

## **Central Government submission to the ONS' public consultation: "The census and future provision of population statistics in England and Wales"**

### **5.1 General comments**

**Q1:** *What are your views of the different census approaches described in this document?*

Government recognises the value of the census and its history as a bedrock of statistical infrastructure for over 200 years. The census in its current form provides a good record of the size, location and characteristics of the population at a point in time based on the fact that it is completed by most households. There are long existing time series, providing detailed local information, and the census enables historical research by both professionals and amateurs.

At the same time central government also recognises that there are clear benefits to modernising the current census approach and that there are improvements to be made. The present, mainly paper-based decennial approach could be improved for the future to take advantage of developments in technology, for example to improve processing times, and to take advantage of potential advances in data-sharing, to better utilise information already collected by government. Whichever census approach is adopted should aim to significantly improve the speed of analysis and outputs.

The census in its current form is expensive and we need to realise the potential time and financial savings from an alternative approach that modern technology could facilitate. Nor are we harnessing the power of the considerable data that government already collects. By running the census just once every ten years we are losing the opportunity for more frequent data collection at closer intervals, and are instead relying on data that can be up to ten years old.

With these overarching considerations in mind we have set out the strengths and weaknesses of the different census approaches below.

#### **Strengths of option A (An online census once a decade)**

##### **1. Methodology**

Having sound population estimates is essential for modern government. The overriding strength of the census is that it provides detailed data which tells us not only about the numbers of people living in particular areas, but also gives detailed information about them (characteristic data) which is used by policy makers across government to determine levels of service provision offered to people.

The proposal to move to a primarily online approach could offer improvements over traditional census methods. ONS already has experience of offering census completion electronically and many other countries have also adopted this approach, so there is a growing community of expertise. The move to an online census has some risks associated with it but these are less significant than option B (using administrative data and compulsory annual surveys). It should be noted that option A

poses fewer methodological challenges. The potential for inbuilt validation checks and prompts to provide the opportunity to improve the data and produce outputs quicker is also attractive.

## 1. Geographical detail

Option A would continue to provide data by 'Output Areas' which are important for some government departments. For example, option A provides detailed cross tabulations of population and characteristic data at Output Area Level needed for rural proofing of Government policy. Option A hence enables policy makers to take a localised approach, for example delivery organisations can understand the numbers and characteristics of disabled people in their local area.

As option A would continue to enable information to be available at Output Area level, non-standard geographies could continue to be built from these base building blocks. It gives an accurate picture of individuals within small area geographies. Increasingly local authorities and Local Enterprise Partnerships need data at small geography level. Furthermore, Output Area level data forms the building blocks for other geographies that are not constrained to standardised geographies, e.g. National Park areas.

An online census once a decade would give accurate point estimates for detailed questions and in particular sub-groups within the population at very low geographic areas.

### Weaknesses of option A (an online census once a decade)

#### 1. Cost

The costs and operation of a full census covering the entire population are significant, and these have to be considered. option B may have greater potential for lowering these costs than option A. However, more detailed work on the full respective costs and benefits of the two options is still required.

#### 2. Frequency

The current paper-based decennial census is infrequent and does not provide regular data and this would continue to be an issue under option A. An online decennial approach also would not address one of the key problems with the current census – that data is not updated regularly, so census data becomes more out of date the further away it is from the date of collection<sup>1</sup>.

---

<sup>1</sup> The fact that data are less reliable towards the end of the 10 year period mainly affects large urban cities where migration into and out of cities is significant. That said, low levels of migration into small rural communities can have a large impact and consequences to the services and needs of that community, and these changes would not be picked up quickly by the census option, online or otherwise.

The delay from census day to publication also means that during the initial two to three years, when the census data is at its most reliable, it cannot be used and instead data from the previous census is updated. By then it is at its most unreliable as it is then between eleven and thirteen years old. A largely online census could potentially improve timeliness and thereby reduce this later period of high unreliability. Improving timeliness could be considered an improvement in value for money.

### 3. Non-response

The census is affected by non-response, and gaining co-operation has become more difficult. We imagine that this is likely to remain a problem even if the census was primarily conducted online. One unknown is whether the potential on-line completion rates will be high enough in 2021. Differentiating between those who ought to respond online and those permitted to use alternative methods will be challenging, particularly where the segments of the population of most interest (for example, migrants) may be the groups least likely to respond online.

For a census conducted primarily online, consideration would need to be given to rural communities. Currently a substantial proportion of the rural population does not have adequate broadband provision. However, this is being addressed by Government, with the intention to provide superfast broadband to at least 90% of premises in the UK and to provide universal access to standard broadband with a speed of at least 2 megabits per second. There is additionally the issue of older people being able to use the internet, and there being proportionally more older people in rural areas, which could risk the reliability of rural data, over and above any risk that there are still rural households unable to access the internet by 2021.

As noted in the ONS consultation document, there would need to be options for paper completion otherwise there is a risk of under-representing the most rural households and other groups unable to respond online.

The current follow up checks and census assistance would need to remain in place and may even need to be enhanced.

