Integrated Population and Characteristics Survey (IPACS)

# 1.0 Introduction

The Office for National Statistics (ONS) is working towards the transformation of Population, Migration and Social Statistics. The Census & Data Collection Transformation Programme (CDCTP) seeks to rebalance ONS’s data collection activity significantly toward wider, more integrated use of administrative and other non-survey data sources, thereby reducing our reliance on large population and business surveys. While this will not eliminate a need for surveys, it does mean ONS’s traditional approach to surveys is likely to change. To support this agenda, ONS is also working towards the development of an Integrated Population and Characteristics Survey (IPACS) which;

1. will support the production of Population, Migration and Social Data from administrative data and other non-survey data sources and;
2. serve as the future sourcing model for any residual population, migration and social data.

Three underlying principles support the development of the IPACS;

* **Administrative data first**: an admin data first approach will allow us to make greater use of existing non-survey sources for statistical outputs, to increase the efficiency of sampling and data collection processes, and to replace and/or enhance survey variables
* **Digital by default**: residual survey data collection will predominantly use online methods, supported by face-to-face and telephone interviews
* **Statistical redesign and rationalisation**: wherever possible we will bring residual survey requirements together so that we minimise the need to ask for the same data across multiple surveys

In this paper we describe the principles behind the IPACS, testing undertaken to date and our future test plans between now and 2022. Key methodological considerations underpinning an IPACS will be highlighted also.

# 2.0 Background

## 2.1 Population & Migration Statistics Transformation

Population & Migration Statistics Transformation (PMST) has developed a Statistical Population Dataset (SPD) to support the development of administrative data population estimates. This is constructed by linking 4 sources together that, when combined, are expected to cover a large proportion of the usual resident population (including some overcoverage). These are:

* The DWP Customer Information System (CIS) – a list of all NINo registrations
* The NHS Patient Register (PR) – a snapshot of current GP registrations
* Higher Education Statistics Agency (HESA) – students registered in the year for higher education courses
* The English and Welsh School Census (SC) – pupils registered in state schools

These datasets are characterised by both undercoverage and overcoverage of population and an adjustment is required to account for them in population estimates. Methods to account for under-coverage are already well established in traditional censuses, for example the Dual System Estimation (DSE) approach that was used for the 2001 and 2011 Census of England and Wales.  By linking multiple administrative sources to construct the SPD the level of population under-coverage is reduced. However, we still anticipate that there is some residual undercoverage and depending on the availability of additional data sources that can account for persons missing on the SPD, it is likely that a similar DSE adjustment may be needed to account for undercoverage.

Measuring and adjusting for over-coverage on the SPD represents a more significant challenge, particularly within a DSE framework. Amongst other assumptions, DSE requires that datasets used in the estimation framework are free from overcoverage. However, attempts to detect and remove over-coverage from the SPD have not been successful to date. Methods for measuring over-coverage are being pursued across a number of countries, and ONS are working with colleagues across NSIs to better understand the properties of estimators that incorporate over-coverage adjustment. As a consequence, future SPDs may be redesigned to best meet the properties of these estimators.

The concept of ‘dependent interviewing’, which involves sampling households and directly checking whether persons listed on administrative records are resident at the address, has been explored as an option for ONS, but ruled out due to sensitivities regarding the disclosure of information about previous residents. It is now proposed that an ongoing Population Coverage Survey (PCS) for collection on a continuous basis be developed to measure and adjust for the both under and over-coverage biases.

## 2.2 Social Survey Transformation

Within the context of the Census and Data Collection Transformation Programme, Social Survey Transformation (SST) is responsible for transforming the statistical design of the current household survey portfolio, from one that has little to no integration across survey designs into an integrated design.

Our vision is to deliver an integrated approach which utilises non-survey data as the primary source for household data outputs. This is dependent upon such data being assessed as meeting quality requirements for the outputs in question. Non-survey data shall also be used as a means of designing better samples for those residual surveys that will be required to complement these data sources.

