Barry Leventhal

is the founder of BarryAnalytics
(www.barryanalytics.com), an
advanced analytics consultancy.
It specializes in helping
companies to generate greater
business value from their data,
either by deeper analytics or
by improving their analytic
processes and/or environments.
BarryAnalytics provides access to
a wide range of experts with the
most comprehensive breadth of
experience in the field of advanced
analytics.

Keywords: census, geodemographics,
market research, direct marketing

200-year-old information source

2011 — the final census?

Barry Leventhal
BarryAnalytics Ltd,
9 Markham Close,
Borehamwood,
Hertfordshire WD6 4PQ, UK.
Tel: +44 (0)7803 231870
E-mail: barry@barryanalytics.com

Census-taking in the United Kingdom: 2011 and beyond

Barry Leventhal

Received (in revised form): 14th December 2012

Abstract

Census-taking in the United Kingdom has never been through a more crucial period — in terms of delivering results from one census, while also researching into options for the future. This paper looks at both activities, starting with a review of the 2011 Census and its key features. Meanwhile, the Beyond 2011 Programme is exploring alternative, more cost-effective options for census-taking — the paper examines its initial options and the latest solutions being examined. Finally, the paper discusses the implications of Beyond 2011 for census users in marketing. This leads to the conclusion that it is in users' own interests to advise the Beyond 2011 team of the benefits that have been derived from geodemographics in order to inform the business case that will need to be made for producing small-area statistics in the future.

Journal of Direct, Data and Digital Marketing Practice (2013) **14**, 201–213.
doi:10.1057/dddmp.2013.4

Introduction

In these times of online digital data, it may seem old fashioned and outdated to extol the virtues of an information source that has been running for over 200 years — the population census. Yet, the census delivers a unique view — a snapshot of the size and demographics of the entire population, residing in every part of the country — and provides a definitive benchmark for planning and evaluating marketing activities.

The census has underpinned research and marketing in the United Kingdom for more than 30 years. Small-area statistics from the census are the main ingredient of geodemographic classification systems such as ACORN and MOSAIC, pioneered in the UK in the mid-70s, using 1971 Census results, and launched into the research industry in 1979.¹

At the time of writing, initial results have been released from the 2011 Census and more detailed outputs are about to be published. However, 2011 may have marked the final traditional census in the UK, as the Office for National Statistics (ONS) has announced that it intends to explore alternative options in order to replace the full population survey by a more cost-effective approach.

This paper has three main objectives. The first is to summarize the key features of the 2011 Census from a user's perspective — if 2011 proves to be the United Kingdom's last full census, then its results are

likely to be referenced for many years to come and we should make the most of them. The second purpose is to review the current status of the Beyond 2011 Programme, which is exploring alternative approaches to census-taking in the future. This project is currently roughly halfway through its research phase, and is exploring alternative possible options. The third and final purpose is to discuss the implications of those options for census users in marketing.

Census strengths and weaknesses

Asking users for their views

The main strengths and weaknesses of the census in the United Kingdom are summarized in Table 1. As part of the Beyond 2011 Programme, users have been asked whether they would be willing to forego some of the strengths, such as small-area geography and multivariate data, and instead obtain more frequently updated statistics — users' reactions are discussed later in this paper.

Key features of the 2011 Census

As with any other survey, a census may be divided into three broad phases — data collection, processing and outputs. This section summarizes the key features of the 2011 Census under those three headings.

Data collection

Income question excluded

In the run-up to 2011, the UK Census Offices — ONS for England and Wales, National Records of Scotland (NRS) and the Northern Ireland Statistics and Research Agency (NISRA) — consulted widely on the question topics that users required. Users in the commercial sector and elsewhere made a strong case for collecting income data — income was considered and tested, but was excluded because of concerns about potential negative effects upon public response.

Table 1: Census strengths and weaknesses

Pros	Cons
<ul style="list-style-type: none"> • Population snapshot at a single time point • Consistent method across the United Kingdom • Small-area building blocks, which are largely unchanged from 2001 to 2011 • Responses to several questions can be cross-analysed • New variables may be derived, for example, lifestages, social grade • Census microdata is available for detailed analysis/research — samples of anonymized records for individuals and households • Standard census outputs are available free of charge 	<ul style="list-style-type: none"> • Ten-yearly snapshot — results become out of date in areas of change • It takes c. 2 years to process results (typically) before full outputs are published • Limited to asking questions that will be acceptable to the public, for example, cannot ask an income question • Outputs are mainly pre-designed tables • Individual-level data cannot be matched with other sources (because of the Confidentiality Act)

Source: Dugmore et al. (2011).