### Strengths of option B (a census using administrative data and surveys)

#### 1. Cost

Option B has good potential for reduced costs and more regular data. Again, it will be necessary to assess the full costs and benefits of this option. This will include ensuring we have an accurate view on the cost to government departments of extracting and transferring administrative data to meet ONS requirements.

Existing work between DWP, ONS and HMRC has demonstrated that census requirements are likely to require new development or extraction processes, cleansing and transformation, additional storage and secure data transfer processes. The costs of these requirements need to be understood.

## 2. Frequency

Many government departments see merit in using administrative information to produce population estimates in the period between censuses. This would provide more dynamic and responsive estimates than currently available. Given large intervals between traditional censuses, more frequently updated data could be beneficial, however this is dependent on the accuracy of the data.

The availability of survey data on an annual basis would have some benefits, in providing a more up to date picture of the population, particularly at Local Authority and Clinical Commissioning Group level.

A large scale annual survey to capture characteristics of the population and households would fill as much as possible of the gap that a decennial census would leave, and if coverage is large enough, may improve on existing sample surveys. A new mandatory annual survey, initially of 4% of the population, is a new burden but if this replaced the census for all individuals it would be a net relief of burden on the public.

Statistics based on three years' worth of survey data would smooth out change, as opposed to point in time estimates produced by the traditional census approach. However the accuracy of these estimates will necessarily vary by age group – for example in 2012 nearly one in four 19 year olds had moved Local Authorities in the previous 12 months. This will impact on the estimated reliability of the statistics. Also, there are challenges with interpreting data collected over multiple years.

## 3. Promoting the use of administrative data

The use of administrative datasets within the Census would hopefully help drive up the quality and understanding around this type of data which is increasingly being used for a range of analyses in departments.

### Weaknesses of option B (a census using administrative data and surveys)

#### 1. Implementation risks from a new approach

Option B is considered to be higher risk due to the uncertainty of outputs for small area geographies (over and above the reduced geographical detail discussed in section 2 below) and the timescale involved for implementation. There is a concern that if option B is not successful, option A or the current census strategy may not be recoverable.

There is a risk in discontinuing the current decennial census without having tested the new administrative data plus surveys method in full, for example by using both methods in 2021. It would be normal practice to seek to fully test the new methods before deciding to change methodology. The methods for population estimates using administrative data may not be sufficiently mature at this point to support a move away from traditional methods, whatever the benefits might be.

Option B also relies on extensive data matching. Matching data is complex and can be imprecise. More work on assessing feasibility would be helpful in considering the options.

Some felt that the timing may be unrealistic to account for such a significant change in methodology and continuity of data, for example it was highlighted that the New Zealand Government have started their census options review with an expectation of introducing a new system by 2030.

More information regarding risks is detailed in section 5.5.

## 2. Geographical detail

Departments had significant concerns about the negative impacts of losing more detailed small area data. The rest of this response provides more detail.

Departments are of the view that further investigation and testing is needed on how more local data needs could be met. One suggestion was made as to whether ONS could increase the size of the sample survey, although this would increase cost.

Under option B the most detailed and reliable data available will be at local authority level. Non-standard geographies would be approximations of Lower Layer Super Output Area (LSOA) level data at best. Parish level data is becoming increasingly important within the localism agenda, e.g. for neighbourhood plans, but some parish level data would not be available if it is smaller than LSOA.

Characteristic data about the population will not be available at Output Area Level, even using 5 years' worth of data. Some clarification is needed as to what small level geography data will be available at OA level – in some cases it is suggested that *limited* estimates by age and sex may be available for small groups of postcodes with *limited* cross tabulations for customised groups of OAs but more research is needed. In other cases it states that population by age and sex may be available for *all* OAs but this has not yet been fully tested.

Another concern is a loss of the additional data that the census provides at OA level. The characteristic data includes categories such as ethnicity, employment status, marital status, household data, health and travel to work. In addition, where estimates are used in place of census data, the precision of estimates for very small population groups would be noticeably lower than present census data.

We understand that this approach also reduces the geographical granularity of data (at least in the short run), which is an important consideration for a number of government departments in the allocation of resources and for the identification of geographic areas for policy and strategy formulation. We would like to be reassured that the proposed solution to that problem (aggregating data across years) provides data of sufficient quality, particularly for two and three dimensional cross tabulations of variables (e.g. ward level data for cross tabulations including religion, ethnicity, nationality, and where born), and not just for the headline figures.

### 3. Drawing on administrative data sources

Many administrative sources proposed are not completely accurate; most have been designed for operational purposes and are therefore susceptible to operational practices – they may lack some of the information required. It is unclear how the administrative data option will be quality assured and what options are available. In addition, it is likely that the compulsory annual surveys would suffer the same issues of non-response for key groups as the full census.

There are some additional issues with administrative data sources to be taken into account. Administrative sources under-estimate groups who migrate and don't register with GPs or use benefits. There are also difficulties estimating for elderly people in care homes and areas with second homes. Administrative sources may change with time, even those owned by government departments, with the potential loss of key data.

