholder base, including the voluntary and community sector, to achieve good coverage across census topics. Our engagement methods will include focus groups, meetings, surveys and correspondence. Some stakeholder engagement will be through overarching bodies and membership organisations. We will assess survey and focus group responses to identify information which requires further action and use qualitative methods such as thematic analysis to understand wider themes to contextualise our findings.

We recognise that there may have been significant changes to uses of census data and to benefits since the data which informed the forecast benefits was collected. We will ask stakeholders whether the uses and assumptions made in the

forecast exercise remain relevant and establish reasons if uses have changed. We will identify any new uses from users who engaged in the forecast exercise and from our expanded stakeholder base.

Benefits to local authorities (LAs) identified in the forecast exercise related to the role of census data in informing spending decisions. The value of big data was calculated as a percentage of the total capital and revenue spend on categories relevant to census, by LAs in England and Wales, per annum. Benefits were estimated by asking selected local authorities to consider the how much of the data relied upon could be attributed to census. We expect to replicate this method.

Central government benefits quantified in the forecast included benefits for funding allocation, policy research and informing evidence-based investments. Benefits related to funding allocation were estimated by considering the welfare loss due to sub-optimal allocation of funds that would arise from relying on alternative sources of data if census data were not available. The cost of research was used as a proxy for the value of policy research. For evidence-based investments we assessed the percentage of the spend that could be attributed to census data, or the delay expected to projects, such as major transport schemes, as a result of not having census data and the associated costs for the delay. We intend to use the same methods to calculate the realised benefits.

To forecast the benefits of census data to the private sector, we engaged with one representative from each industry known to use census data, drawn from a membership organisation of commercial users of government demographic datasets. Each representative indicated the value of data to each relevant decision or activity, and the proportion attributable to census. We used this to calculate an initial benefit estimate for the representative organisation and scaled this up for their sector.

We expect new benefits, including those for our widened stakeholder base, to fall into similar categories to existing forecast benefits, so we anticipate using similar methods to assess them.

We will share our initial estimates with contributing stakeholders and overarching bodies to confirm that they are reasonable and generalisable. We will consider whether those benefits that we identify but cannot directly quantify can be accounted for in our overall benefit assessment. We will adjust benefits for optimism bias, sensitivity and decline over time.

3.0 Our approach to collecting data and calculating benefits

In the Census 2021 white paper, we stated that we expected to realise around £5 of benefit in the wider economy for every £1 spent on Census 2021. The total benefits to central government, local government and private sector users were expected to be £5.5bn over the ten-year appraisal period. This was based on evidence from the assessment of benefits from the 2011 Census, reviewed and refreshed between 2017 and 2018 in order to assess the forecast benefits of Census 2021.

We have outlined our intended approach to assessing the realised benefits of Census 2021 data. To develop our approach, we reviewed the assessment of the benefits of the 2011 Census and work to produce a forecast of the benefits of

Census 2021 data. We looked at similar benefit assessment exercises, including by Statistics New Zealand in 2013 and 2021, and the Australian Bureau of Statistics in 2019. We also considered a report presented to the Conference of European Statistics about measuring the value of official statistics.

Our previous publication summarised the methods used to develop forecasts for the expected benefits of Census 2021. We intend to replicate the methods used to create the forecast, so we can validate the forecast benefits and understand if they have been realised. By using the same methods to assess the benefit, we will help to ensure that the results can be compared to the forecast benefits. Our approach also takes into consideration the available resource to complete the data collection and assessment.

We are mindful that the research and data environment today is different to 2018, when the forecast benefits were developed, and the uses and benefits of census data may have changed too. Because we are using a variety of data sources to provide more frequent, relevant and timely statistics, it may be that the degree to which value is attributed solely to census data has changed since forecasts were made. Our conversations with stakeholders will help us to understand these changes and indicate whether adjustments such as how we account for changes over time need to be revised.

The method will be subject to an internal and external review process, and if this indicates that updates are necessary, particularly to align with best practice, we will make changes to the method as required.

4.0 Identifying our stakeholders

We identified a list of census users who we will approach to obtain evidence to inform our assessment of the realised benefits of Census 2021 data: our contributor stakeholders. This includes users who engaged with previous benefits assessments, users who already engage with or are known to the ONS and additional users identified from desk research.

We have prioritised stakeholders according to their importance in validating the forecast, and the degree of quantifiable evidence we expect them to provide. Our final list of users also considers our capacity to engage, whilst maintaining good coverage of census users across all topics.

Benefits Priority 1

- Previous Engagement with Benefits Quantification. We must engage with these stakeholders to validate whether the forecast benefits have been realised.

Benefits Priority 2

- Not part of previous benefits quantification but likely to contribute quantifiable or very valuable qualitative evidence of benefits

Benefits Priority 3

- May contribute some evidence, may be captured as part of other groups

We have prioritised quantifiable benefits, as this evidence will be most useful in calculating and monetising the benefits. We recognise the value of those benefits that cannot be quantified and intend to identify and describe these in our final report. We will also explore whether those benefits that cannot be directly quantified, the 'long tail', can still be accounted for in our benefit calculations.

4.1 Sampling

We selected our priority one LAs based on those who engaged with the forecast exercise, which we understand was selected on the basis of participation in ONS stakeholder groups. Similarly, priority two LAs were chosen to provide wider representation across different LA tiers (such as unitary authorities and district councils) and to provide good geographical coverage, but this selection is still influenced by existing engagement. This is pragmatic, in that LAs with generally strong engagement with ONS may be more receptive to taking part in this exercise. There is a risk that bias would be introduced, in that those LAs who most value ONS data may be more likely to be engaging with ONS. By collecting feedback from all other LAs on how representative the initial estimates are for their organisation, we hope to mitigate against this.

Business stakeholders were purposefully selected because they are part of an industry known to use census data, have already engaged with ONS, or desktop research suggests they are one of the leading organisations in a sector. Voluntary sector stakeholders were similarly purposefully selected.