A number of other less contentious topics were added to the Census questionnaire for the first time, including:

- Second address (including reason)*
- Passports held (citizenship)
- Month/year of entry into the United Kingdom (for the non-UK born)
- Intended length of stay in the United Kingdom (for non-UK born)
- National identity
- Main language and proficiency in English
- Number of bedrooms*
- Visitor information

*England and Wales only

A summary of the 2011 question topics, differences between countries of the United Kingdom and changes since 2001 are provided by Martin van Staveren.²

Innovations in data collection

The three Census Offices have followed a harmonized approach where possible, and all three took a census on Sunday, 27 March 2011. The data collection operation involved several significant innovations, albeit with some variations in different parts of the United Kingdom. For the first time, forms were sent out by post, rather than being distributed by a field force of enumerators. Postal delivery required a highly accurate address register to be developed in order to avoid the problem of undercoverage — failure to count legitimate households and population — and this register became a central part of the 2011 Census operation. ONS³ describes its address register development and provides an assessment of its quality.

Households had the option to return their completed forms by post or to answer the questions online — overall 16 per cent of returns were made online, with higher proportions in urban centres, as can be seen from Figure 1. The census field force focused its efforts on following up non-respondents and improving the response rates in the worst areas, guided by a new questionnaire-tracking system. As a result, ONS achieved its targets of 94 per cent response rate nationally and over 80 per cent in every local authority district — improving on the rates achieved in the worst districts in 2001. Such response rates would be completely unachievable on any other survey — for the census, they are essential for ensuring the accuracy of the published results.

Processing

Imputing missing records

Following the data collection, the Census Offices carried out the capture and processing tasks that are necessary for any survey operation. In addition, they applied a process known as ‘coverage adjustment’, which imputes data for households and individuals missed during data collection. This process drew on a coverage survey of 350,000 households that was carried out 6 weeks after Census Day to obtain an independent measure of the population in each area. Using the two surveys, a statistical procedure known as dual system

Figure 1: 2011 Census in England and Wales: Percentage of household forms completed online
For local authorities in England and Wales
Source: ONS licensed under the Open Government Licence v.1.0

estimation was applied to estimate the numbers of missing records, and then the gaps in the database were filled by imputation.

All outputs produced from the Census are therefore estimates for the total population, rather than containing ‘holes’ where households and individuals were missed. This will be an important strength of the data when, for example, used to measure the potential market size in any catchment area.

Deriving variables via proxies

Another key strength is that the census collects multivariate data — this means that the Census Offices are able to derive additional variables and classifications using the information supplied by each individual and household. For example, social classifications may be derived, which can provide measures of affluence and lifestyles (eg as a proxy for income data) using variables such as working status, occupation and housing tenure. The National Statistics Socio-Economic Classification is produced as a standard derived variable and is mainly applied by users in the public and academic sectors. Users in the commercial sector tend to employ social grade — the ABC1 system used in marketing and advertising — as their classification of choice. However, it is not feasible for the census to measure social grade. Therefore, for 2001, an approximation model was developed⁴ that predicted each individual’s social grade according to their working status, occupation and other attributes. ONS applied the model to the census database in order to derive an ‘approximated social grade’ and a series of output tables that analyse this variable. For 2011, the social grade approximation has been

As at 11th December 2012

Geography	July-November 2012 First release phases			December 2012 - February 2013 Second release phases			March - June 2013 Third release phases			July-Oct 2013 Fourth release
	1	2	3	1	2	3	1	2	3	
England and Wales	PHE (SY)			KS and QS			LC			DC, T and AF
England	PHE (SY)			KS and QS			LC			DC, T and AF
Wales	PHE (SY)			KS and QS			LC			DC, T and AF
region	PHE (SY)			KS and QS			LC			DC, T and AF
unitary/local authority	PHE (SY) 2nd ADD			KS and QS			LC			DC, T and AF
middle layer super output area		PHE		KS and QS			LC			DC and T
lower layer super output area		PHE		KS and QS			LC			
output area		PHE		KS and QS			LC			
ward				KS and QS				LC		DC and T
health area					KS and QS				LC	DC and T
national parks					KS					DC and T
parishes/communities				KS and QS					LC	
parl. constituencies				KS and QS					LC	
postcode sector					KS					
Welsh Gov. electoral area					KS and QS					