Data on dwellings and household structures available through the current census methodology may not be as richly replicated via administrative data alone. These are important for understanding the regional economies.

### 4. Potential impacts on the health system

Option B could have adverse impacts on the health system, which uses data for the central allocation of budgets and local service delivery. The population estimates derived from administrative data differ markedly from the census figures for a significant minority of areas, particularly for some age groups. Changes to administrative systems, processes and incentives will need to minimise any discontinuities in the data which prevent meaningful comparisons over time. Resource allocation formulae make use of information from Lower Super Output Areas. A 4% survey option would only offer limited information at this level even when results are produced by combining three or five years' data. The resultant allocations would be less accurate and may be more volatile. The lack of availability of information for very small areas would also impact on local service planning and delivery, and on epidemiological work. Also of concern is the timeliness with which the first set of small area characteristics data would become available.

### 5. Under-represented groups

Several departments have a particular interest in how well this new approach will provide information on some subgroups. For example, current administrative sources are extremely poor in identifying migrants, whether these are residents who were once born abroad or foreign nationals, and it therefore seems unlikely that this approach will deliver data on these categories of people. There has been a significant benefit from the new questions in the 2011 traditional census, reflecting the increased public interest in migration as a topic over the past decade.

Some under-represented groups in the existing census may become more so by using administrative sources. Some groups may limit their visibility and are likely to

be as unresponsive to the new surveys as they are to the census and will almost certainly be even less visible in the administrative data. The evidence from other countries who use registration systems at present indicates that migrants in particular are difficult to track via administrative systems. This is because emigrants are often slow to deregister from systems and immigrants are often slow to register.

#### Combining options A and B – ‘dual running’

Pursuing administrative data options and an online approach need not be mutually exclusive. In addition, there are concerns that ONS should be cautious about adopting a “Big Bang” approach. A period of dual running, as envisaged in page 13 of the consultation document, was suggested as a way of managing the risks around changing to an administrative data census.

Another option would be to retain the full census and introduce interim updating, perhaps using administrative data together with a method for correcting for known gaps in that data and with the proposed annual surveys of 4% of the population.

## 5.2 Uses and benefits of population and housing statistics

**Q2:** *Please specify any significant uses of population and housing statistics that we have not already identified.*

A number of government departments identified additional significant uses of population and housing statistics that ONS had not already identified. These are set out by department below:

### Department for Business, Innovation and Skills

- BIS uses census data to feed into the UK Assisted Areas map. This map identifies those areas of the UK where regional aid may be given under EU legislation. The work relies on detailed geographical data. Guidance from the European Commission specifies which level of geography has to act as a building block and for the most recent exercise this was wards. Census 2011 data has been used for the current exercise.
- Census data also provides detailed small area information which is important for research on social mobility, in particular for determining participation rates in Higher and Further Education. The spatial dimension is particularly important in looking at geographic as well as social mobility.

### Cabinet Office

In common with many other users of population statistics, population estimates (and by extension census data) are predominantly used as a denominator for a range of statistics and management information. Cabinet Office also relies on this data for the weighting of the Community Life Survey.

### Department for Communities and Local Government

In addition to the DCLG uses which the ONS have already included in their supporting documentation, DCLG has the following uses for Census data:

- census and census-based population data by age and sex at national and sub-national level are used for the sampling frame and calibration of the English Housing Survey (EHS);
- net housing supply analysis uses the number of dwellings as reported in the Census year at the LA level and then adjusts in subsequent years for net gains and losses;
- LA level estimates of households, and population estimates and projections, are used in household projections. Detailed household reference person cross-tabulations from the census at the national and sub-national level are also used in these projections; and
- data on tenure and household characteristics are also used for other research and analysis.



## Department for Culture, Media and Sport

While DCMS does not often use census data directly, it is the source of weightings in large scale participation surveys.

Government Equalities Office is part of DCMS. Equalities requirements need to be considered, and census data are used both within GEO and in the equalities research community. ONS should comply with the Public Sector Equality Duty (PSED) by ensuring that they understand what the impact of the different decisions would be on data that could affect the equality outcomes of policy making and take that into consideration in reaching a decision on how to proceed. Small area analysis of demographic characteristics is very important for equalities analysis.

Some equalities groups and communities are widely geographically dispersed and some clustered in particular locations, and this has implications for robust sampling of the option B 1% and 4% samples.

For creative industries analysis and benchmarking, detailed census data at the 4 digit Standard Industrial Code level and 4 digit Standard Occupational Code Level are required for accurate benchmarking of “creative intensities” and small area analysis. A large single dataset can clearly be obtained from a more traditional census approach but there could also be value in obtaining more frequent updates as “creative intensities” change with business structures.

## Department for Education

The Department for Education (DfE) makes considerable use of population estimates and projections based on the census. These are important for understanding pressures on spending and the performance of the education system. This data, with our own Schools Census data, helps inform how many pupils need to be educated, the number of teachers needed to teach the pupils, and the number of schools needed.