Bias may be introduced here; by drawing volunteers from a membership organisation of commercial users of government demographic datasets, we are more likely to find users that are experienced and knowledgeable about how census data can be used. Larger organisations within a sector may use census data differently to smaller ones. The data we collect from our stakeholders may not therefore be fully representative of the sector of which they are part. We do intend to survey an expanded stakeholder base and we are continuing to explore how we can address this further when we seek to scale up our initial estimates to wider sectors.

5.0 Collecting evidence from stakeholders

5.1 Our approach

We have identified our expected first and second approaches, by stakeholder group and priority.

Organisations within each priority category are listed in Annexe B.

Stakeholder group	Priority	Expected primary approach	Expected Secondary Approach
Local Government	1	Direct – focus group	Correspondence
	2	Direct – focus group	Correspondence
	3	Survey (likely via LGA for LAs)	Follow up by meeting or correspondence if indicated
Central Government	1	Direct - meeting	Correspondence
	2	Survey	Follow up by meeting if indicated
	3	Survey	Follow up if indicated (meeting/correspondence)
Private Sector	1	Direct - meeting	Correspondence
	2	Survey	Follow up by meeting if indicated
	3	Survey	Follow up if indicated (meeting/correspondence)

Voluntary/Community/ Academic/Other	1	n/a	n/a
	2	Survey	Follow up if indicated (meeting/correspondence)
	3	Survey	Follow up if indicated (meeting/correspondence)

Focus groups will follow ONS best practice for qualitative work. Topic guides will be developed in advance and reviewed by qualitative experts to check for biases in the structure or wording of the questions. Structured questions will be developed to ensure that the required data to inform our analysis are captured. Probing questions will be considered in advance to help to maintain neutrality. We will seek qualitative experts to co-moderate the focus groups.

All surveys will be electronic, using the Citizen Space platform. Questions will be reviewed by qualitative experts and tested by colleagues within ONS. We will ask the overarching organisations and membership bodies we are engaging with to review the questions prior to stakeholders being invited to complete the survey.

For all priority one stakeholders, we will validate the assumptions and values used in the forecasts to check they still apply, and that the data is being used as anticipated. We will also check whether there are any new uses or benefits of Census 2021 data that were not included in the benefit forecast.

For all other stakeholders, we will validate the uses of census data we are already aware of, collect further information about these uses and identify any new uses.

Where use of Census 2021 data is identified, we will explore what stakeholders would do in the absence of Census 2021 data. This will enable us to establish a counterfactual to understand how any benefits from census data are generated.

5.2 Local Government

We expect that the types of benefit generated by census data will be common across local authorities (or, at least, LAs of the same tier and rurality). We consider we have a good understanding of the uses of census data in local government, based on the forecast benefits and ongoing ONS engagement with users. It is unlikely that any evidence shared will be business sensitive. We therefore intend to hold a focus group. This will include priority one LAs, and priority two LAs, .

Participants will be provided with a pre-group information pack at least two weeks before the meeting, outlining the uses, assumptions and values used to create the forecast benefits. This will allow them time to consider the information and collect any relevant evidence. In the focus group, participants will be asked to consider, for each spending decision:

- Is the decision still reliant upon census data? If not, why not?
- Does the estimate of importance in the forecast still seem reasonable? If not, what would you now consider the importance to be (high, medium or low)?

- If Census 2021 data were not available, what would you need to do? Is there an alternative data source you would rely upon and what would the impacts be of using this?

They will also be asked to consider:

- Are there any additional spending decisions that are informed by census data?
- Do you use census data for any other purposes that have not already been described?

Any clarification or follow up required will be conducted either by correspondence or meeting. A note of the meeting will be shared with participants to confirm our understanding of what has been discussed. The data collected will be inputted into the benefits model and an initial estimate of the benefits associated with each use will be calculated. This will be shared with the focus group participants for review.

Following the focus group participants' review, the findings of the assessment will be shared with all remaining local authorities and the local government association. We intend to use a survey to confirm that calculated benefits are reasonable and representative, and to collect information about uses that have been missed.

5.3 Central government

We consider we have a good understanding of the uses of census data in the priority one central government departments, based on the forecast benefits and ongoing ONS engagement with users. Although there may be similarities in data uses between departments, to enable detailed uses and benefits to be understood we intend to meet representatives from each department separately.

Participants will be provided with a pre-group information pack at least two weeks before the meeting, outlining the uses, assumptions and values used to create the forecast benefits. This will allow them time to consider the information and collect any relevant evidence. In the focus group, participants will be asked to consider:

Policy research benefits:

- Does the research or analysis identified in the benefits forecast still take place? If not, what are the reasons for this?
- Is the research or analysis still reliant on census data? (where indicated in forecast, is the percentage attributed to census data still reasonable?) If not, what are the reasons for this? What is the value (commissioned or indicative cost) of the research or analysis?
- If Census 2021 data were not available, what would you need to do? Is there an alternative data source you would rely upon and what would the impacts be of using this?
- Is there any additional analysis/research/modelling that you use census data for that is not described in the forecast?

Evidence-based investments:

- Are the investments identified in the benefits forecast still made by the department? If not, what are the reasons for this?
- Do the investments still rely on Census data? (where indicated in forecast, is the percentage attributed to census data, and the proportion to estimates and attributes, still reasonable?) If not, what are the reasons for this?
- Are the assumptions applied in the forecast still reasonable?
- If Census 2021 data were not available, what would you need to do? Is there an alternative data source you would rely upon and what would the impacts be of using this?
- Are there any additional investments that are, or will be, informed by census data that are not described in the forecast?