Key

KS

QS

LC

DC

T

AF

PHE (SY)

2nd ADD

Key Statistics (counts derived from more than one variable)

Quick Statistics (univariate counts)

Local Characteristics (cross-tabulations on two or more variables)

Detailed Characteristics (detailed cross-tabulations on two or more variables)

Theme (a range of results on a particular variable)

Armed Forces

Population and household estimates by single year of age and sex

Information about second addresses by age, sex, and type of second address

Figure 2: High-level Release Plan: 2011 census in England and Wales

Source: ONS licensed under the Open Government Licence v.1.0

redeveloped and, at the time of writing, users are looking forward to seeing the results.

Outputs

A surprise in population growth

Census outputs are due to be issued in four releases, starting in July 2012 and running through to late 2013. The ONS high-level release plan for England and Wales is shown in Figure 2 — further details and current status may be reviewed in their online Output Prospectus.⁵ NRS and NISRA have corresponding output plans for Scotland⁶ and Northern Ireland,⁷ respectively.

The initial release in July 2012 has been used to update the population estimates for local authority districts. These 2011 Census results have already provided one surprise for demographers in terms of the extent of population growth — the new population estimate for England and Wales is 476,000 higher than the population estimate rolled forward from the 2001 Census. It is thought that just over half of this difference might be due to underestimation of net international migration over the decade, and that the remainder could be due to a shortfall in the 2001 Census estimate.⁸

As can be seen from Figure 3, the regional growth rates range from 2 per cent to 12 per cent, with the greatest increase being in London.

Harmonized output for the United Kingdom

First-time alignment of outputs

In the past, each Census Office was responsible for producing a set of output tables for its coverage area and little was done to create a single set of results for the United Kingdom as a whole. However, this situation has changed for 2011 as a result of an agreement on harmonization⁹ between the National Statistician and the Registrars General of Scotland and Northern Ireland.

Figure 3: Regional growth since 2001
Population (in millions) for regions of England and Wales in 2001 and 2011
Source: ONS licensed under the Open Government License v.0.1.

Therefore, the Census Offices have been working together to create UK-wide census outputs, which will contain tables that are either fully harmonized or can be released with suitable guidance. A series of tables for the United Kingdom is being published, starting with population estimates by sex and age groups, down to local authority level, in December 2012. Details of further UK releases will be published in an online prospectus.¹⁰

Beyond 2011 programme

Introduction

In April 2011, ONS launched the Beyond 2011 Programme in order to explore alternative options for the future production of census-type data in a more cost-effective way than a full population survey. Representatives from NRS and NISRA have been closely involved in order to initiate similar developments for Scotland and Northern Ireland, respectively, at the appropriate time.

Looking outside the United Kingdom, a number of other countries have been moving away from the traditional full population survey approach, so there is much to be learnt from their experiences. For this reason, Ralphs and Tutton¹¹ produced a review of international practice on census taking early on in the Beyond 2011 Programme. Their review concentrates on the methods used in countries that operate any of the following:

- A traditional census with enhancements using administrative sources, for example, Australia.

Moving away from full-population surveys

- Approaches based on administrative or register data, for example, the Nordic countries.
- Variations on a traditional census, for example, the rolling census in France.

European Union moving away from traditional census taking

Dugmore *et al.*¹² examined the trends in census taking over the last decade, between the 2000 and 2010 census rounds, and found that the trend away from a traditional census is most marked in European Union countries. Of the 18 countries that conducted a traditional census in the 2000 round, seven had moved to a different method for the 2010 cycle, often taking a combined approach involving administrative registers together with a survey or population enumeration.

Developing and testing for 2021

Timescale for decision making

The Beyond 2011 Programme is aiming to make its recommendations by September 2014. This timescale is required in order to allow time for a new approach to be fully developed and tested for 2021. Should the Government decide to retain the traditional approach, then a similar timescale would be needed for development, testing and rehearsal before a 2021 Census.