Demographic change has specific implications for policies, their deliverability and their costs; for example, the raising of the participation age. Census data is used for calculating the 16-18 participation rate and the share of 16-18 year olds not in education, employment or training (NEET).

Census data also matters for measuring the participation rate for under 5s. The take-up level of the free early years entitlement is measured using DfE’s own Early Years Census, divided by the size of the eligible population which comes from the census.

The DfE uses ONS’s national population projections, which are based on census data, in the production of our National Pupil Projections series. Census data is also used for calculating rates of Children in Need / Children Looked After in the wider population.

## Department for Energy and Climate Change

Along with other government departments, DECC uses census data for policy development and research and analysis. Specific examples include:

- to develop, focus and target policies associated with fuel poverty. Better information on households and people helps DECC's understanding of those susceptible to fuel poverty, as well as the impact of energy efficiency measures, and contributes to wider work on improving the thermal efficiency of the housing stock as well as assessing trends in energy consumption. Data from the census supports analysis of sub-national energy consumption and contributes to the model used to derive sub-regional estimates of fuel poverty. It is also used for ongoing methodological and analytical work to improve fuel poverty modelling, and to develop new and improved indicators; and
- small area data may be used to impute household characteristics as part of work to develop a new National Household Model. The department makes extensive use of the English Housing Survey as well as the Indices of Deprivation and household projections, all of which depend directly or indirectly on the census. For example, socio-economic deprivation data which is part based on census data informs analysis to support the allocation of areas which may be targeted for elements of the Energy Company Obligation. DECC therefore need sufficient data provided at the LSOA level of geography to allow these deprivation indicators to continue to be constructed. The department has an increasing need for up to date data at local authority and small area level because new models rely more and more on small area data. Population and household size characteristics are key.

## Department for Environment, Food and Rural Affairs (DEFRA)

Defra has cross-government responsibility for championing rural issues in the development and implementation of policies. Information from the census is used extensively both to define and characterise rural communities. There are around 9.3 million people in England living in rural areas, representing 17.6 per cent of the population. Age demographics alone suggest that it is important to consider separately rural areas, as there are proportionately more older people living in rural areas than in urban areas, which has implications for service provision. The opportunities, challenges and barriers for people and businesses can be different in rural areas compared with urban areas.

The Rural-Urban Classification is used to identify rural areas and the people living there, and the census underpins the Rural-Urban Classification (RUC). The RUC categorises the OAs according to settlement form and context and taking into account population density. At OA level, RUC categorises areas to one of six rural or four urban settlement types. The RUC can be used in aggregated form to classify larger geographic areas according to rural and urban definitions.

The main uses of census data in Defra are by the Rural Communities Policy Unit as mentioned below, but census data are also used by Defra and its agencies as follows:

- understanding the rural population permitting the rural proofing of policies across government;
- understanding the bounds of deprived communities;
- through the Environment Agency, planning and securing approval for flood defences; auditing water companies' water resources and drought plans; securing water supplies to meet local population needs, and planning of waste provision; and
- through Natural England, planning and delivering local policies on the environment e.g. National Parks, Areas of Outstanding Natural Beauty. A recent example is the use of small area statistics for the Lakes-Dales National Parks Extension public enquiry.

### Department of Health

Work on benefit quantification has focussed to date on allocation of the healthcare and public health budgets. Census-type data has a number of additional important direct national uses, including for the allocation of the adult social care budget and setting GP pay. At a local level, the data are used to inform the commissioning, design and delivery of services, and for the investigation and handling of disease clusters and outbreaks.

The annual population estimates that are derived from the census form the basis for a mass of other important statistics, such as mortality rates, disease rates, immunisation, vaccination and screening rates, and conception and abortion rates. These are critical to both national policies and local operations. For example, the mortality rates included within the NHS and Public Health Outcomes Frameworks are calculated using local age and sex population estimates.

Census information is also used to adjust for bias in responses to many surveys, such as the Health Survey for England, ensuring that the results of those surveys are meaningful and able to inform policy development.

### Home Office

The Home Office's responsibilities include reducing and preventing crime, controlling immigration and keeping the UK safe from the threat of terrorism.

The census and related population statistics enable the Home Office to:

- support funding decisions, most notably the money distributed to police authorities via the Police Main Grant;
- develop strategies, as well as the design and evaluation of policies, including those on migration and integration;
- understand the needs of specific communities and promote community cohesion;
- carry out research and analysis to understand the impact of migration and assess, at a local level, the extent and speed of integration; and

- support methodological work including ‘Most Similar Groups’ analysis for the presentation and comparison of police force data between areas.

### Department for Transport

DfT and the transport sector use a range of census data including personal and household characteristics, and deprivation, urban/rural and travel to work classifications, for a wide range of purposes. These have been largely captured by ONS, but the key points are summarised here for completeness. The most distinctive uses are in transport modelling and in the appraisal of transport schemes, which is required for all transport schemes involving public funding. This work typically uses small area data, often requiring MSOA or LSOA level data, but sometimes OA level. It informs investment decisions with a major long term impact on the economic wellbeing of the UK. To put this in perspective, the whole life cost of the part of the Government major projects portfolio for which DfT is responsible is valued at over £46bn – most of this relates to major transport projects, and this is before including most of the lifetime costs of HS2, or support for local transport schemes amounting to hundreds of millions per year.