Surveys will be ‘digital by default’ where possible but will utilise an online only mode for an initial period, followed by concurrent mixed mode methods (online, face-to-face and telephone). This vision and the plans that underpin it align with the Authority’s strategic objectives of increasing on-line collection, survey rationalisation and greater integration of non-survey data sources into statistical outputs.

SST is responsible for the transformation of the Labour Force Survey (LFS), the Household Financial Surveys (encompassing the Living Cost & Food Survey (LCF), Survey on Living Conditions (SLC), Wealth and Assets Survey (WAS) and the Opinions and Lifestyle Survey (OPN). Work to date has focussed on transformation of the Labour Force Survey; a prototype version of a questionnaire including core labour market variables has been developed – this is referred to as the Labour Market Survey (LMS). Work is also ongoing for the transformation of the Household Financial Surveys and the non-labour market elements of the LFS.

2.3 Integrated Population and Characteristics Survey (IPACS)

PMST and SST are working together to meet future requirements for the production of Population, Migration and Social statistical systems, via the development of an IPACS.

A key feature of the IPACS is a wave 1 questionnaire with a large sample of 500,000 households, which driven by the required level of precision for the PCS. This is referred to as the ‘Master Wave’, and is designed to collect socio-demographic data, core labour market data, and to provide an assessment of the coverage of administrative data that will be used to derive population estimates.

The ‘Master Wave’ would be collected via a concurrent mixed mode questionnaire, with an online first approach. Non-responding households would be invited to take part in a face-to-face interview. The responding households to the wave 1 ‘Master Wave’ would then act as the sampling frame for subsequent waves or other social surveys which containing topics to meet residual social survey output requirements.

*Figure 1* illustrates the key components of the IPACS. It would be designed with an administrative data first approach in mind, with the idea being to utilise administrative data sources wherever possible to satisfy population, migration and social data requirements. Surveys would be used to capture any residual requirements, including population undercoverage requirements.

Our current thinking for the collection of residual survey requirements is described well by the work of Karlberg, Reis, Calizzani, and Gras (2015). They propose the creation of a series of ‘modules’, 10 to 15 questions that ‘logically belong together’ (i.e. address various aspects of a certain topic). Modules would then be grouped into ‘instruments’, where an instrument in defined as a sequence of modules. Each instrument would allow the joint observation of several modules from the same statistical unit and thus the crossing of variables from different modules. Modules could be present in one or more instruments, thereby providing cross-survey consistency. This approach is also being considered by Statistics New Zealand for their own integrated household survey model.

*Figure 2* provides an illustration of how residual survey requirements might be met beyond the ‘Master Wave’ using the instrument and module design, and provides an example of how cross-instrument analysis could be performed. These subsequent waves / surveys would again be online first, but would use telephone collection in addition to face-to-face interviewing.

The proposed design would utilise AddressBase+ as the sampling frame; the current sampling frame for Social Surveys, the Postal Address File (PAF) has an eligibility rate circa 91%, compared to a 95% eligibility rate for the AddressBase product. Administrative data would be used to supplement the AddressBase sample frame to identify more ineligible addresses which will increase the efficiency of data collection operations.

*Figure 1: Proposed Design of the Integrated Population and Characteristics Survey (IPACS)*



*Figure 2. Illustrative example of Wave 2 survey structure*



2.4 Key **benefits** of integration

There are key benefits to moving from a non-integrated to an IPACS:

