Statistical Disclosure Control in the 2011 UK Census: Swapping Certainty for Safety

Joe Frend[*], Carole Abrahams[*], Angela Forbes[**], Paul Groom[**], Keith Spicer[*], Caroline Tudor[*] and Peter Youens[*]

[*] Office for National Statistics, Segensworth Road, Titchfield, Fareham, PO15 5RR, United Kingdom, {Joe.Frend, Carole.Abrahams, Keith.Spicer, Caroline.Tudor, Peter.Youens}@ons.gsi.gov.uk

[**] Statistics New Zealand, Statistics House, The Boulevard, Harbour Quays, PO Box 2922, Wellington 6140, New Zealand, {Angela.Forbes, Paul.Groom}@stats.govt.nz

Abstract: In this paper we will describe the targeted record swapping methodology that has been developed by the Office for National Statistics (ONS) as a statistical disclosure control strategy for 2011 UK Census tabular outputs and how 'sufficient uncertainty' has impacted on this. We discuss the key features of the household methodology and detail how it has been adapted to create the communal establishment methodology. This paper will discuss how we are using population thresholds and explain how we can measure the uncertainty created by the disclosure control methodology.

1 Introduction

The 2011 UK Census was conducted on the 27th March 2011. With detailed information on the socio-economic characteristics of the UK population, the data gathered from the census will allow governmental departments and other organisations to better understand how to allocate their resources for public services such as transport, health services and education. Census outputs will be released in different formats; pre-planned tables, user requested commissioned tables and census sample microdata. With the release of aggregate and individual data through these outputs, there comes a risk of identification of an individual or other entity (e.g. households, families, businesses, communal establishments[1]) and discovery of confidential information about them. With the production of official statistics dependent on the cooperation and trust of the population, individuals' information must remain protected to maintain this trust. The UK Census Offices of England and Wales, Scotland and Northern Ireland must also ensure that the confidentiality of the census respondents is upheld for all outputs to comply with their legal requirement. With an increased number of outputs to be released and with a constant demand from users for greater detail and accuracy, Statistical Disclosure Control (SDC) plays an important part in the success of this, potentially the last, UK Census. Indeed, in the last ten years in the UK, if not across all National Statistical Institutes (NSIs), we

[1] Communal establishments are managed residential accommodations. Examples of communal establishments include prisons, hospitals, hotels and university halls.

have seen a divergence of opinion between the data providers (demanding greater levels of privacy) and the data users which, coupled with improvements in technology and data mining, requires more advanced SDC methods. SDC aims to ensure the confidentiality of respondents is protected and is achieved either through the treatment of the microdata (pre-tabular), the output tables (post-tabular) or a combination of both. There will inevitably be some damage to the utility of the output and it is a key objective for the chosen SDC method to minimise this utility loss whilst maximising the protection against disclosure. This paper will outline the primary SDC methodology of targeted record swapping that will be used to protect the 2011 UK Census data. The paper will also describe how the concept of 'sufficient uncertainty' has resulted in a shift towards exploiting the ambiguity already surrounding the data and how this has affected the SDC methodology.

The disclosure risks inherent in frequency tables surround cells of low counts such 1s, 2s and 3s. It is, however, the position of 0s in a table that may lead to information about an individual or group of individuals being learnt. It is therefore important for the census SDC method to add ambiguity to both small cells and zero cells in tabular outputs. A common pre-tabular SDC method is record (data) swapping between households matched on control variables (Willenborg & de Waal, 2001). Random record swapping was selected as the SDC method for the 2001 UK Census. However, at a late stage, concerns were raised that the public would not perceive that the data had been properly protected. In light of this, a decision was made by the Office for National Statistics (ONS) and Northern Ireland Statistics and Records Agency (NISRA) to employ the post-tabular method of Small Cell Adjustment (SCA) which rounded small cells either up or down to a base number. Because the General Register Office of Scotland (GROS) (which has since been renamed the National Records of Scotland (NRS)) decided against SCA there was a lack of harmonisation across the UK. Furthermore, because each table was independently adjusted so the sub-totals and totals took account of the SCA, there were inconsistencies between the cell counts in different tables. The lack of harmonisation, inconsistency between tables and damage to utility (especially with sparse tables) led to considerable criticism from users. In some circumstances the SCA was found to be possible to unpick or provided little protection against disclosure by differencing (Shlomo, 2005).