### Department for Work and Pensions (DWP)

The main uses of census data in DWP are:

- Local delivery: population statistics are used to inform DWP’s delivery organisations as to what services to provide, and how to provide them. This includes an understanding of the age and/or ethnic group of disabled people living in their local area.
- Travel to Work Analysis and Workplace Zones: The Independent Review of Methodology for the Beyond 2011 Programme by Chris Skinner raises concerns about the future validity of travel to work area analysis and workplace zones under option B. This information is used in DWP Operations and by labour market analysts to help frame conditionality policy. It is also used to plan estate strategy for the DWP, and to consider siting of offices to maximise their accessibility.
- The Longitudinal Study is part of the evidence base for reviewing State Pension Age. If option B is chosen the fate of the Longitudinal Study (LS) is uncertain. The Longitudinal Study is used to estimate healthy life expectancy according to the National Statistics Socio-economic Classification. If option B was chosen and the Longitudinal Study was discontinued this would seriously affect consideration of the appropriate State Pension Age for the UK by future reviews, to be held by the Government every 6 years. Healthy life expectancies at age 65 will be a key consideration when Parliament is debating whether and when to raise the State Pension Age (SPA). This decision is worth billions of pounds to the exchequer for a small difference either way. From a SPA point of view, a reduction in ability to understand differential health and longevity within different socio-economic groups and

regions could potentially reduce the insight of future reviews looking at SPA increases.

- **Model Development:** accurate population projections (including by marital status) are vital to DWP's long-term forecasts, especially of PENSIM2 (DWP's model of pension-age incomes). Any diminution in their accuracy would increase the risk around estimates of expenditure on pension-age benefits. DWP also values census questions on country of birth and date of arrival, as understanding the demographic characteristics of those who migrate to the UK is important to modelling of future pension entitlement. Similarly, DWP is interested in monitoring the numbers in civil partnerships (and in the future, same-sex marriages), behaviour in terms of partnership formation and dissolution, and the likelihood of having children. The concomitant effects on labour market participation and pension inheritance rights may be significantly different in modelling terms from other groups in the population. (Evidence already shows, for example, that married and cohabiting couples are different in terms of likelihood of having children in the household.)

**Q3:** *Please specify any significant additional benefits of population and housing statistics that we have not already identified.*

N/A

### 5.3 Impact of different census approaches on statistical uses

**Q4:** *What would the impact be if the most detailed statistics for very small geographic areas and small population groups were no longer available? High, medium, low, or no impact?*

Impact ratings vary across government, in part this is dependent on the importance and use of small population statistics in the formulation of policy or further analysis. Where small geographic/population statistics are not important to the formulation of policy then understandably the impact is assessed as low. However some departments do make use of small population group and micro-geography outputs in the formulation of policy, therefore the impact on these activities is high. The table below summarises the different impact ratings of government departments of the withdrawal of small area and small group statistics. The survey grossing issues highlighted by DWP below would also affect similar household surveys and their users as the census derived population estimates are the standard reference for grossing to population totals.

Low	Medium	High
DCLG	DWP – disability policy	DWP – Grossing household surveys and model development
Cabinet Office	DECC	DH – Health and Social Care
BIS	DfE	DfT
		Defra
		DCMS/Government Equalities Office (GEO)
		Home Office

While DCLG rate the impact on their own work as low they note that for local authorities and voluntary sector organisations where service delivery and risk management is focussed, there could be a wider impact on the implementation of departmental policy. Furthermore, while there would be some impact on DCLG's Indices of Deprivation these are no different to the current impact of a decennial census. There also currently exists an indicator of population sparsity, which is based on OA level population data, in the Business Rates Retention system, and therefore a lack of population data at this level would require a less specific measure to be used.

While BIS rate the impact as low, the availability of data at low levels of geography that can be aggregated into new geographies (such as Local Enterprise Partnerships) is important for local policy design.

DECC's medium impact rating depends on the availability of alternatives, such as Experian's datasets, and how the Indices of Deprivation will be constructed and calculated in the absence of detailed census data.

**Q4.1** *If medium or high, please give further information.*

Medium impact ratings have been provided where there are concerns about the availability of particular statistics or alternatives, for example the availability (and cost) of private sector geodemographic datasets. DWP highlighted a specific concern about the availability of information to support disability policy. If it was not possible to provide LA-level data by age group and disability, then there would be a medium-level impact.