* A more **cost efficient** system; An integrated system for the production of population, migration and social statistics will significantly reduce data collection operations when compared to the costs for the decennial Census, and the regular collection of Social Surveys. Despite the large size of the ‘Master Wave’, the greater use of non-survey data, online first data collection, and survey rationalisation will result in the overall cost of the data collection operations being significantly lower
* The **simplifying and standardising** of data collection and processing systems across the integrated system will also reduce costs and will provide more consistency in production and analysis.
* As a result of merging the PCS and LMS, the large sample size required for population statistics will have the likely benefit of providing core labour market outcomes at a **more precise and lower geographic breakdown** than is currently available from the existing LFS.
* A more **efficient** and **flexible** system: Karlberg et al. (2015) describe how non-integrated survey systems, such as our current system, lead to rather large ‘overloaded’ questionnaires and oversampling of certain items that are piggy-backed onto a larger survey without sub-sampling. They describe the problem of ‘excessive sampling’ (i.e. the collection of data on a variable from a sample that is bigger than it needs to be to obtain results with the desired level of precision). They state that by considering surveys simultaneously, better decisions can be taken on where to add new variables, or whether they can be added at all, on the basis of the overall priorities of the statistical system. Although the IPACS would involve sampling 500,000 households at Wave 1, by satisfying the principles described by Karlberg et al. (2015), coupled with an administrative data census, the IPACS will satisfy core population, migration and labour market requirements simultaneously. Further, it is intended that the Wave 1 questionnaire is limited in length, containing only ‘key’ core questions. Respondents would then be routed to subsequent wave 2 modules. In this way respondents would not be subjected to lengthy questionnaires or asked unnecessary questions, though some sample redundancy would remain.
* A **reduction in respondent burden** will result from rationalising individual household and population surveys into the IPACS. Rationalisation of questionnaires and the greater use of non-survey data sources for question replacement will result in shorter, more focussed residual surveys which take significantly less time for households to complete.
* An IPACS has the potential to provide higher levels of **flexibility** than the current system and a **more responsive** tool to meet policy needs – increasing value to users. The only way to satisfy additional survey requirements in our current system is to add to existing surveys, which can lead to large ‘overloaded’ questionnaires and sometimes oversampling of certain variables or to set up a new survey, a resource intensive exercise. Considering all survey data collection requirements simultaneously should allow for better and more flexible decision making on the content of the surveys.
* **Improved opportunities for cross-analysis** - a limitation of our current system of non-integrated surveys is that it is difficult to provide cross survey statistics. This problem is overcome to a certain extent by excessive sampling (i.e. including key analytical variables on all surveys). However, this does not allow for cross-analysis in any depth. Designed in the correct way, an IPACS would allow for the production of statistics that cross statistical domains (Karlberg et al., 2015). It is acknowledged that the benefits of this come with some significant challenges, namely the potential complexity of the analytical processes that would be required to produce such outputs.

2.5 Key **challenges** of integration

* **Harmonisation of sample and survey designs**; for example, PCS uses an unequal probability stratified, clustered sample compared to LFS which uses an equal probability random sample. Work is underway to determine the most efficient and effective sampling technique for the IPACS.
* **Harmonisation of questions and definitions** between surveys; for example, LFS currently captures information on earnings in a different way to HFS. This is an opportunity to provide more harmonised stardards across outputs where possible, but will present challenges when considering breaks in time series or different legislative requirements (e.g. EuroStat outputs).
* **Understanding user requirements** – there are a vast array of users of the LFS and APS who use the data in a multitude of ways; an exercise is underway to gather core user requirements from Government departments and to categorise these into common themes.
* **Reference weeks** – to aid respondent recall for the online mode, a ‘rolling’ reference week has been used for core labour market questions. The impact of this, compared to the fixed reference week approach currently used by LFS, is currently being explored.
* **Interviewer challenges** – the job of a field interviewer may become more challenging as an online first approach may be capturing those households who are more amenable to participating in a voluntary survey; interviewers may therefore encounter the ‘harder to reach’ households
* Modernisation and integration of surveys could result in contextual or mode effects; the thorough approach to survey rationalisation, redesign and testing is closely considering these potential issues. This will allow us to fully understand and minimise any such impact.
* **Attrition** could introduce additional **bias** in social survey estimates. A mulit-wave attrition test is currently underway which will allow us to explore this issue, with further testing planned for 2020.
* **Non-response.** The Census Coverage Surveys in 2001 and 2011 achieved response rates higher than 90%. This is largely due to the collection taking place 6-8 weeks after Census day and the perception that it is mandatory to participate. From a PCS perspective, it will be much more apparent to respondents that participation is voluntary, and response rates are expected to be significantly lower than 90%. Depending on the relationship between survey non-response and admin data coverage, this may result in biased population estimates. Work is continuing to assess required response rates and how these could be addressed.