In November 2006, the Registrars General of England and Wales, Scotland and Northern Ireland agreed upon the key aims of the UK Census Outputs including a "[c]ommon UK SDC methodology for Census 2011 outputs that minimises disclosure risk whilst maximising utility" (Office for National Statistics, 2006). The agreement on the SDC UK Policy was based on the principle of protecting confidentiality as set out by the National Statistics Code of Practice. This has since been superseded by the Code of Practice for Official Statistics following the introduction of the Statistics and Registration Service Act (SRSA) 2007, though the UK SDC Policy is still valid as it is in line with Section 39 of the SRSA (The

National Archives, 2007). Given the lack of harmonisation and the criticism caused by SCA in the 2001 Census, an important step taken through the Registrars General agreement was the definition of what constitutes a disclosure risk. First, the Registrars General acknowledged that discovering new information about an individual through outputs (attribute disclosure) is the key disclosure risk, rather than just the identification of an individual. This decision was particularly beneficial for SDC relating to Communal Establishments (CEs) as protecting specifically the identification of CEs would greatly damage the data. Given the impracticality of entirely removing risk caused by small cells, the Registrars General also concluded that the Code of Practice would be met if no outputs were produced that would allow the identification of an individual (or information about them) with a high degree of confidence. It was therefore agreed that small counts (0s, 1s, and 2s) could be included in publicly disseminated census tables provided

a) there is sufficient uncertainty as to whether the small cell is a true value; and
b) that creating this uncertainty did not significantly damage the data.

Whilst this decision had great benefits for the quality of the census outputs, avoiding the large amount of information loss caused by small cell adjustment or other rounding methods, a challenge was faced to create a SDC methodology which would ensure the same levels of data protection without significant damage to the data. It is well known that even before SDC is applied, there exists a level of ambiguity created through non-response, coverage adjustment, item non-response, imputation errors, respondent errors, capture errors and even the length of time from data collection to data dissemination. Whilst the census processes are used to clean the data and ensure a higher level of data quality they will still distort the collected data before outputs. With knowledge of where and to what extent ambiguity is created, the selected SDC method can be limited and tailored to ensure the data are not overprotected. Aptly, 'sufficient uncertainty' is a rather ambiguous expression; how can you calculate the uncertainty of a cell value, and at what point is this uncertainty 'sufficient'? A quantitative measure of uncertainty has been developed for this purpose and will be discussed in *Section 5*.

An initial review of a wide range of SDC methods was completed in 2007 to create a short-list of methods for the 2011 UK Census. With an emphasis on selecting a method which would be accepted by users and assured additivity and consistency across tables, the three methods short-listed for further evaluation were record swapping, over-imputation and the ABS Cell Perturbation Method. Spicer & Tudor (2009) describe the quantitative and qualitative evaluation process that led to the decision to use record swapping across the UK for the 2011 Census. As well as in the 2001 Census, record swapping has successfully been used for protecting census tables at the US Census Bureau. The expected weaknesses of the method were how to deal with population uniques and special populations such as CE residents. Importantly, the key strengths of record swapping are that a) no person or

data items are removed ensuring that outputs for national and high level geographies are unaffected, and b) that the method has been used before for national censuses.

2 Targeted Record Swapping for Households in the 2011 UK Census

Record swapping is a pre-tabular method of perturbation where the geographical variables of a sample of households are swapped with those of other households matched on a set of key variables. By swapping the geographical information of paired households the method ensures that the relationship of variables within the households will remain unchanged and also ensures the marginal distributions are preserved at higher geographies. The sample of households can either be selected randomly or by targeting households that pose a greater risk of disclosure. Using 2001 data, Shlomo *et al.* (2010) compared the effects of targeted and random record swapping methods. Their analysis showed that targeted record swapping only required approximately half the swap rate of random record swapping to attain the same level of protection, whilst generally preserving higher levels of utility.