Funding allocation:

- In relation to the funds identified in the forecast, have there been any changes to the funds allocated or the formulae or models applied? Are any changes anticipated? If there are changes, what are these attributed to?
- Do the formulae or models still rely on census data? If not, what are the reasons for this?
- Is the estimate of the contribution of census data/ estimates/ attributes still reasonable?
- If Census 2021 data were not available, what would you need to do? Is there an alternative data source you would rely upon and what would the impacts be of using this?
- Are there any other funds which utilise census data to inform allocation which were not captured in the forecast exercise?

Any clarification or follow up required will be conducted either by correspondence or meeting. A note of the meeting will be shared with participants to confirm our understanding of what has been discussed. The data collected will be inputted into the benefits model and an initial estimate of the benefits associated with each use will be calculated. This will be shared with the relevant department for review.

We have identified priority two and three central government stakeholders which may also be able to provide further evidence of benefits. Because our understanding of their uses of census data are more limited, we intend to initially approach these departments with a survey to gain initial information about their uses. This survey will establish:

- If the department uses census or census-based data to inform any research.
- If so, what this research is and what the value (commissioned or indicative cost) of the research is.
- If the department uses census or census-based data to inform any investments.
- If so, what is the investment, the budget and the frequency.
- To what extent does the investment rely on data.
- What proportion of this data is census based?
- What proportion of the calculations involve census data?

- If the department uses census or census-based data to inform any allocation of funds.
- If so, what are the funds and how frequently is the allocation made
- To what extent does the allocation rely on data.
- What proportion of this data is census-based.
- What proportion of this is attributed to estimates data and attributes data
- What proportion of the calculations involve census data?
- If Census 2021 data were not available, what would you need to do? Are there alternative data sources you would rely upon and what would the impacts be of using them?

Where quantifiable benefits are identified and further information is required, we will follow up by meeting and/ or correspondence. We will then input information into the benefits model and share initial estimates of benefits with the department for review, as described above.

5.4 Private sector

The approach taken when producing the benefits forecast was to engage with one representative from industries understood to use census data, drawn from a membership organisation of commercial users of government demographic datasets:

- Retail grocery
- Leisure
- Real estate
- Market research (including Big Data analytics and geodemographic resellers)
- Utilities
- Banking
- Insurance
- Management consultancy
- Advertising
- Direct marketing

We consider that the information in the forecast calculations, coupled with our ongoing engagement with users, means we have a relatively good understanding of the uses of census data in the priority one private sector stakeholder group. However, there have been many factors that may have impacted on uses and benefits since the forecasts were developed, including the Coronavirus (Covid-19) pandemic, Brexit and the growing green economy. We therefore consider that we need to have detailed conversations with priority one stakeholders to explore the benefits. Due to the potentially business-sensitive nature of these conversations, they will be conducted as separate meetings.

Participants will be provided with a pre-group information pack at least two weeks before the meeting, outlining the uses, assumptions and values used to create the forecast benefits. This will allow them time to consider the information and collect any relevant evidence.

For each sector, we will ask:

- Are the business decisions still reliant upon census data or census derived data?
- Are the assumptions applied still reasonable?
- Are there factors that may affect the realisation of the benefit over the appraisal period (2022-2031)?
- If Census 2021 data were not available, what would you need to do? Is there an alternative data source you would rely upon and what would the impacts be of using this?
- Are there any other ways that you use census or census derived data which are not described in the forecasts?

Where quantifiable benefits are identified and further information is required, we will follow up by meeting and/ or correspondence.

For priority two and three organisations, where we have less prior knowledge of their uses of census data, our initial approach will be a survey. We will ask:

- Can you identify any areas of your business which use census data or geodemographic segmentation systems?
- What do you use it for?
- How reliant is that decision/action on data in general?
- What proportion of the data can be attributed to Census?
- If Census 2021 data were not available, what would you need to do? Is there an alternative data source you would rely upon and what would the impacts be of using this?

We will follow up by meeting or correspondence with any priority two or three organisation that provides evidence of a quantifiable benefit, unless sufficient information is included in the survey response.

We will input information into the benefits model and share initial estimates of benefits with the relevant industry representative for review.

Once initial estimates have been reviewed by the contributors and any feedback incorporated, we will scale up the initial estimates to reflect the sector. We will share the initial estimates and scaled up estimates for the private sector with the membership organisation of commercial users of government demographic datasets for review.

5.5 Voluntary sector

Benefits to the voluntary and community sector were not directly assessed in the benefits forecast. We intend to follow the approach described for priority two and three private sector organisations. We will also seek to engage with a membership organisation representing the sector.

Where quantifiable benefits are identified and further information is required to develop assumptions, we will follow up by meeting and/ or correspondence. We will

then input information into the benefits model and share initial estimates of benefits with the contributor.

6.0 Analysis of survey and focus group data

We will assess survey and focus group responses using charting and thematic analysis to identify information that needs to be built into assumptions, figures to include in our model and information that needs further follow up. We will also use qualitative methods such as thematic analysis to understand wider themes relating to the benefits of census data which will help to contextualise our findings.

7.0 Additional sources of data

7.1 Business critical models

Some government decisions are highly model led. Each government department maintains a list of business critical models and is required to document the inputs to these models. In addition to our direct engagement with government departments, we will review published inputs to business critical models to identify key domains based upon census outputs.

7.2 Web scraping

We will explore the use of website scraping techniques to identify additional uses of census data. This may be useful particularly when considering voluntary and community sector organisations, where less detail is already known about their uses of census data. There may be challenges in narrowing down what websites to use the technique with. Findings may still need significant follow up with users to understand the detail of the use and associated benefit.

8.0 Calculation of initial benefit estimates

The benefits forecasts made use of published data, including budgets and external reports, to inform the assumptions made. Where available, the most recent version of this information will be collected. Where a more recent version of a report is not available, we will seek to find an equivalent source.

8.1 Local government

8.1.1 Informing spending decisions

To understand the value of data in general to informing local authority spending decisions, an assumption for the value of big data to the UK economy across central and local government was drawn from a published report. It was assumed that local government accounts for 50% of this value. This was then adjusted for England and Wales (based on the percentage of UK local authorities' GVA accounted for by England and Wales). This was then calculated as a percentage of the total capital and revenue spend on relevant categories, by LAs in England and Wales, per annum.