Support for traditional fieldwork

Initial consultation

In October 2011, the programme launched its first Beyond 2011 public consultation, which focused on two areas:

- Current and future requirements for population and demographic statistics.
- The trade-off between accuracy, geography and frequency at which the statistics are produced (by topic themes such as housing, labour market, etc).

The summary report from this consultation¹³ presented a number of general findings and then went into more detail for each topic theme. The main points were:

- Although users supported the principle of a review, there remained strong support for a traditional fieldwork-based census, together with concern about losing multivariate data for small areas, which was seen as the key benefit of a census.
- Data requirements have increased across all themes — there is a need for high-quality data to support local decision making.
- There is a need for more responsive statistics about new and emerging topics (eg income).
- Users would like results to be produced more quickly and supplied more regularly. However, most users are not prepared to trade accuracy or small-area geography for frequency.
- For many users and topics, there appears to be a dual requirement — frequently updated estimates at local authority level, combined with less frequent updates at a lower level of geography.

Eight options for the next census

Options

The Beyond 2011 Programme started off with a number of census and non-census options,¹⁴ which may be summarized as follows:

Census options

Option 1: Traditional census

Continuation of the current approach, with greater use of online data capture.

Option 2: Short- and long-form census

Similar to the current census approach in that information would be collected from the full population at a single point in time. A subset of the population would complete a traditional long-form questionnaire; the remainder would complete a short-form questionnaire collecting basic demographic, household and family information. Estimates would be derived by combining the two sources.

Option 3: Short-form census plus continuous survey

A short-form census would be carried out to collect basic demographic, household and family information from the full population. Socio-demographic statistics would be collected through a large-scale continuous survey.

Option 4: Rolling census

A census would be conducted in different areas on a rolling basis. For example, 10 per cent of the country might be selected each year so that, over a decade, the whole population would be counted.

Survey options

Option 5: Address register plus survey

Data from an address register would be combined with a survey to estimate the population and its characteristics. Estimates could be updated on an annual basis, but a longer time period may be required to produce estimates for smaller geographic areas.

Administrative data options

Option 6: Administrative data — aggregate

Aggregate data from a variety of administrative sources would be used to produce initial population counts. Data to produce socio-demographic statistics would be collected through survey(s), with potential for including information from administrative sources in the future.

Option 7: Administrative data — full linkage

Record-level administrative sources (including the address register) would be linked together to produce initial population counts. Data to produce socio-demographic statistics would be collected through survey(s), with potential for including information from administrative sources in the future.

Option 8: Administrative data — partial linkage

This option would make use of both aggregate and record-level administrative data approaches. The aggregate administrative data approach would be used to produce initial population counts. Data to produce socio-demographic statistics would be collected through survey(s), with potential for including information from administrative sources in the future.

Investigating and trialling alternatives

Initial research results

Much of the research effort has been going into examining options 6–8, which make efficient use of existing administrative sources such as the National Health Service Patient Register and the Department of Work & Pensions/HM Revenue & Customs Customer Information System. The survey-based approaches (options 1–5) require less investigation, as these are already well understood.

In Autumn 2012, the first set of research results were published¹⁵ and meetings were held with users — progress was reported in a number of areas, including the following:

- Developing a set of criteria to be used in assessing the various methodological options — this helps in building options that will achieve the stated goals.
- Trialling estimation methods using aggregate administrative data (option 6 above) — results were encouraging, but highlighted that full linkage of administrative sources (option 7) is likely to be necessary.
- Investigations into the options for designing a system to produce small-area demographic statistics, based on a sample survey.
- Research into the design of a coverage survey, for use with an ‘administrative data’ model — coverage adjustment will be essential in order to compensate for both ‘over’- and ‘under’ estimation.

Administrative data option is attractive

The latest position on choice of option was summarized by Alistair Calder (ONS) in his presentation¹⁶ to the meetings held in November 2012. Calder advised that two main options were thought to exist:

- Estimation of population size and demographic attributes using a combination of short and long surveys (options 1–4 above).
- Estimation of population size from linked administrative sources (option 7 above) in combination with a sample survey to measure demographic attributes.

Calder reported that if the ‘administrative data’ option can be proven to work and is acceptable to the public, then this type of solution looks very attractive.