2.1 Calculation of Risk Scores

As mentioned before, the cells which require the most protection are those with small counts which will have been aggregated from rarer individuals or households. By targeting the records that pose a greater risk of disclosure we are able to create more ambiguity for those small cells and limit the damage to utility for the rest of the data. Whilst the focus is on protecting against attribute disclosure, to be able to glean new information about an individual an intruder does first have to identify an individual in a table, and to do so, the intruder must have some prior knowledge. To determine which variables should be used to assess the rarity of a household, we considered which of these could be described as 'visible' and used to identify a record. Whilst the risk of attribute disclosure will increase with 'visible' variables, the presence of 'sensitive' variables in a table may increase the probability of intruder motivation and subsequently the risk of disclosure. The census variables were classified as sensitive and/or visible, using the following definitions:

- **visible:** if their attributes are 'constantly or frequently in the public view'
- **sensitive** if they involve 'information of a delicate nature that could cause harm or distress if revealed'

With the census variables defined as sensitive and/or visible a number of variables were selected from which to calculate the risk scores. ONS will not release specifics of the SDC methodology such as swap rates, matching variables and risk variables in order to maintain the level of confidentiality protection offered by the methods.

The 2011 UK Census record swapping methodology will be carried out in large geographic blocks called Delivery Groups (DGs). Within these DGs are three levels

of nested geographies; Local Authority Districts (LADs), Middle Super Output Areas (MSOAs) and Output Areas (OAs), where OAs are the smallest geography for which detailed outputs are produced, consisting of, on average, 300 persons and 120 households. Because the majority of Census tables consist of individual persons, households are identified as high risk dependent on the individuals residing there. Individuals are given a risk score reliant on frequency counts on univariate distributions on the chosen set of *risk variables*. A risk score, as described by Shlomo *et al.* (2010) is calculated at all three geographical levels for each individual record (not including imputed records), as follows:

- For M risk variables with k_m, $(m = 1,...,M)$ categories at the geographical level g, the frequency $N^g_{k_m}$ is calculated.
- For every individual with values of categories $k = (k_1, k_2, ..., k_M)$, a risk score (*HR)* is calculated at geography level g by taking the average of the reciprocals of the counts:

$$HR^g_k = (\sum_{m=1}^{M} 1 / N^g_{k_m}) / M$$

A threshold level is set for each of the three geographies, whereby all records with a risk score higher than the threshold at that geography are deemed to be high risk. Any household containing a high risk individual is then flagged as a high risk household.

2.2 Selection of Samples and Matches

As mentioned in *Section 1*, prior to the application of the SDC methodology, there will often be uncertainty created in the data by other census processes. One of the most impacting procedures to the data is the imputation of records to account for non-response. Non-response rates, and concurrently, imputation levels will vary between delivery groups. In areas where there are high levels of uncertainty gained through imputation, there is less of a need for SDC than areas with lower levels of imputation. Therefore, the percentage of households selected to be swapped within the delivery group will have an inverse relationship to the imputation rate, though all areas will have some swapping. Whilst the sample size is set for each DG, it is allocated across the OAs dependent on a) the number of non-imputed households in the OA, and b) the percentage of high risk households in the OA, though a threshold percentage of households selected in any one OA is employed. The high risk households will be given a higher probability of being selected for swapping, though all non-imputed households will have a non-zero probability. No imputed households will be selected for either sampling or matching, though real households with additional imputed persons are eligible.

With the sample households selected, each sample household will be matched with another household outside of the sample. To ensure counts of households and

individuals remain correct for each level of geography, households will be matched on household size as well as a set of *control* or *matching variables*, selected to preserve certain characteristics. Initially, those non-sampled households flagged as high risk are considered as possible matches to again place the emphasis on protecting the more risky records. If an individual is unique at geography, on any one of the risk variables, the individual and subsequently the household is flagged as unique at that geography. If then selected for swapping, the household would be swapped outside of the geography in which it is unique. So, for example, a household unique at the MSOA level will be swapped out of its MSOA but still within the LAD. A unique flag does not necessarily imply that the household will also be flagged as high risk. Aside from these exceptions, we will always look to minimise the distance between the sample household and a match, but the sampled household will always be swapped outside of its OA. This better preserves the utility of the data at higher levels of geography. It will not be uncommon for some sampled households not to be matched at a low geography and with the set of matching variables. If this occurs, we initially relax the detail in the matching variables before looking at swapping at a higher geography. Once a match is found all geographical variables will be swapped.