To inform the forecast benefits, we then sought to determine how much of the value of data could be attributed to census data. Five local authorities in England and Wales were asked to rank relevant areas of capital and revenue investment according to how reliant they were on the census data, as follows:

Lower importance (15% of the value of data assumed to be attributed to census data):
Planning and Development
Waste collection, disposal and recycling
Library Services
Emergency Services

Medium importance (25% of the value of data assumed to be attributed to census data):
Education
Transport
Housing

Higher importance (50% of the value of data assumed to be attributed to census data):
Public Health
Social Care

We intend to confirm with selected local authorities (representing those involved in the forecast development and additional LAs selected to ensure representation across different LA tiers in England and Wales) that the importance of census attributed to each area of spend remains appropriate. We will update calculations with the most recent published capital and revenue investment figures. We will seek to update the value of big data overall by locating a more recent source. If this is not possible, we will increase the original value in line with inflation.

Once we have agreed assumptions with our selected local authorities and they have reviewed our initial benefit assessment, we intend to invite all remaining local authorities to review the assessment to confirm whether it can be generalised to them.

8.2 Central government:

8.2.1 Funding allocation

In previous work, government departments used their modelling to assess the sub-optimal allocation that would result from using the next best available data (rather than census data) to inform funding allocations. This was used to calculate net welfare loss i.e., what is the welfare loss of a pound being spent where it is not needed, compared to where it is needed. Based on this previous work it has been assumed that the net welfare gain when census estimate data is used to inform funding allocations is 0.015% of the total allocated and that the net welfare gain when census attribute data is used to inform funding allocations is 0.053% of the total allocated.

The funding considered for each department was:

Department for Education
Higher needs funding formula
Early years
Higher education
Further education

Department for Business, Energy and Industrial Strategy (BEIS)
Local Enterprise Partnerships funding allocation

Department for Transport
Integrated Transport Block Grant allocation

Department for Communities and Local Government – now Department for Levelling Up, Housing and Communities (DLUHC)
Communities and local government funding

Department for Environment, Food and Rural Affairs
Internal Drainage Board flood defence
Flood defence- Environment Agency

Department of Health - now Department of Health and Social Care (DHSC)
Public health funding
Clinical Commissioning Group allocation
General Practitioner pay funding
Social care funding

Home Office
Police budget allocation

Welsh Government
Local authority funding allocation

Welfare loss was calculated as follows:

Department	Funding allocated	Welfare loss	Justification
Department for Education	Higher needs funding formula- attributes data resource allocation	0.053% of the total funds allocated	Assumption based on previous work
	Higher needs funding formula- estimates data resource allocation	0.015% of the total funds allocated	Assumption based on previous work
	Early years	0.015% of the total funds allocated	Assumption based on previous work
	Higher education	0.015% of the total funds allocated	Assumption based on previous work

	Further education	0.015% of the total funds allocated	Assumption based on previous work
BEIS	Local Enterprise Partnerships funding allocation	0.015% of the total funds allocated	Assumption based on previous work
Department for Transport	Integrated Transport Block Grant allocation (estimates)	change in allocation is 1% of total funds allocated, 3% of this is welfare loss, 70% attributed to estimates	
	Integrated Transport Block Grant allocation (attributes)	change in allocation is 1% of total funds allocated, 3% of this is welfare loss, 30% attributed to estimates	
DLUHC (prev DCLG)	Communities and local government funding (estimates)	0.78% annual budget sub-optimally allocated. 0.33% of this is net welfare loss	Funding model run using alternative data to establish difference in funding allocation.
	Communities and local government funding (attributes)	Est. 15% additional benefit of attributes data 0.78% annual budget sub-optimally allocated. 0.33% of this is net welfare loss	

DEFRA	Internal Drainage Board flood defence	0.1% of the total funds allocated	Based on transport capital allocation
	Flood defence-Environment Agency	0.1% of the total funds allocated	Based on transport capital allocation
DHSC (prev DH)	Public health funding	2% of budget sub-optimally allocated, 1.9% of this is net welfare loss	
	Clinical Commissioning Group allocation	0.067% of total budget is net welfare loss	
	General Practitioner pay funding	0.015% of the total funds allocated (estimates)	Assumption based on previous work
	Social care funding	0.053% of the total funds allocated (attributes)	Assumption based on previous work
Home Office	Police budget allocation	0.015% of the total funds allocated (estimates)	Assumption based on previous work
Welsh Government	Local authority funding allocation	1.8% of total budget sub-optimally allocated. 0.8% of this is welfare loss	Funding model run using alternative data to establish difference in funding allocation.

We will engage with stakeholders to verify that the funding is still in place and that it still relies on Census data. We will seek to understand whether there have been any changes in the funding formulae and whether the assumptions are still reasonable. If not, we will adjust the assumptions as required.

We will explore with stakeholders the feasibility of running funding models with Census 2021 data and the next best data source to establish the difference in allocation of funds and inform the calculation of the welfare loss as detailed above. If this is not possible, we may need to rely on expert opinion to estimate the likely allocation difference, considering the difference seen when the exercise was previously undertaken and what is known about Census 2021 data compared to the next best data source.

8.2.2 Policy research

The granular level data census provides can improve the productivity of the policy making process by reducing time spent by analysts, reducing costs in collecting data and acquiring commercial data, or reducing risk. In the forecast benefits, the commissioned cost of research based on census or census-derived data was used as a proxy for this value, in the absence of other information. In instances where stakeholders indicated the absence of census data would require the department to undertake their own research or rebuild their models, the benefit was assumed to be the cost of doing so.