<b>High</b>	
DWP – Grossing household surveys and model development	The impact could potentially be high – the key for survey weighting is accurate population counts split by a few key variables (age, sex, local geography). There is potential for this list to expand as the demand increases for assessing (and adjusting for) survey non-response, for example, ethnicity and nationality could become more important. Accurate survey weighting necessitates cross-tabulation of these variables, creating a demand for highly accurate estimates of small populations. These surveys are key to evidence based policy making.
DH- Health and social care	The lack of availability of detailed statistics at LSOA level and of all characteristics information at Output Area level could have a significant adverse impact on health and social care with particular concerns regarding: <ul style="list-style-type: none"> <li>• the allocation of health care, public health and adult social care budgets; and,</li> <li>• the planning and delivery of local health and social care services as local partners may be less able to identify need within their area.</li> </ul>
DfT	Transport is intrinsically spatial, therefore many analyses require a geographical element. Often MSOA or LSOA level data may be sufficient, but in some cases OA level data are required, for example to identify the localised impacts of a transport scheme, or the impact on a socio-demographic group that is geographically very clustered.
Defra	Defra requires Output Area estimates both to define and characterise rural areas. In particular, the loss of characteristic data would be of concern. It would be possible to use higher geographies as a ‘best fit’ for the analysis. This would allow some estimates to be formed, but it has the potential to miss out completely or to provide poor information for some rural areas, in particular households in the most sparse rural areas. Our understanding of the needs for these groups will be lacking and may impact on policy development and the

	<p>outcomes for rural households in England.</p> <p>At OA level there are six classes of rural settlements. By using LSOA level we immediately lose the finer grained information about hamlets and isolated dwellings as this is lost within the 'villages' classification. This means that for the most rural households we have no or little information about them.</p>
DCMS/Government Equalities Office (GEO)	<p>Small area analysis, for example, of demographic characteristics, is very important for equalities analysis as variation within a local authority area, for example for a particular characteristic, may be much higher than between local authorities. This holds down to the smallest Output Area level provided by census data. Small area analysis also allows these equalities characteristics to be analysed accurately against area level characteristics such as housing or deprivation. The Cathie Marsh Centre for Census and Survey Research, for example, devotes much of its programme of research to small area analysis around equalities.</p>
Home Office	<p>The movements, socio-economic features and behaviours of short-term and long-term migrants are of key interest for government because this provides information on integration and demand for services. For example, very small area mapping of the population of migrants by nationality or groups of nationalities (cross tabulated with other variables e.g. English language ability or year of arrival) can identify community needs, particularly with regard to integration. Being able to separately identify EU and non-EU migrants is important given the different rights each group has, the differential impact of immigration policies on them and the potential impacts these differences have on wider society.</p>
Department for Education	<p>In terms of measuring participation, the primary geographic level of data needed is the local authority area. What is more important is getting robust national age breakdowns of the population – ideally in one year age bands.</p> <p>More generally, detailed data on low level geographies will become increasingly important as the number of pupils and demand for pupil places continues to rise, so the DfE and other stakeholders such as local authorities can understand where the school capacity pressures are. The DfE can get a reasonably good understanding</p>



	of this using our own School Census and other data, but low level census data will help provide a more complete picture.
--	--------------------------------------------------------------------------------------------------------------------------

**Q5:** *What would the additional benefit be if more frequent (ie annual) statistics about population characteristics were available for areas like Local Authorities and Electoral Wards? High, medium, low, or no impact?*

There is an intrinsic benefit from more frequent statistics, in that they would enable policy makers to understand the changing nature of the British population closer to “real-time”, which is critical at a time of increasingly dynamic population movement and compositional changes.

For some, however, the utility depends on what detailed variables are additionally provided as part of such a dataset.

Low	Medium	High
DWP – Delivery outputs	DH – Health and Social Care	
DWP - Grossing household surveys and model development (potentially low) <sup>2</sup>	DCLG	
DfT <sup>3</sup>	DfE	
DECC		
Cabinet Office (low or no impact)		
BIS		

**Q5.1** *If medium or high, please give further information.*

Medium	
Department of Health – Health and Social Care	The availability of survey data on an annual basis would have some benefits, in providing a more up to date picture of the population, particularly at Local Authority and Clinical Commissioning Group level and, more importantly, MSOAs within those areas. This would enable the more frequent assessment and monitoring of health and care to inform local needs assessment and service delivery. This would be of even more benefit if a broader range of health-related information was collected. This information could potentially be used in the NHS, Public Health and Adult Social Care Outcomes Frameworks.
Department for Communities and Local Government	Annual statistics about population characteristics at a local authority level could result in more use of census data in the formulation of grant allocations, policy development and monitoring, service planning and

<sup>2</sup> Leaving detail aside, the benefits of more frequent statistics depend entirely on how much more accurate they are (if at all) than the current approach and data currently available for survey grossing.

<sup>3</sup> The loss of small area detail is liable to significantly outweigh the potential gain of more frequent statistics for broader area

	<p>delivery and research and analysis. This is because often alternative sample surveys, such as the Labour Force Survey, are too small to provide robust local authority level data for all authorities. This is especially the case for small or sparsely populated authorities. The availability of robust, timely and frequent data is often a key consideration in what indicators are used. Depending on what questions on housing and household are included in the sample surveys, there could be the potential to reduce or eliminate current DCLG housing related data collections, especially where there is duplication, which could reduce the cost and burdens on local authorities and DCLG.</p>
Department for Education	<p>As noted, the number of pupils and demand for pupil places is projected to rise for the foreseeable future, so more frequent statistics would enable DfE and other stakeholders better understand where the capacity pressures and forecast future need.</p>

#### **5.4 Impact of different census approaches on historical research**

ONS has worked with The National Archives and genealogists to understand how census information is used in historical research. These questions ask you to tell us about any uses or benefits of census information that we have not yet fully understood and to share your views on the potential impact of the different census approaches.