3 Targeted Record Swapping for Communal Establishments in the 2011 UK Census

As previously mentioned, one of the weaknesses of the record swapping methodology was thought to surround persons living in Communal Establishments (CEs). However to provide consistency in the SDC methodology an adaptation of the household record swapping code is used to protect the Census 2011 CE data. The location of a large number of CEs such as prisons or general hospitals, are very much public knowledge. Hotels and care homes are also likely to be known within the local area. Given this, it is difficult and in many cases pointless, to protect the actual establishments from identity disclosure. To do so would either mean perturbing the CE counts or the counts of persons in CEs, neither of which are suitable in a record swapping methodology. Regardless, it is not a priority to protect the identity of a CE given the Registrars General agreement and the limited amount of information the census outputs produce for the CE itself. However, if a CE is identified within an output, it can essentially be viewed as a smaller geography and subsequently would lead to a greater disclosure risk for the individuals living within the establishment (referred to from now as residents). The primary difference between the household record swapping method and the CE method is that, rather than swapping whole CEs, individual records are swapped. Whilst this is mostly out of necessity (in the majority of cases, there will not be a suitable match to swap whole CEs without greatly damaging the data and/or making the swap obvious in outputs), it ensures that the emphasis of record swapping is placed on protecting the

information of the residents and also preserves both the counts of CEs and persons at all geographies.

A problem that is posed when adapting the household record swapping methodology is not the transition from swapping whole households to swapping individuals, but the aim to maximise utility. At the heart of this difficulty is the homogenous nature of the residents of certain CE Types and the impact on utility that a distortion of particular characteristics will cause. A prison, for example, will have different resident characteristics to those in university halls. Within the resident CE population, there is a separation into three resident groups:

- **Clients:** Non-staff residents for which the CE caters (e.g. patients of a hospital, clients of a hotel).
- **Staff:** Staff and owners who live in the CE. This does not include staff of the CE who live at another location.
- **Family:** Family members or partners that live in the CE with either a member of staff or a client.

In the same way that we wish to preserve the characteristics of the CE Type[2], maintaining the characteristics of the different resident groups is also a key aim. Whilst the preservation of characteristics is dependent on the matching process of record swapping, it is also possible to maximise utility by minimising the amount of swapping. Because of the variety of establishments falling under the definition of CEs, it is clear that there are some CE Types which have a greater risk of disclosure than others. We therefore look towards varying the swap rate dependent on how great the risk of attribute disclosure is for a CE Type within a particular geography. A final restriction on the record swapping methodology for CEs is the logistical requirement to limit the swapping of records to within the DG. The three aims of the CE record swapping methodology are therefore a) to keep record swapping within delivery groups, b) to ensure a greater proportion of individuals are swapped in CE Types which would have a higher impact attribute disclosure, and c) to develop the matching process to minimise utility loss.

3.1 Communal Establishment Protection Scores

It was observed in the 2001 UK Census that the response rates and data quality for the CE population were significantly worse than for the household population but we cannot assume by default that the response rates and data quality of every CE will be poor. There will be CE Types with significantly higher or lower response rates and much like with household response rates, these rates will also vary across the UK. Because we cannot assume low response rates for all CE Types, we are not using the imputation rates to vary the swap rates. The CE methodology will instead vary the

[2] CE Type relates to the category in which the CE will appear as in the Census output tables.

swap rate based on a calculated protection score (PS), which is taken as the sum of different risk factors and is calculated for every CE Type in each MSOA. MSOAs, which contain 7000-8000 persons, have been the geography chosen at which to calculate the scores because this is the lowest geography at which tabular outputs are produced which combine CE Type with other variables, besides counts of staff, clients and family residents. The first three risk factor scores are solely dependent on the CE Type and its rarity within the area:

- **Counts of CE Type in MSOA:** As stated in *Section 3*, if a CE is identified, it can essentially be viewed as a smaller geography, in which an individual may be identified. A single establishment of a particular type is easier to identify and therefore the risk of attribute disclosure is increased.
- **Uniqueness of the CE Type in the LAD:** Because the higher geography has outputs of more detail, a unique CE Type within the LAD requires greater protection.
- **Impact of identity disclosure in a CE Type:** Although the emphasis of SDC for the 2011 UK Census is placed on protecting against attribute disclosure, identifying an individual as living in a particular CE Type is in itself an attribute disclosure. Identity disclosure of an individual that lives in a 'high impact' CE (such as a prison) will be more damaging than that of identity disclosure of an individual in a 'low impact' CE (such as a hotel).