The funding considered for each department was:

Department for Education
public policy research dependent on census data
Benefit from avoiding spend on assessing complaints and business cases on appeals about pupil projections

Department for Business, Energy and Industrial Strategy (BEIS)
Sub regional fuel poverty modelling
UK Research and Innovation/ Economic and Social Research Council
- Spatial Economic Research Centre
-Census sponsored research
-Centre for Urban and regional development studies

Department for Transport
National Travel Survey

British Social Attitude Survey

Research to replace census data

Evaluations using census data

Updates to national trip end model- spend is cost of maintaining the model, completed every three years to three times within the assessment period

Transport Social & Distributional Impact analysis- benefit equal to value of the research

Transport other analysis (estimates)- 10% of annual analytical budget-50% of value of census data attributed to estimates

Transport other analysis (attributes)- 10% of annual analytical budget-50% of value of census data attributed to estimates

Transport general national scenario modelling- one off cost to rebuild model of, split over 10 years.

Department for Communities and Local Government – now Department for Levelling Up, Housing and Communities (DLUHC)
English Housing Survey
Other research

Department for Environment, Food and Rural Affairs
Rural Evidence Research Centre
Inequalities in Health and Rural Areas
Rural Evidence Hub reports
Carbon Reduction Commitment Energy Efficiency Scheme (CRC) research and publications

Department for Digital, Culture, Media and Sport (DCMS)
Taking Part Survey
Arts Council
English Heritage
Sports England

Department for Work and Pensions (DWP)
Policy research

Home Office
Public Sector equality duty
Think Tanks and policy research unit- cost is taken to be wage bill for the unit
Cost of migrant survey to replace Census Data- survey would be completed every 3 years

Samples of Anonymised Records (SARs)
Policy research

We expect to use the same methodology to assess the realised benefits. The data we collect from stakeholders about what they would need to do in the absence of census data will help us to build a counterfactual which may enable us to better understand and calculate the benefit, rather than relying on the cost of research as a proxy.

8.2.4 Evidence-based investments

We assessed the delay expected to projects such as major transport schemes as a result of not having census data and estimated the associated costs for the delay.

Department for Education
Funding for teacher training- benefit of data is 10% of the spend, 10% of the value of the data can be attributed to census data

Department for Transport

Transport LA major schemes (estimates)- estimate average two week delay due to need to gather equivalent data, leading to inflation costs, increased appraisal costs and lost opportunity cost. Estimates contribute 50% of value of Census data

Transport LA major schemes (attributes) Attributes contribute 50% of value of census data used in appraisal

Transport HE major schemes (estimates)- as above

Transport HE major schemes (attributes) as above

Transport national Rail Model (estimates) as above

Transport national rail model (attributes) as above

Transport Analytical support for HS2 (estimates)- 10% of cost of two week delay- 50% of value of census data attributed to estimates

Transport Analytical support for HS2 (attributes)- 10% of cost of two week delay- 50% of value of census data attributed to attributes

8.3 Private sector

Some benefits to the private sector can be attributed directly to census data. In other cases, census data is an intermediate input to geo-demographic resellers. We assumed in our benefit forecast, based on information collected from the sector, that 20% of any benefit generated from using commercial geo-demographic software can be allocated to census data.

To inform our benefit forecast, we engaged with one representative from each industry (Insurance, Leisure, Advertising, Real Estate, Utilities, Geo-demographic resellers, Direct marketing, Banking, Management consultancy, Retail and Market research) drawn from a membership organisation of commercial users of government demographic datasets. Each representative identified a percentage value of data in relation to each relevant decision or activity. They then identified the portion of data attributable to census and to census-based geo-demographic software. We used this to calculate an initial benefit estimate for the representative organisation and scaled this up for their sector, taking account of how homogenous the sector was.

The detailed methods used to calculate each forecast benefit are detailed in Annexe A.

Overall, we expect to replicate the same method to calculate the benefits to the private sector. We recognise that may have been considerable change in some sectors and we will ensure detailed verification of all assumptions with stakeholders. Where assumptions are no longer reasonable, we will work with stakeholder representatives to revise or formulate new assumptions.

We recognise that the level of homogeneity within an industry or sector will impact on how readily an initial estimate based on a small number of organisations can be

scaled up to reflect the wider industry. Survey responses may help us to understand more about how heterogenous each industry is, at least to some degree. We are currently exploring how this scaling up might be best approached, given the limited capacity for the project to collect data from a larger number of organisations.

8.4 Capturing changes to forecast benefits and new benefits

The most recent stakeholder engagement exercise to develop the forecast benefits was in 2016 and 2017. Some of the assumptions were based on information collected in 2013. Many factors may have impacted on uses and benefits of census data since these forecasts were developed. These include the Coronavirus (Covid-19) pandemic, Brexit and the growing green economy. In our engagement with users who took part in the forecast benefit exercise, we will establish whether any of the existing benefits in the forecast have substantially changed or are no longer applicable. We will also seek to identify uses of Census 2021 data that have emerged or were not captured in the benefits forecast. These may arise from stakeholders who engaged with the benefit forecast exercise and have begun to use census data in additional ways. New benefits may also arise from our expanded stakeholder base.

Census 2021 data will be available to users in different ways to previous censuses. The integrated data service will expedite the availability of underlying data from Census 2021, in more accessible and interoperable formats, so that statistical outcomes can better inform policymakers. These data will be made available with more detail, and many years earlier, than for previous censuses, which will significantly increase their use, and the benefits generated.

We anticipate that most new uses will still fall into the primary categories of benefit identified in the forecast exercise, such as policy research, informing business or spending decisions, evidence-based investment or funding allocation. We therefore intend, in the first instance, to adopt similar approaches to those used for existing comparable benefits. All assumptions will be verified with the users before calculating the benefit.

9.0 Adjustments

9.1 Long tail

Some international assessments of the benefits of census data have included an estimate of the additional monetary value of the long tail of smaller benefits that have not been directly quantified. We will give consideration, based on the findings of our stakeholder engagement, to whether this may be appropriate. This would be a change to the method used in the forecast, which recognised that the forecast benefits may be underestimated, but did not seek to account for benefits that were not quantified directly.