3.0 Integrated Survey Testing

An agile and iterative approach to testing has taken place over the past 2 years, with each test designed to provide information and understanding of operational and statistical processes, and to build upon the outcomes of each previous test. Cumulatively the tests provide assurance towards the development of the IPACS from both quantitative and qualitative perspectives.

### 2017

In February 2017, ONS commissioned Ipsos MORI to conduct a series of response rate tests to establish the optimum design for a future LMS. The experiments focussed on a prototype LMS questionnaire – composed of a core set of LFS questions transformed by the Research & Design Team within SST. However, these experiments also offered an opportunity to explore response rates to an integrated survey. An integrated survey was also created composing both Labour Market and Population Coverage questions.

The first test involved administering an **online only** LMS to 37,000 households across Great Britain. It explored the effects of alternative combinations of survey materials (i.e. invites, pre-notification letters) and conditions (i.e. day of dispatch, length of time between invites) on uptake rate[[1]](#footnote-2) to the survey. The most effective communications strategy was found to be an invite letter followed by two reminders, issued one week apart. This approach was difficult to put into operation, so the second most effective strategy using one reminder was chosen.

The second test investigated the effects of different incentive strategies on uptake rate. Again, this test was **online only** and was administered to 40,000 households. When both cost and uptake were taken into consideration, the most effective incentive was found to be the non-monetary incentive of a tote bag, achieving **27.8%** uptake rate.

The second test also explored response rates to both the integrated survey and individual LMS and PCS. Figure 2 illustrates the three comparative response rate experiments that were undertaken as part of this test.

*Figure 2: Concepts underpinning the ONS Integrated Population and Characteristics Survey (IPACS)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tests | | Online | Follow Up | | | |
| Telephone Capture | Face to Face | Paper | Survey Enquiry Line |
| Test Two | Tranche One | 15 Mins - LMS | **×** | **×** | **×** | ✔ |
| Tranche Two | 5 Mins - PCS | ✔ | ✔ | ✔ | ✔ |
| Tranche Three | 15 Mins - LMS  5 Mins PCS PCS  + | ✔(PCS Questions Only) | ✔ | ✔(PCS Questions Only) | ✔ |
| Run by Ipsos Mori | | | Run by ONS | | | |

*Table 3* shows the PCS alone achieved the highest take up rate (67.6%). This may be due to the shorter questionnaire length. However, a longer fieldwork period and different agencies collecting the data may also have contributed. It also shows that integrating the LMS and PCS reduces the take up rate by around 10%. The integrated survey achieved a lower partial rate compared to the PCS alone, a partial being defined as a household who did not fully complete the survey. The integrated questionnaire used in these experiments was not put through the rigorous research and cognitive testing process that the LMS alone underwent. With proper research and testing it is thought this gap can be reduced. Detailed cognitive testing of PCS questions is currently progressing and recommendations to improve questions will follow. The results of the 2017 test were taken as positive support to continue with the development of an IPACS.