**Q6:** *Please specify any significant uses of census information for historical research that we have not already identified.*

The Government is aware of the importance of census to genealogists, and would want to ensure that ONS do everything possible to support this community under option B. ONS could perhaps further assist this group by using existing records better.

**Q7:** *What advantages or disadvantages for genealogical or historical research can you see from a move to a solution based on archiving administrative data sources?*

N/A

## 5.5 Managing risks

**Q8:** *What are your views of the risks of each census method and how they might be managed?*

Overall, the Beyond 2011 Independent Review of Methodology by Chris Skinner for the Programme makes a number of valuable points which we hope ONS will consider thoroughly before making a final recommendation.

We have set out our analysis of the risks associated with each option below.

### Risks associated with option A

Risks associated with option A are lower, at least in the short to medium term, in that the method is well established, and it is known that it can deliver the outputs currently relied upon, especially small area data. The risks are mostly those of missed opportunities in the longer term – e.g. failing to improve statistics by using evidence from administrative sources, and the increasing challenge of falling response rates.

An obvious risk is related to non-response, from those who have no secure access to the internet, those who find internet access difficult or those who find form completion difficult. This may introduce additional bias into estimates, particularly if they are similar in characteristics, for example, the elderly or people with disabilities. However, as the document states on page 9, other methods of responding will be made available. It would be useful for the ONS to research this further. Non-response could also be affected by concerns about the security of completing an on-line form.

While option A retains key methodological similarities to the existing decennial census (i.e. it remains a collection of data from all households) by moving to an online data collection mode there is a fundamental change to the methodology which could impact on the quality of outputs and estimates. ONS should actively explore what differences (if any) a switch in data collection mode will have to underlying estimates. The government accepts methodological differences as a natural by-product of changing data collection mode, however prior assessments of mode differences will prepare data users for potential impacts to time-series data. As the census collects factual data and is already a self-completion exercise these mode differences may be smaller than if the census collected attitudinal data where social desirability biases can more strongly affect estimates. However concerns over online privacy may result in stronger reluctance for households to provide information or for individuals to share information within households for the purposes of an online census.

### Risks associated with option B

Option B is a significant change for the ONS and there is a risk of not achieving the desired outcome (accurate, timely and reliable data). The risks need to be managed closely and it is essential that there is a clear back up plan if data sharing agreements are not achieved, and sampling frames are not satisfactory. There will

also be issues relating to non-response and ensuring hard-to-count populations are adequately covered in the sample surveys and this has potential to have a greater effect on the overall results than with a traditional census.

In addition, there are a number of concerns with option B – for example, the methods of obtaining population estimates are described as partially untested and access to the administrative data would require legislation. Combining data from a number of different sources may also raise privacy concerns and would require specific safeguards.

## 1. Geographical detail

Several departments make extensive use of the fine level of geographical detail provided by the census. Option B envisages providing as much as possible of this very fine level of detail by model-based small area estimation (SAE) combining administrative and survey data. There is a question of whether ONS will know how effective this SAE would be by the time a decision is made about the future of the census.

More research is needed to confirm what level of data will be available at OA level and to consider the implications of using administrative data for basic population estimates at OA level. Alternative reliable sources for characteristic data to approximate to OA level will be required by some Departments. We also suggest that consideration might be given to increasing the size of the additional survey, although there would be cost implications of doing so.

Accordingly, mitigating strategies might include:

- carrying out a cost benefit analysis looking at increasing the sample sizes for both the 1 per cent and 4 per cent admin top-up surveys to see whether this would improve estimates at OA level and other small area geographies;
- investigating the potential for producing finer disaggregations in urban areas. In these areas, high density, high capacity public transport networks may require especially detailed data. The consultation also noted that there may be particular problems with the use of administrative data in urban areas;
- giving particular attention to maintaining statistics for travel to work areas and workplace zone statistics; and
- giving particular attention to the 'Output Areas as a building block' approach outlined in the consultation document.

## 2. Benchmarking

An important function of the census is to provide an accurate benchmark for weighting and calibrating surveys that are used to produce a range of Official and National Statistics.

If judged against the criterion of accurate benchmarking, the risks around option B appear greater. Administrative sources can build up large parts of the overall picture, but accurate benchmarking is over-dependent on the sorts of marginal differences not captured by administrative sources (e.g. people not captured or double-counted). Crucially ONS would have less control over the collection and content of these sources. Therefore, accuracy would be very dependent on the accuracy of the 'surveys' aspect of option B. There are particular issues associated with this which we discuss below.

### 3. Response rates

Sampling error associated with a 4% sample survey should be very small, but the potential for non-response bias could be much higher.