Another important influence on disclosure risk in a CE is the count of residents. Since the counts between the three resident groups will vary, we require different swap rates for clients, staff and family residents. We calculate both a Client Protection Score (CPS) and a Staff Protection Score (SPS). Family residents will be discussed in *Section 3.4*. A final factor affecting the SPS is whether the CE has a high or low client turnover[3]. Because the detailed outputs will be released over a year and a half after the day the Census was conducted, many CEs with a high client turnover will naturally have higher uncertainty around apparent attribute disclosures. Greater SDC protection will therefore be needed where there is low client turnover. The protection scores are calculated by multiplying together the relevant scores from the risk factors affecting the disclosure risks. The factors and respective scores used in calculating the CPS and SPS are shown in *Tables 1* and *2*.

[3] If it is very unusual for the length of stay of a client to be more than a year, the CE Type is considered to have a high client turnover. Otherwise, the CE Type has a low client turnover.

Score	A CE Type count in MSOA	B Is the CE unique within the LAD?	C Is the CE Type high impact?	D1 Number of clients within the CE Type	E Client Turnover
4	-	-	-	1-15	-
3	1-2	-	-	16-40	-
2	3-5	Yes	Yes	41-100	Low
1	6+	No	No	101+	High
0	-	-	-	0	-

Table 1: Scores for each of the factors affecting the Client Protection Score (CPS).

Score	A CE Type count in MSOA	B Is the CE unique within the LAD?	C Is the CE Type high impact?	D2 Number of staff within the CE Type
3	1-2	-	-	-
2	3-5	Yes	Yes	1-10
1	6+	No	No	11+
0	-	-	-	0

Table 2: Scores for each of the factors affecting the Staff Protection Score (SPS).

For example, the CE Type 'Prison Service establishment', considered to be a high impact CE with a low client turnover, with a count of 1 in the MSOA and the LAD, containing a total of 250 clients and 12 staff, would have a CPS and SPS calculated as follows:

- A = 3, B = 2, C = 2, D1 = 1, D2 = 1, E = 2
- CPS = 3 x 2 x 2 x 1 x 2 = 24 SPS = 3 x 2 x 2 x 1 = 12

By contrast, the CE Type 'University', considered to be a low impact CE with a high client turnover, with a count of 6 in the MSOA, containing a total of 24 clients and 1 staff, would have a CPS and SPS calculated as:

- A = 1, B = 1, C = 1, D1 = 3, D2 = 2, E = 1
- CPS = 1 x 1 x 1 x 3 x 1 = 3 SPS = 1 x 1 x 1 x 2 = 2

3.2 Communal Establishment Swap Rates

With the protection score calculated for both staff and clients of a CE Type in an MSOA, a swap rate is assigned based on CPS and SPS, as shown in *Tables 3* and *4*, where A% is the lowest percentage swap rate and C% is the highest. Regardless of what swap rates are chosen for the Census, if we were to round to the nearest integer the number of records to be sampled for each CE Type population, it would more

than likely result in cases of CE Types having no records swapped. To ensure that swapping occurs for every CE Type for each MSOA, the number of records to be swapped will always be rounded up to the next integer. However, this will imply that no matter how risky a CE is, if there is only one staff or client record, it will always be sampled. Whilst this should not occur too regularly for client records, it may be a relatively common occurrence for staff records. To overcome this and avoid significant damage to the staff resident data we will attach a probability to the record being sampled which will depend on the SPS score.

Client swap rate =	0%	if CPS is	0
	A%		1-5
	B%		6-25
	C%		26+

Table 3: The client swap rates vary depending on the Client Protection Scores.

Staff swap rate =	0%	if SPS is	0
	A%		1-5
	B%		6-11
	C%		12+

Table 4: The staff swap rates vary depending on the Staff Protection Scores.

Returning to our examples, the prison establishment with a CPS score of 24 and an SPS score of 12 would have B% of client records sampled and C% of staff sampled. With a CPS of 3, the 'University' CE Type would have the lowest percentage (A%) of client records selected for the swap sample. With only 1 staff record, there would be a non-zero probability of the record being sampled. With an SPS of 2, this probability will be lower than if the SPS was, for example, 12.

3.3 Selecting the Sample

The sample selection process will be similar to that of households whereby a risk score is calculated for each individual record, however, whilst unique records will be flagged at each geography, the risk score will only be calculated at MSOA level and in relation to the other records within the resident group and CE Type. Rather than setting a threshold and flagging high risk records, the risk scores are used as the weighting when selecting the sample of records to be swapped. Those records with a higher risk score will then have a greater chance of being sampled for swapping.