*Table 3: Take up rates for the LMS, PCS and Integrated Survey*

|  |  |
| --- | --- |
| **Tranche** | **Take up rate (%)** |
| 1 – LMS (online only) | 22.5 |
| 2 – PCS | 67.6 |
| 3 – LMS and PCS | 56.8 |

### 2018

In October 2018 ONS commissioned Ipsos MORI to conduct a further test of an integrated labour market and population coverage survey. This was a mixed mode (online and face-to-face) test administered to a sample of 14,149 households across Great Britain. The aim was to investigate the mixed mode uptake rate, the responding sample composition and any response bias that may arise. Data collection is complete, and analysis of the data will commence imminently. Comparisons will be made between key employment variables collected on both the LMS and LFS at a national level. This will provide evidence towards the quality of the data collection from the integrated survey approach. In addition to these aims, the analysis of the test will provide evidence on whether the survey can provide sufficient information to support an admin data population estimation system. The aim is to use the data collected to test methods for estimating population at the national level for five-year age-groups. In addition to the core sample, 3,000 further households have been added to the 14149, in addresses where the admin data suggests that an international migrant is resident. Flag 4 registrations (an indicator of migration to UK) on the NHS Personal Demographic Service have been used to select addresses where migrant families have registered during 2018. . This is part of early testing to determine whether a boost to increase the coverage of international migrants in the sample can help support future international migration statistics.

### 2019

Tests to date have demonstrated that, in principle, we can integrate surveys to satisfy both social and population data collection requirements without having detrimental effects on the response rate of either. However, this is just the start of a longer-term research and testing strategy for the development of an IPACS. In particular, research and testing are required to explore the quality of data we obtain from an integrated survey, the optimal design for an integrated survey and how best to build on the integrated survey prototype to the scale that will satisfy wider residual social and population survey requirements.

A further test, launched in April 2019, is exploring online uptake and attrition rates across 3 waves of an integrated labour market and population coverage survey. It will test the efficacy of different between-wave engagement strategies, incentivisation at Waves 2 and 3, and the effect these have on attrition rates across waves. This test will be completed by December 2019.

A larger scale test is also planned for 2020. This will be a multi-mode, multi wave test on a larger scale which builds upon the statistical and operational learning from all previous tests. The current proposal is for this test to iteratively build towards a large scale parallel run, which will then become the IPACS and result in the decommissioning of the existing LFS.

# 4.0 Key methodological considerations underpinning an Integrated Population and Characteristics Survey (IPACS)

4.1 Sample Design

In preparation for the 2020 test, it is critical to establish a sample design that can meet both labour market and population coverage requirements. ONS has published research outputs on estimating population size using administrative data, highlighting that errors in coverage vary across geographic areas and age-sex groups. It is assumed therefore that a stratified design, similar to hard-to-count concepts used in the Census, will be required to optimise the allocation of sample to estimate population size.

While stratification may improve the accuracy of LMS outputs if the stratifier is correlated with labour market variables, we have not yet defined the overall targets on which to optimise a multi-purpose survey. For the 2020 test, we have made the decision to design a sample for population estimation and are in the processing of researching what the design effects might be on labour market statistics, which are currently produced from a random systematic sample.

ONS have started working with University of Southampton on a proposed sample design for the 2020 test, with the aim of producing an estimate of population size at national level, with precision +/- 1%.

This is considerably less precise than the anticipated requirement from 2022 onwards, however limiting the test objective to a national level precision reduces the sample size requirements in early testing. Over time, samples will be designed to produce estimates for more detailed geographies with specified precision, as we increase operational capacity to conduct the survey on a bigger scale.

A key difference in designing a coverage survey for admin data, is that stratification needs to take into account variation in both under-coverage and over-coverage. Non-registration (for example with GP’s or applying for National Insurance Numbers) and failures to de-register when leaving the UK result in both types of coverage errors occurring on the SPD. By linking administrative data to the 2011 Census, we have explored variables that potentially predict the prevalence of under-coverage and over-coverage at small area level (output areas) in the SPD.

For **over-coverage** we have derived a variable based on the concept of address churn. This has been calculated by comparing multiple years (2014-2016) across SPDs and counting the people that move in and out of each address. We then sum up the movers and leavers for each address and divide by the number of bedrooms and number of years included in the comparison. The ‘mean-churn’ indicator used as the basis of over-coverage stratification is calculated as the mean of all address churn measures in each OA.