9.2 Optimism bias

We will assess the degree of optimism bias by considering the level of confidence over the assumptions for each benefit. This will consider factors such as whether the assumptions and data were drawn directly from the stakeholder, how recent

figures are and whether the assumptions calculated directly or based on expert opinion.

We will attribute degrees of optimism bias to each benefit as follows:

High confidence	10% optimism bias
Medium confidence	20% optimism bias
Low confidence	40% optimism bias

Scottish Enterprise (a sponsored non-departmental public body of the Scottish Government which encourages economic development, enterprise, innovation and investment in business) recommends that where projects are discrete and financed by the public sector, an optimism adjustment of 0-40% is likely to be appropriate. This was the rationale used in the benefits forecast.

9.3 Sensitivity

For each benefit, the level of prudence applied will be assessed, based upon whether it is likely there is related activity that has not been captured, where figures were drawn from and the likelihood of the actual costs or expenditure being higher or lower than assumed.

We will attribute the degree of prudence as follows:

High	130%
Medium	115%
Low	100%

We will then consider the benefits assessment at varying levels of optimism bias and prudence, as well as those modelled, to consider the range within which the actual realised benefits are likely to fall.

9.4 Taking account of time

9.4.1 Decline in data quality

It was assumed in the benefits forecast that the value of census data decays at a rate of 5% per annum on a straight-line basis. This was thought to be a reasonable estimate as a lower limit would be 0% and an upper limit 10% (given that the Census takes place every ten years). It is our expectation that we will similarly apply this assumption. We will, however, review this if evidence collected from users suggests that the expected data quality over time has changed.

9.4.2 Discounting

The calculated forecast benefits were discounted at 3.5%, applied at an aggregate level. We expect to apply the same approach. The values are nominal, and reduce over time, so discounting could be inferred to be implicit.

Annexe A: Detailed method for private sector benefits

Industry	Benefit	Justification	Forecast				Proposed method
			Assumption 1	Assumption 2	Assumption 3	Assumption 4	
Banking	**Location planning/ refurbishment**	Census and census-based data help to make better decisions about where to invest in refurbishment and relocation. Absence of data would lead to poor investment decisions and ultimately branch closure, so investment lost. The benefit is equal to the proportion of lost investment attributable to census data.	£40-50m per year spent by one industry representative [drawn from stakeholder engagement]. Scaled up for sector, based on market share.	adjusted to the yearly spend in England and Wales based on the England and Wales share of Financial Services Industry (SIC64).	Absence of direct census data will lead to 1% of branches closing [drawn from stakeholder engagement].	Absence of indirect data (of which 20% is reliant on census) will lead to a further 10% closing [drawn from stakeholder engagement].	Base method expected to remain the same, but assumptions may need considerable review
	Location divestment planning	Census and census-based data inform bank closures, creating efficiencies. In the absence of data, some of these efficiencies will not be realised. The benefit of the data is equal to the amount that would not be realised	A single bank closure was estimated to create 100k efficiencies per annum, (£1m over 10 year period) [drawn from stakeholder engagement].	Based on published rates of yearly bank closures, it was assumed that banks and building societies would shut at rate of 300 per year. An even profile of 300 additional closures per year was assumed, in order to calculate total efficiencies achieved.	Absence of direct census data will lead to 1% of these efficiencies not being realised [drawn from stakeholder engagement].	Absence of indirect data (of which 20% is reliant on census) will lead to a further 10% of these efficiencies not being realised [drawn from stakeholder engagement].	
Cross-sector advertising	Cross-sector advertising	Census data informs decisions about targeting advertising in the B2C market. A proportion of the advertising revenue for defined advertising is assumed to be attributable to census	The advertising revenue for defined advertising in the UK was adjusted for England and Wales using the England and Wales share of the GVA of the advertising and	The B2C share of magazine advertising was identified, and the same split was assumed for all advertising.	It was assumed that data was worth 3-4% of the value of the business decision (an average figure of 3.5% was used in calculations).	It was assumed that 5% of this could be attributed to census data because it does not rely on small area data [drawn from stakeholder engagement].	We expect to replicate the same method, following verification of assumptions with stakeholders

		data, which is equal to the benefit.	market research industry		[drawn from stakeholder engagement].		
Cross Sector Direct Marketing	*Cross Sector Direct Marketing*	Census data is used to target direct marketing efforts. The benefit is equal to the proportion of the total spend on B2C marketing that is attributable to census data.	The total UK spend of direct marketing was identified and adjusted for England and Wales based on the England and Wales share of the GVA of the Professional, scientific and technical sector	53% of companies use direct marketing to customers, spending on average 70% of their direct marketing budget on consumer marketing	3-4% (an average figure of 3.5% was used in calculations) of the value of business decisions related to B2C marketing was dependent on data analysis	20% of this was attributable to census data.	We expect to replicate the same method, following verification of assumptions with stakeholders
Insurance	*Life insurance*	The life insurance industry relies on CMI life tables, which are reliant on census data. The benefit is assumed to be equal to the proportion of GVA of the industry that is attributable to CMI tables.	The GVA of the life insurance sector in England and Wales was assumed to be 37% of the England and Wales share of the GVA of the Insurance and Pension Funding Industry (SIC 65), given that BRES employment data indicated that 37% of the sector (by employment) is Life Insurance.	0.1% of this GVA is reliant on CMI tables [drawn from stakeholder engagement]. .			We expect to replicate the same method, following verification of assumptions with stakeholders
	General Insurance	Census data is used to prevent fraud through analysis to identify traits predictive of fraud. The benefit is assumed to be equal to the proportion of the value of detected fraud cases attributed to census and census derived data.	The value of detected fraud cases in the UK was identified and adjusted for England and Wales using the England and Wales share of the GVA of the Insurance and	2.5% of this value is directly attributable to Census data [drawn from stakeholder engagement].	a further 10% is attributable to indirect data (of which 20% is reliant on census)		We expect to replicate the same method, following verification of assumptions with stakeholders