Option B relies on the compulsory return of the 4% sample, however it is unclear how well this compulsion to complete it will be enforced. While the existing census and option A also rely on this method, the whole population coverage enables a strong brand and communications exercise that option B is unlikely to be able to replicate. Significant resources may need to be made available to encourage or enforce response.

The enforcing of a law requiring people to complete the census once a decade is very different to enforcing a law that requires surveys to be completed on an annual basis. This risk needs to be managed through evidence from other countries with the factoring in of cultural differences.

A lower response rate may result in bias. If option B proceeds, consideration needs to be given to how the sample is distributed between areas and for follow-up action to be targeted at those areas and groups who are otherwise less likely to respond.

There is a further risk that if the cultural expectation of responding to the 'census' is lost due to the switch a compulsory sample survey, it may be difficult to recover participation should there be a return to a 100% census.

### 4. Interaction with existing household surveys

There is a risk that other existing household surveys would be affected by a move to option B, especially if the accuracy of the estimates is reduced – as described above. There is also a potential risk that regular annual estimates of key statistics through option B reduces the need for some existing household surveys, which could result in a loss of data for questions not covered by the ONS survey.

We would welcome more details of the future of existing government surveys which currently give both high level data and detailed data for sub-groups of the population. For example, the Labour Force Survey (LFS) gives headline employment and unemployment data as well as very detailed labour market information from the wide ranging questions in the survey. With an administrative source combined with a large scale survey we could see competing measures of unemployment thereby potentially undermining trust in one or both sources. However the large scale survey could not deliver all of the detail offered by the LFS itself. There are likely to be similar scenarios where new data sources offer an 'implied' net migration estimate (by a

residual method), but this would not offer the detail currently delivered by the International Passenger Survey (e.g. immigration and emigration by nationality and by reason) and could potentially represent a step back from the international standard which we currently meet.

An area which would benefit from more research would be the inter-linkages between data sets where it might not be apparent that one of these relies upon census data. If that source changed then it could have a knock on effect on other surveys or data sources which do not in themselves directly rely upon census.

Relevant stakeholders in existing government surveys need to be involved in any discussions if option B is taken forward so that implications across government can be considered and understood before changes are made. If ONS aims to move other household surveys like LFS towards substantial internet completion in the next few years, then this could inform the census approach.

## 5. Discontinuities

Changes to the systems, processes and incentives for the data sources may create discontinuities. Attempts to manage or deal with this through change controls and coverage surveys may not wholly rectify the problem. In assessing option B, it is important to factor in the cost of such restrictions on the delivery of public services.

Option B also presents some more general risks by virtue of its novelty. In some respects the method is as yet unproven, and there is a risk that if it does not work as well as hoped, the country may be left with data gaps, or with estimates that are no longer as widely trusted or accepted as those from the census have been in the past.

## 6. Implications for departmental administrative data collection systems

Departments collecting administrative data which will be used for any replacement to the census will need to be aware of the impact of any change in their data collection systems.

As discussed above it will be also be necessary to fully explore the costs to departments of the use of administrative data.

### Combining option A and B – ‘dual running’

Many departments felt that the risks of option B might be better identified and mitigated by running a full online census in 2021 and double running with administrative sources. The ‘transition period’ outlined on page 13 would be a sensible way of managing the risks of this and help ensure that an approach that is sustainable over the long term is implemented. Both approaches would differ from the enumerator based approach used in 2011 and ONS should consider any potential discontinuities.

Dual running offers a good opportunity to explore both methods fully, however as mentioned in the consultation document the respondent burden will be increased which could increase non response. In addition, it may be that limited lessons can be



drawn from the option B approach if it takes place in the wider context of full census as noted above in relation to enforcing a law for completion of these surveys. In the case of the survey envisaged by option B, to minimise the increase to respondent burden the survey could be implemented in a staged way, for example by building up the frequency or size of the survey over time.

Moreover, it would be advantageous to consider how administrative data can be used to provide more timely national and local estimates of official statistics irrespective of the Beyond 2011 programme – for example using administrative data to produce regular LSOA population estimates.

**Q9:** Are there any other issues that you believe we should be taking into account?

It is critical when evaluating options that the full benefits and full costs are taken into account. This includes the full costs to the UK of the opportunities that would be lost (for example through sub-optimal planning of major infrastructure projects) as a result of not having the data that have been available from previous censuses.

It is important that there is consistency between population censuses in different parts of the UK. In Scotland the population census is a devolved matter and the responsibility of the National Records of Scotland. From a UK perspective there are two main reasons for consistency. Firstly many UK departments deliver services in Scotland and need good quality population statistics to enable them to deliver those services. Secondly the Coalition Government has rightly devoted resources to publishing large public datasets to driving the knowledge economy and in that context there are clear benefits from consistent cross UK data. Full and proper liaison with the National Records of Scotland is therefore important.

ONS should comply with the Public Sector Equality Duty (PSED) by ensuring that they understand what the impact of the different decisions would be on data that could affect the equality outcomes of policy making.