For **under-coverage** we have derived a composite indicator using principal components analysis. This has been derived from three predictors; the proportion of non-white British pupils in each OA (from School Census data), proportion of Flag 4’s (migrant indicator flags) in each OA (from NHS Patient Register), and proportion of males and females in the OA that have some evidence of interacting with administrative data, but have not been included on the SPD.

Based on first iterations, the variation explained by these models using R2 values is 0.44 for the over-coverage measure, and 0.48 for under-coverage. This is similar to the reported level of variation (0.47) explained by the hard-to-count index used to model non-response to the 2011 Census (ONS, 2011).

At this stage, our research suggests that over-coverage and under-coverage are not highly correlated, therefore separate stratification indices are needed within the design. The Dalenius -Hodges method has been used to determine the optimum number of strata for both the over-coverage and under-coverage measures. Initial analysis suggests that six strata are optimal for both measures, and these strata are proposed to be used in the designed of the 2020 test.

Simulations are being developed to estimate the required sample sizes for each stratum. We have largely focused on over-coverage to date. The sample uses optimum allocation, which results in unequal slection probabilities across strata. A two-stage stratified cluster design is proposed, with OAs selected from each stratum at random as the primary sampling unit. Within the selected OA’s, a fixed number of households will be randomly selected to take part in the survey.

Stratification of the sample will have implications for the LMS:

* Having different sampling fractions across the strata leads to unequal design weights, reducing the effective size of the LMS sample compared with the current design of LFS where all sampled addresses have the same design weight.
* Selecting the primary sampling units (OAs for PCS) using simple random sampling and then selecting a fixed number of cases in each OA leads to unequal weights within a stratum. This reduces the effective size of the LMS sample.
* Selecting OAs in the first stage of sampling, instead of sampling addresses as LFS currently does, will result in a reduction of the effective size of the LMS sample because of the effect of clustering.
* The clustering effect increases linearly with the number of cases selected in each OA
* The clustering effect varies between variables: variables with a higher homogeneity within an OA (eg ethnicity), have a higher clustering effect.
* The homogeneity coefficient of a variable can be estimated from data

4.2 Simulations

We assessed the combined effect of stratification and the sample design that is being considered for the 2020 PCS test by computing the design effect that results from weight variation.

The list of OAs is partitioned into 6 over-coverage strata and the sample of OAs is allocated optimally for PCS use across the strata; the sampling fractions relative to that in the first stratum are given in *Table 4*. We can see that strata 2 and 3 are nearly equally sampled whereas stratum 6 is heavily sampled - its sampling fraction of OAs is six times higher than that in stratum 1.

Six different designs, varying in terms of the total number of selected OAs and the number of achieved cases per OA, were considered. The design effect values from weight variation are given in *Table 5*. The design effects of the different designs are quite similar, varying only between 1.27 and 1.30.

*Table 4. Relative Sampling fractions across strata*

|  |  |  |
| --- | --- | --- |
| Stratum | Total No. of OAs | Relative Sample Proportion |
| 1 | 30854 | 1 |
| 2 | 54014 | 1.08 |
| 3 | 38261 | 1.09 |
| 4 | 31595 | 1.61 |
| 5 | 19622 | 2.37 |
| 6 | 6923 | 6 |

*Table 5. Design effect from weight variation in proposed PCS test samples*

|  |  |  |  |
| --- | --- | --- | --- |
| Design | Sample Size | Precision | Design,effect (LMS) |
| (Cases, Clusters) |
| (3, 1700) | 5,100 | 0.99% | 1.28 |
| (6, 950) | 5,700 | 0.96% | 1.30 |
| (9, 650) | 5,850 | 0.97% | 1.29 |
| (12, 500) | 6,000 | 1.00% | 1.27 |
| (15, 429) | 6,435 | 0.96% | 1.27 |
| (20, 350) | 7,000 | 0.96% | 1.27 |
| (25, 312) | 7,800 | 0.98% | 1.28 |
| (28, 273) | 7,644 | 1.00% | 1.30 |