3.4 Family Residents

Unlike families in households, there is no family composition matrix to connect the family members either to each other, or to any of the staff or client residents. This enables us to swap the family residents individually rather than as a whole family. The number of family residents living in CEs is very low; in the Census 2001 most

LADs had between 0-15 family residents living in CEs. Because of this, we shall set one single swap rate and select the sample from across all CE Types in the LAD.

3.5 The Matching Process

Because the characteristics of clients will differ from those of staff within CE Types, it is important to ensure that swapping is carried out within resident groups. In the same way, the amount of matching on CE Type is to be maximised. It is an aim to keep swapping within CE Type, however, because swapping must be kept within the DG, in many circumstances this will not be possible (e.g. there is only one prison establishment in the Bristol DG). Therefore, the matching variables play a crucial role in ensuring that where swapping does occur between CE Types, the key variables that define the CE Type populations are preserved as much as possible. Because individuals aged under 16 do not answer particular questions on the census form, this has in turn forced the client records to be split between those aged 16 and over, and those under 16. We therefore have three matching groups, each with a unique set of matching variables. Because of the lack of records, it is required to swap family records with either staff or clients. Family records are therefore allocated out to one of the three matching groups depending on the age of the individual and whether they fit into a particular variable category associating them more with either staff records or client records. As with households, we may find that a match is not found using the specified matching variables in which case, we will use a coarser set of variable classifications before looking for matches at a higher geography and then outside of the CE Type. As with households, imputed records will not be swapped.

4 The Use of Thresholds

The SDC methodology of record swapping targets rare records; these will be the records that form the cells with small counts in output tables. As we increase the detail of an output table either through the detail of the constituent variables (which increases the number of cells of a table as the variables become more detailed) or the size of geography (which reduces the records in a table as we focus on smaller areas), we increase the number of small cells and subsequently the number of attribute disclosures. Because it wouldn't be possible to protect them all, to ensure the tables produced are not to a detail which will cause too large a risk of disclosure we employ both a population threshold to the geography and the sparsity rule to the table (which ensures all tables have an average cell size greater than 1). Thresholds are in place for OAs to make them a consistent size and also to ensure they are not too detailed to produce outputs for, but where else can we use population thresholds to allow for the safe production of outputs?

4.1 Workplace Zones

Under Section 39 of the SRSA (The National Archives, 2007), statistics on workplaces must be protected just as it is provided for statistics based on residents. Even though it is not the businesses themselves that have provided the information, but individuals who work for them, the aggregation of the individual statistics generates a set of business statistics and as much protection should be afforded to individual businesses as individual persons. In previous censuses, workplace statistical outputs have been released on the same geographies as those used for private households and household residents, even where workers and workplaces may be spread unevenly. The result for the 2001 Census was that workplace tables were not released as Standard Tables at geographies below district level, and those released at district level were done so without further protection only at a low industry variable detail (12 categories). The primary geography for 2011 Census outputs will be the output area hierarchy that has been designed on the basis of household and resident populations. There was an expectation that this would again cause problems for outputs regarding workplaces because of the uneven distribution of workplaces and workers by OAs, so to improve outputs provision for workplaces, a new geography has been proposed. These workplace zones will be created by aggregating or splitting OAs to reach a set threshold level and, as well as providing greater levels of protection, will also have the benefit of allowing the release of more detailed and more consistent statistics on workplaces. Because the workplace zones will be created to lie within MSOA boundaries, statistics will be able to be produced unaffected at higher geography levels. Whilst much of the protection of workplace based statistics will come from the design of the workplace zones the data are also afforded the protection of record swapping and further uncertainty created by the other processes and respondent error, which tends to be greater for workplace statistics.

4.2 Small Population Tables

Thresholds are being used predominately in connection with geography, however, for the 2011 Census we will also be using population thresholds to allow for the production of outputs in connection with *small* or *minority populations* on variables such as ethnicity, religion and country of birth. Though not possible for every MSOA, a minority population that meets a specified population threshold size will be considered large enough to enable detailed outputs to be produced at MSOA level and higher geographies. These tables show an alternative use of thresholds and reflect the changing outlook of national statistical agencies to make data more available.