However, the size and frequency of the accompanying demographic survey would depend on the quantified benefits of small area data. ONS was working to assess these benefits, and Calder invited users to contribute existing census uses that require small-area data, together with the benefits that they have brought.

Need for economic value

Business case development and benefits quantification¹⁷

Whichever solution is recommended by Beyond 2011, it will need to be underpinned by a robust business case in order to secure funding to proceed. The economic value of benefits will need to adequately cover the implementation costs in order to convince HM Treasury that the investment is sound.

Some parallels may be drawn with the business case that ONS produced for proceeding with the 2011 Census. However, the situation was clearer then, and the benefits did not need to be fully quantified. In effect, the case was justified from two applications:

- The estimated value of more efficient allocation of funding to local authorities (£446m over 10 years);
- The expected value of census data to private sector researchers (eg in geodemographic products) and to retailers for store location (£274m).
- As the total benefit from these two sources exceeded the cost (£432m), there was a clear case for conducting the 2011 Census.

Identifying value of small area data

Looking beyond 2011, there is likely to be a strong benefit case for producing population estimates down to local authority level. However, a business case is required for producing lower-level population estimates and demographic attribute data for small areas. At the time of writing, ONS is looking for help from users in order to make this case by asking questions such as:

- Where does small-area data add real value?
- What would users do if ONS did not provide small-area data and what would the financial consequences be?
- How much is the data really worth to an organization and what is the value of the business decisions that it informs?
- What difference does the data really make and at what geographical level does benefit really accrue?

Three areas of marketing application

Implications for users in marketing

As Peter Sleight¹⁸ explained, users in marketing employ census data in a variety of ways, either as raw variables to describe areas or, more often, as geodemographic classifications such as ACORN and MOSAIC. Sleight describes three main areas of applications:

- Retail — for use in store location projects, using techniques such as market share analysis and branch performance modelling, and also in local store marketing. Heavy users include supermarket chains such as Sainsbury's¹⁹ and Tesco.²⁰
- Direct marketing — for targeting and door-to-door distribution. For example, Peter Mouncey²¹ summarized a variety of ways in which geodemographics is used to manage customer relationships.
- Media — for helping to target local media such as regional press.

Potential impact on geodemographics

Following on from the last section, it seems likely that the Beyond 2011 Programme will have a significant impact on census use in marketing, which will depend on the strength of the business case that ONS constructs for small-area data.

If ONS is completely successful in putting together a strong business case that is accepted by HM Treasury, then it should be able to continue producing population estimate and attribute data down to small-area level. In this case, geodemographic classifications should continue — possibly with changes to the raw variables that are available — and it may even be feasible to obtain more frequent area updates.

If ONS is partially successful with its business case, then it may need to cut back on the level of geography and/or the richness of detail that can be produced. This would severely impact on the geodemographics industry, which relies heavily on census data for its main input — and for which there is no replacement source. Alternative market analysis methods would need to be devised, driven by a more limited amount of census data.

If ONS is unsuccessful with its business case, then the geodemographics industry (in its present form) would most probably disappear and users would have to find completely different methods of market analysis.

Conclusions

Revealing the quality of administrative data

The 2011 Census operation seems to have been highly successful, and this will provide the basis for a valuable demographic snapshot of the UK population for many years to come. The Census Offices are producing a wide range of national, regional and local outputs, designed to meet the needs of users and extract maximum benefit from the data.

The Beyond 2011 Programme is researching into alternative, more cost-effective methods for the future that could replace the traditional census by estimates derived from administrative sources, in combination with sample surveys.

The review is to be encouraged — much is being learnt about the quality of administrative data and best practice methods for deriving population estimates using these sources. However, any recommendation for a new approach will need to be fully tested and proven before moving away from the established census methodology. The research alone will be of value even if a traditional census approach continues to be followed in 2021.

Beyond 2011 needs to deliver its recommendations by September 2014, supported by a strong business case that quantifies the benefits in order to justify the costs. This implies that a benefit case will be required for producing small-area population estimates and demographic data — the input needs to come from all census users, but most importantly from those employing geodemographics in marketing.

Failure to identify the significant uses of small-area data and their benefits could have serious implications.