Potential ways of reducing the contribution of differential weighting to the design effect include selecting the OAs with a probability proportional to their size (PPS) and then selecting a fixed number of cases in each OA; or selecting the OAs using simple random sampling, and within each OA selecting a number of cases that is proportional to the size of the OA. Achieving equal weights in the design (15,429) will see its design effect decrease from 1.27 to 1.20. We plan to sample OAs PPS within each stratum in the 2020 test as it is more convenient for fielwork planning.

The effect of clustering is expected to vary between variables; it should be high for ethnicity and low for health. We will use data from Census 2011 to compute the intra-correlation coefficients, which will allow us to compute the clustering effect, for all census variables that are equivalent, or proxy, to varables collected in social surveys.

Obtaining the combined design effect, from weight variation and clustering, will allow us to choose which design has the highest effective sample size for LMS. The designs (3,1700) and (28,273) should have the lowest and highest combined design effects, respectively as the effect of clustering increases with the sample size in an OA. Given that the different designs vary little in terms of precision for population estimates, the design with the highest effective sample size, which is defined as the ratio of the total sample size to the combined design effect, will be chosen for the 2020 test, provided it is cost-effective for fieldwork.

## 4.3 Length of field period

Introducing an online mode to surveys requires a longer fieldwork period to allow for the transition between modes. Currently, the LFS has a fieldwork period of 2 weeks. In the most recent mixed mode integrated test the field period was 8 weeks (2 weeks online only, followed by 6 weeks Face-to-Face and Online). There is a concern that lengthening the field period will result in a clustering effect in the distribution of data collection which will have a negative effect on the quality of the data. Data from the recent test will offer an opportunity to explore the impact of lengthening the reference period and provide further evidence to determine the optimum field length, both for each mode and overall.

## 4.4 Reference period

Current social surveys use various reference periods within their design, for example, the current LFS design uses a fixed reference week, whereas the Crime Survey uses a reference period of the last 12 months.

Introducing an online mode to the integrated LMS and PCS requires a longer fieldwork period, meaning a fixed reference week would be cognitively challenging in terms of re-call for those completing the survey at the end of the fieldwork period. We have therefore taken the decision to use a rolling reference week, based on the week prior to the date the household commences data collection.

Data from the recent IPACS test will offer an opportunity to explore the impact of this change. For example, is there an equal distribution of reference weeks, are there seasonal variations and are there variations between modes?

In addition to this, we also need to consider if an IPACS would require all subsurveys to have a harmonised reference period, and if/how this would be possible? What implications would different reference periods have for cross-analysis?

## 4.5 Calibration under an Integrated Population and Characteristics Survey (IPACS)

Calibration is the method that adjusts the weights assigned to survey sample members in order to satisfy (or approximately satisfy) some pre-determined constraints. The most commonly used constraints involve (assumed) known **fixed** population totals. In the LFS, and all other ONS surveys, the response dataset is calibrated to mid-year population estimates (MYEs) by age, sex and geography (down to Local Authority level). MYEs are based on Census data which are updated using mostly administrative data and are hence independent of ONS household surveys. The key idea is that estimates formed from the weighted sample should replicate the **known** values from other sources.

In the proposed IPACS, the survey is used to estimate the population totals, so they are **estimated** (not fixed) and not independent of the survey. We have sought initial methodological assurance on our proposals and have received advice that our approach is possible provided some key assumptions are satisfied. The critical point is that the estimated population totals, derived from the population estimation framework and used in calibration, have to be nearly **unbiased (we aim to achieve levels of bias that do not exceed 0.5%, which is a census quality aim)**. While the methodology is unspecified it is hard to comment on how far this can be achieved. However, it is known that bias is particularly problematic when there is nonresponse. Work is planned to explore the correlation in non-response between administrative data sources and surveys. If there is no correlation between the two sources, then lower response rates will be less of an issue. However, if the two are correlated a high response rate in the master sample is essential, as it is the case with current Census Coverage Survey. This will be problematic for a **voluntary** survey. Testing to date has achieved a response rate around 60% suggesting a need to consider mandating the survey.