5 Measuring Uncertainty

With a SDC methodology based on creating 'sufficient uncertainty', we require a measure for uncertainty to ensure we have protected the outputs adequately. By allowing small counts in the tables we are also allowing cases of attribute disclosure, however, these attribute disclosures may, or not, be real. If a user is aware of a person who 'should' be in a cell that is shown as zero, or *vice versa*, it is likely to throw considerable doubt as to whether *any* apparent disclosure is real and reduce the chance of a claim. We have created a theoretical measure of doubt, devised as a combination of two success measures.

The first success measure calculates the proportion of real cases of attribute disclosure removed in the output table either by the imputation of records or the record swapping methodology. Where cases of attribute disclosure have been removed, the chance of a user making a claim is reduced.

We have recognised that levels of uncertainty are raised where false cases of attribute disclosure are created through either imputation or record swapping. Therefore, our second success measure calculates the proportion of the apparent attribute disclosures that are false. Where cases of false attribute disclosure have been created, the likelihood of a correct claim of attribute disclosure is reduced. With these success measures defined, the *doubt measure* can be calculated as follows:

- Doubt = [1 – (1-S1)(1-S2)], where
- S1 = Proportion of real attribute disclosure cases protected
- S2 = Proportion of apparent attribute disclosure cases that are not real

With this doubt measure, we will be able to assess the uncertainty created by imputation and record swapping for an output table. The value the doubt measure is required to meet for the uncertainty to be deemed 'sufficient' is not released to maintain the protection of the methodology. Whilst we shall be using the doubt measure to assess the effectiveness of the record swapping methodology, we shall also be assessing the utility loss, primarily, using the Average Absolute Distance measure as detailed by Shlomo *et. al* (2010).

6 Summary

This paper has provided an overview of the SDC process for the 2011 UK Census. There has been a shift towards providing as much data as possible for users, within the legislation, with the aim to release small cells in outputs this time round. We have discussed the development of the record swapping method for households including key decisions to use targeted record swapping and varying the swap rates according to the imputation rates of the DGs. A communal establishment swapping

methodology has been developed which has in past censuses been overlooked because of the difficulties involved such as the homogenous nature of certain CE Types and the unique CEs in delivery groups. We have developed a protection score method to reduce the swap rate for those CE Types that have less of a risk of disclosure. The matching process has been advanced to preserve specific population characteristics and to allow for swapping individuals between CE Types where required. Data quality is expected to be lower in CEs than in households and thus offer some natural uncertainty. However, we cannot rely on this assumption for the privacy of individuals in CEs and this methodology will provide the necessary further protection without significant damage to the data. A quantitative measure of *doubt* has been developed to assess whether a sufficient level of uncertainty, and therefore, protection has been reached.

Striking the balance between utility and protection is at the heart of effective disclosure control. Users may be forgiven for thinking that this balance was skewed against data utility in the 2001 UK Census. However for the 2011 Census, we have created an improved SDC methodology which is harmonised across the UK Census Offices for both households and communal establishments, that focuses disclosure control where it is needed most. It allows for additivity and consistency between tables, limits the damage to utility and, crucially, protects the information of census respondents.

References

The National Archives (2007) *Statistics and Registration Service Act*. Available at: http://www.legislation.gov.uk/ukpga/2007/18/pdfs/ukpga_20070018_en.pdf (Accessed: 27 March 2012)

Office for National Statistics (2006) *Registrars General's Agreement*. Available at: www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/producing-and-delivering-data/confidentiality/index.html (Accessed: 27 March 2012)

Shlomo, N. (2005). *Assessment of Statistical Disclosure Control Methods for the 2001 UK Census,* Joint UNECE/Eurostat work session on statistical data confidentiality, Geneva, 9-11 November 2005.

Shlomo, N., Tudor, C. & Groom, P. (2010). *Data Swapping for Protecting Census Tables,* Privacy in Statistical Databases, Corfu, 22-24 September 2010.

Spicer, K. & Tudor, C. (2009). *Balancing Risk and Utility – Statistical Disclosure Control for the 2011 UK Census,* Joint UNECE/Eurostat work session on statistical data confidentiality, Bilbao, 2-4 December 2009.

Willenborg, L. & de Waal, T. (2001). *Elements of Statistical Disclosure Control, Lecture Notes in Statistics, 155.* Springer Verlag, New York.