			Pension Funding Industry.				
Leisure	Leisure	Census data is used to inform leisure location planning. The benefit could be considered to be equal to the proportion of investment in new outlets, or the proportion of the GVA of the industry, attributed to census data. The forecast benefit was calculated as the average of these estimates.	Low estimate: the total value of the UK leisure estate was adjusted for England and Wales using the England and Wales share of the GVA of the Sport, Amusement and Recreation Activities Industry (SIC 93)	investment in new outlets would be the same as in retail, and census data would contribute 1.2% of this			We expect to replicate the same method, following verification of assumptions with stakeholders
			High estimate: The England and Wales share of the GVA of the arts, entertainment & leisure industry (SIC R)	0.1% is attributable to census data. [drawn from stakeholder engagement].			
Market Research (inc Big Data Analytics and Geodemographic Resellers)	Market Research (inc Big Data Analytics and Geodemographic Resellers)		The GVA of the UK market research sector (identified from a published source), less the GVA associated with the private sector, central government and local government to prevent double counting. We also disregarded the percentage of research attributable to business research.	adjusted for England and Wales based on the England and Wales share of the share of the GVA of the Advertising and Market Research Industry (SIC 73).	Data was assumed to contribute 30% of the remaining GVA [drawn from stakeholder engagement].	between 5 and 15% (10% used in calculations) of which is assumed to be attributable to census data. [drawn from stakeholder engagement].	We expect to replicate the same method, following verification of assumptions with stakeholders

Real Estate	Real Estate	Census data is used in house planning, to provide evidence that it meets policy and local need.	The number of houses built in England and Wales, multiplied by the average product fee for a small-scale scheme of around 10 houses [drawn from stakeholder engagement], is used as a proxy for the overall value of the sector	data contributes 10-50% of this value (30% used in calculations) [drawn from stakeholder engagement].	30-70% (50% used in calculations) of the value of data is attributed to census [drawn from stakeholder engagement].		We expect to replicate the same method, following verification of assumptions with stakeholders
Retail Grocery	Retail Grocery	Census data informs where to locate new stores, avoiding lost capital investment.	The average capital investment in the UK (sourced from competitor tracking database for the major grocery retailers) was adjusted for England and Wales based on the England and Wales share of the GVA of the retail trade industry.	2.5% stores would close without direct census data [drawn from stakeholder engagement].	a further 25% would close without indirect census data, (of which 20% is census based) [drawn from stakeholder engagement].		We expect to replicate the same method, following verification of assumptions with stakeholders

Utilities	Water and sewerage companies	Census data is used in planning investment and monitoring leakage	total capital investment in England and Wales was calculated from planned water expenditure by UK water and sewerage companies, excluding on waste water, adjusted for England and Wales using the England and Wales share of the GVA of the Water supply and Sewerage industry (SIC 36-37).	contribution of Census data was assumed to be 1% of the total capital investment [drawn from stakeholder engagement].			We expect to replicate the same method, following verification of assumptions with stakeholders
	Energy Company Obligations	Energy companies use census data to identify areas with a high proportion of vulnerable households to target their Energy Company Obligations.	The total investment in energy efficiency schemes under ECOs was adjusted for England and Wales using the England and Wales share of the GVA of the Electricity, gas, steam and air conditioning supply industry (SIC 35)	value of data was assumed to be 10% of the investment [drawn from stakeholder engagement].	33% of the value of data was attributed directly to census data [drawn from stakeholder engagement].	57% of the value of data is attributed to indirect census data, 20% of which is based on census [drawn from stakeholder engagement].	Base method expected to remain the same, but assumptions may need considerable review

	Gas and Electric Energy Saving Solutions (e.g. location of electric charging points)	Census data informs locations for investment	One of the “Big 6” energy suppliers estimated that £10m per annum was spent by the supplier on solutions in the UK. This was adjusted for England and Wales using the England and Wales share of the GVA of the Electricity, gas, steam and air conditioning supply industry (SIC 35)	The value of data was assumed to be 10% of the investment [drawn from stakeholder engagement].	33% of the value of data was attributed directly to census data 57% attributed to indirect census data, 20% of which is based on census [drawn from stakeholder engagement].	This was scaled up to account for the six “Big 6” energy suppliers	Base method expected to remain the same, but assumptions may need considerable review

Annexe B: Contributor Stakeholder List by Priority

Stakeholder group	Organisation	Benefits Priority
Local Government MVP	Camden	1
	City and County of Swansea (Cyngor Sir a Dinas Abertawe)	1
	Hampshire County Council	1
	Manchester City Council	1
	Sheffield City Council	1
Local Government Extended Base	Brent	2
	Ceredigion County Council (Cyngor Sir Ceredigion)	2
	Cornwall Council	2
	Coventry City Council	2
	Norwich City Council	2
	Oxford City Council	2
	West of England	2
Local Government (Priority 3)	All other LAs	3
	NHS Integrated Care Boards	3
Central Government MVP	Department for Business, Energy and Industrial Strategy	1
	Department for Education	1
	Department for Levelling Up, Housing and Communities	1
	Department for Transport	1
	Department for Work and Pensions	1
	Department of Health and Social Care	1
	Home Office	1
Central Government Extended Base	Cabinet Office	2
	Cabinet Office - Government Equalities Office	2