## 4.6 Multivariate analysis requirements

As user engagement progresses and we begin to understand more about the residual survey requirements, it will be important to work out the required sample size for the “master wave” that will satisfy sub-sampling and longitudinal requirements, with unbiased estimates and small variance. It will also be necessary to consider the ‘construction’ of our IPACS. For example, whether we will follow the approach described by Karlberg (2015) (i.e. the ‘instrument’ and ‘module’ approach) or an alternative.

## 4.7 Communal Establishments

Sampling the Communal Establishment (CE) population is an essential requirement for Population & Migration Statistics. Information on CEs is currently collected via the Census every 10 years through the building of its own CE address register, a costly and time-consuming exercise. The CE population is not currently captured in ONS Social Surveys. In 2008 ONS conducted a pilot Communal Establishment Survey. However, the pilot found considerable difficulties in identifying CEs as there was lack of an adequate sampling frame at the time. Furthermore, the practical difficulties of accessing respondents within the CEs resulted in an unacceptable level of non-response. However, it would be worth re-visiting this approach based on AddressBase and alternative survey modes.

## 4.8 Estimation Framework

We are yet to determine the estimation framework for producing admin-based population estimates. A review of methods is in progress, largely focusing on variants of DSE. While measures of variance can be obtained from the survey, we still need to consider methods for bias assessment. We also need to consider the impact of linkage errors, both in constructing the SPD and linking the SPD to the integrated survey. Similar to the requirement for very high-quality Census and CCS matching, our assumption is that linkage between the SPD and the survey will need to be undertaken with very few false positives or false negatives minimum errors. This may be more difficult to achieve with lags on administrative records regarding people’s address information.

## 4.9 Summary

The IPACS is a central part of plans to transform social surveys at ONS. While there is considerable focus on modernising the collection of social surveys to support the ONS ‘admin data first’ approach, the decision to move towards a multi-purpose survey that combines LFS topics with new requirements for coverage assessment poses significant methodological challenges. A robust test plan is needed to prove the viability of the survey and provide the necessary evidence to support the recommendation for future Censuses in England and Wales which will be made in 2023.

We are currently aiming to commence testing for a parallel run in 2020, with the IPACS running concurrently with the existing LFS. The IPACS test will need to start relatively small in 2020, with the aim of collecting sufficient data to produce research outputs on admin based population estimates at the national level. Between 2020 and 2023 the IPACS will gradually increase up to full scale, with 500,000 households sampled per annum. In addition to the methodological challenges described in this paper, the operational delivery of a test of this magnitude also presents a challenge. We are exploring options for delivering the operational systems needed to support collection as well as the strategy for increasing the level of field force capacity needed to run two large surveys concurrently over a three year period.

# References

Karlberg, M., Reis, F., Calizzani, C., & Gras, F. (2015). *“A toolbox for a modular design and pooled analysis of sample survey programmes*”. Statistical Journal of the IAOS, 31, 447-462

Hall, Gibbs, Carter, Das, You, Minshall (2017) “*Options for integrating Statistics New Zealand’s household surveys with a large-scale attribute survey*”

Predicting patterns of household non-response in the 2011census, ONS (2011), <http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/processing-the-information/statistical-methodology/predicting-patterns-of-household-non-response-in-the-2011-census.pdf>

1. Uptake rate is defined as the proportion of households who enter any amount of data into the online collection instrument (all partial household completions and full household completions). This is different to response rate which includes The uptake rate calculation excludes ineligible households. [↑](#footnote-ref-2)