References and Notes

1. Bermingham, J., Baker, K. and McDonald, C. (1979) 'The utility to market research of the classification of residential neighbourhoods', MRS Conference.
2. Van Staveren, M. (2011) 'The questions asked — what's new since 2001?' Presentation to MRS Census and Geodemographics Group Seminar, available at http://www.mrs.org.uk/pdf/Martin_van_Staveren_presentation.pdf
3. Office for National Statistics. (2012) '2011 Census Address Register', 2011 Census Evaluation Report, available at <http://www.ons.gov.uk/ons/guide-method/census/2011/how-our-census-works/how-did-we-do-in-2011-/index.html>.
4. Meier, E. and Moy, C. (2004) 'Social grading and the census', *International Journal of Market Research*, Vol. 46, No. 2, pp. 141–170.
5. Office for National Statistics. (2012) 'Output Prospectus (England and Wales)', available at <http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-prospectus/index.html>.
6. General Register Office for Scotland. (2012) 'Output Prospectus Scotland', available at <http://www.gro-scotland.gov.uk/census/censushm2011/producing-delivering-data.html>.
7. Northern Ireland Statistics and Research Agency. (2012) 'Northern Ireland output prospectus', available at <http://www.nisra.gov.uk/Census/2011CensusProposedOutputs.html>.
8. Office for National Statistics. (2012) 'Explaining the difference between the 2011 Census estimates and the rolled-forward population estimates', available at <http://www.ons.gov.uk/ons/rel/census/2011-census/population-and-household-estimates-for-england-and-wales/stb-e-w.html#tab-How-the-census-estimates-compare-to-the-2010-mid-year-population-estimates>.
9. General Register Office for Scotland. (2012) 'The Conduct of the 2011 Censuses in the UK: Statement of agreement of the National Statistician and the Registrars General for Scotland and Northern Ireland', available at <http://www.gro-scotland.gov.uk/census/censushm2011/policy-and-methodology/uk-harmonisation.html>.
10. Office for Population Statistics. (2012) 'UK population estimates', available at <http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-prospectus/release-plans-for-2011-census-statistics/uk-population-estimates/index.html>
11. Ralphs, M. and Tutton, P. (2011) 'Beyond 2011: International models for census taking: Current processes and future developments', Beyond 2011 Project, Office for National Statistics.
12. Dugmore, K., Furness, P., Leventhal, B. and Moy, C. (2011) 'Beyond the 2011 Census in the United Kingdom: With an international perspective', *International Journal of Market Research*, Vol. 53, No. 5, pp. 619–650.
13. Office for National Statistics. (2012) 'Beyond 2011 Public consultation on user requirements — report', available at <http://www.ons.gov.uk/ons/about-ons/what-we-do/programmes---projects/beyond-2011/news/reports-and-publications/consultation-reports/index.html>.
14. Office for National Statistics. (2012) 'Beyond 2011: The options explained', available at <http://www.ons.gov.uk/ons/about-ons/what-we-do/programmes---projects/beyond-2011/news/reports-and-publications/methods-and-policies/index.html>.
15. Office for National Statistics. (2012) 'Reports and publications', available at <http://www.ons.gov.uk/ons/about-ons/what-we-do/programmes---projects/beyond-2011/news/reports-and-publications/index.html>.
16. Calder, A. (2012) 'Beyond 2011 user workshop', Presentation to Seminar at Royal Statistical Society on 28 November.
17. Townsend, N. (2012) 'Business case development and benefits quantification', Presentation to Seminar at Royal Statistical Society on 28 November.
18. Sleight, P. (2004) 'Targeting Customers: How to use geodemographic and lifestyle data in your business', 3rd edn, World Advertising Research Centre, Henley-on-Thames.

19. Dellevoet, A. (2011) 'The value of census information in location planning at Sainsbury's Presentation to MRS CGG Seminar', available at http://www.mrs.org.uk/pdf/Annette_Dellevoet_presentation.pdf.
20. Pavey, G. (2008) 'Customer insight in action at Tesco', DUG Conference, available at http://www.demographicusergroup.co.uk/resources/2008conf-GILES_PAVEY.pdf.
21. Mouncey, P. (2004) 'Using geodemographics to manage customer relationships', Presentation to MRS CGG Seminar, available at http://www.mrs.org.uk/pdf/jul04_mouncey.pdf.