	Cabinet Office - Race Disparity Unit	2
	Cabinet Office- Equality Hub	2
	Cabinet Office- Disabilities Unit	2
	Cabinet Office- Office for Veterans Affairs	2
	Chief Operating Officer's Group	2
	Climate Change and Rural Affairs Group	2
	Covid Recovery & Local Government Group	2
	Data Cymru	2
	Department for Environment, Food and Rural Affairs	2
	Digital Health and Care Wales	2
	Economy, Treasury and Constitution Group	2
	Education, Social Justice and Welsh Language Group	2
	Equality and Human Rights Commission	2
	Health and Social Services Group	2
	Ministry of Defence	2
	Ministry of Justice	2
	National Infrastructure Commission	2
	National Infrastructure Commission for Wales	2
	NHS England	2
	NHS Wales	2
	Senedd Cymru (Welsh Parliament)	2
	WG Data and Digital team	2
Central Government (Priority 3)	Academy for Social Justice	3
	Advisory Committee on Resource Allocation for Health	3
	Association of Ambulance Chief Executives (AACE)	3
	Association of Police and Crime Commissioners	3

	Boundary Commission for England	3
	Boundary Commission for Wales	3
	Cabinet Office - Geospatial Commission	3
	Cabinet Office - Government Digital Service	3
	Children's Commissioner for Wales	3
	Climate Action Strategy Cabinet Committee	3
	Community Housing Cymru	3
	Council for Economic Development	3
	Department for Digital, Culture, Media and Sport	3
	Department for International Trade	3
	Dept BEIS - Cities and Local Growth Unit	3
	Domestic and Economic (Strategy) Cabinet Committee	3
	Education and Skills Funding Agency	3
	Foreign, Commonwealth and Development Office	3
	Future Generations Commissioner for Wales	3
	Government Actuary's Department	3
	Government Office for Science	3
	Government Priorities Delivery Cabinet Committee	3
	High Speed Two (HS2) Limited	3
	Higher Education Funding Council for Wales	3
	HM Revenue & Customs	3
	Home Affairs Committee	3
	Institute for public policy research	3
	Migration Advisory Committee	3
	National Fire Chiefs Council	3
	National Infrastructure Commission	3

	Network Rail	3
	NHS Business Services Authority	3
	Office of the Immigration Services Commissioner	3
	Office of the Secretary of State for Wales	3
	Older People's Commissioner in Wales	3
	ONS Sustainable Development Goals Team	3
	Prime Ministers Office - via Cabinet Office	3
	Public Health Wales	3
	Rural Services Network (LGA)	3
	Social Mobility Commission	3
	The National Archives	3
	Transport for London	3
	Transport for Wales	3
	UK Health Security Agency	3
	Water UK	3
	Welsh Ambulance Service NHS Trust	3
	Welsh Language Commissioner	3
Business MVP	Demographic User Group (M&S, Boots etc.)	1
	CACI	1
	Barclays	1
	Co-Op	1
	Deloitte	1
	Eon	1
	Experian	1
	Faculty of Institute of Actuaries	1
	GSK	1

	Market Research / Big Data Independents Jame Furness	1
	Market Research / Big Data Independents: Barry Levenhall	1
	Marks and Spencer	1
	MRS Census and Geodemographers User Group	1
	Sainsburys	1
	Swinton Group	1
Business extended base	British Chambers of Commerce	2
	Confederation of British Industry (CBI)	2
	Dunnhumby	2
	Google	2
	Livin	2
	Wood Mackenzie	2
Business Priority 3	Axa PPP Healthcare	3
	British Property Federation (BPF)	3
	Call Credit	3
	Data Talk	3
	Equifax	3
	Federation of Small Businesses	3
	Gartner	3
	Ipsos Mori	3
	Kantar	3
	MDDUS	3
	National Association of Property Buyers	3
	National Residential Landlords Association	3
	RWE group	3
	Transunion	3

Voluntary/Community/Academic/ Other Extended Base	British Heart Foundation	2
	Cancer Research	2
	Centre for Cities	2
	CoE	2
	Core Cities	2
	Health Data Research (HDR UK)	2
	Institute for Fiscal Studies (IFS)	2
	Inter-Faith Network	2
	Local Enterprise Partnerships Network	2
	Macmillan	2
	NCVO	2
	Overseas Development Institute (ODI)	2
	Renewable Energy Association	2
	Renewable UK	2
	Resolution Foundation	2
	Salvation Army	2
	WCVA	2
Voluntary/Community/Academic/ Other Priority 3	Academic Network of Disability Experts (ANED)	3
	Age UK	3
	British Society for Population Studies	3
	Care International UK	3
	Carers Trust	3
	Carers UK	3
	Centre for Population Change	3
	Chwarae Teg	3

	Connecting generations partnership	3
	Countryside Alliance	3
	Disability Rights UK	3
	E3G Think tank	3
	Economic and Social Research Council	3
	Economic Statistics Centre of Excellence (ESCOE)	3
	Higher Education Statistics Agency	3
	International Organisation for Migration	3
	Kings College London (KCL)	3
	Kings Fund	3
	Liverpool Hope University (Centre for Culture & Disability Studies)	3
	Llamau	3
	London School of Economics (LSE)	3
	London School of Economics (LSE) Care Policy and Evaluation Centre	3
	Migration Policy Institute	3
	Migration Watch UK	3
	National Institute of Economic and Social Research (NIESR)	3
	Nuffield Trust	3
	Oxford DataLab (Nuffield Department of Primary Care Health Services)	3
	Queens University Belfast	3
	Royal British Legion	3
	Race Council Cymru	3
	Royal Economic Society	3
	Royal Society	3
	Royal Statistical Society (RSS)	3
	Royal Town Planning Institute	3

	Shelter	3
	Shelter Cymru	3
	Society of Genealogists	3
	Southampton University	3
	Swansea University	3
	Tai Pawb	3
	The Alan Turing Institute	3
	The Equality Trust	3
	The Migration Observatory	3
	The Wallich	3
	The Young Foundation	3
	UK Research and Innovation (UKRI)	3
	University of Kent - Personal Social Services Research Unit (PSSRU)	3
	University of Leeds (Centre for Disability Studies)	3
	University of St Andrews	3
	Urdd Obaith Cymru	3
	Wellcome Trust	3