Statistical Disclosure Control for the 2011 UK Census

Jane Longhurst, Nicola Tromans and Caroline Young,
Office for National Statistics

1. Introduction

Every 10 years since 1801, the UK has set aside one day for the census, whereby information is obtained on every member of the population. It is the most complete source of information about the population that we have with details of family composition, health, employment and other socio-economic characteristics. The information provided allows central and local Government, health authorities and many other organisations to target their resources more effectively and to plan housing, education, health and transport services for years to come. The next census is due to take place in 2011.

Census data is released in a number of different formats; standard pre-planned tables, commissioned tables requested by users and Census sample microdata. In addition in 2011 the aim is to release user defined tables via flexible table generating web-based software. Publishing aggregate or individual data carries the risk that individuals or entities could be identified and confidential information about them could be released. The UK Census Offices need to protect the confidentiality of census respondents for a number of reasons. The production and use of official statistics depends on the cooperation and trust of citizens. Such trust cannot be maintained unless the privacy of individuals' information is protected. There are also legal and policy obligations that must be respected. The Census Act 1920 as amended by the Census (Confidentiality) Act 1991 and the Census Act (Northern Ireland) 1969 as amended by the Census (Confidentiality) (Northern Ireland) Order 1991, make it an offence for the Registrar-Generals (or any person under their control or a supplier of any services to them) to disclose any personal census information to another person without lawful authority. The National Statistics Code of

Practice sets out principles for protecting confidentiality. These include the principle that:
'The National Statistician will set standards for protecting confidentiality, including a guarantee that no statistics will be produced that are likely to identify an individual unless specifically agreed with them.'

The aim of Statistical Disclosure Control (SDC) is to ensure that statistical outputs provide as much value to the users while protecting the confidentiality of information concerning individuals or entities. SDC methods modify, summarise or perturb the data and there are a range of different methods that can be used to protect different census outputs. SDC methods can be pre-tabular (applied to the underlying microdata) or post-tabular (applied to tables).

A pre-tabular method of disclosure control, random record swapping, was initially planned for the 2001 UK census tables. This method of disclosure control was followed up by applying population thresholds to the tables. The General Register Office for Scotland (GROS) adopted smaller thresholds than the Office for National Statistics (ONS) and the Northern Ireland Research Agency (NISRA). Prior to releasing tabular outputs from the 2001 Census concerns were raised that the public would perceive that no disclosure control method had been applied. ONS decided that the additional method of small cell adjustments was required for tabular outputs. The small cell adjustments added more uncertainty and removed small cells from tabular outputs. NISRA also applied the additional method of small cell adjustment but GROS did not. This late change in SDC methodology and lack of UK harmonisation caused a number of problems for users. A different SDC technique was used to protect the microdata samples or Sample of Anonymised Records (SARs) from the 2001 Census. The disclosure risk was reduced by recoding variables and applying PRAM (Post-Randomisation Method), a perturbative microdata disclosure control technique for categorical variables.

This paper describes the strategy that is being employed to develop an SDC solution for the 2011 Census. The key aim is to ensure a harmonised UK SDC strategy for all outputs (pre-defined tabular outputs, microdata samples and possibly flexible user defined tabular outputs) which ensures that the public interest in the figures is met while managing data confidentiality risks. The most desirable qualities for the SDC strategy are;

- Maximum data utility
- Minimum disclosure risk
- Acceptable to users
- Simple to understand and transparent
- Easy to implement

The next section of the paper provides a high level description of the project. Section 3 provides a high level overview of possible SDC methods that could be used to protect census tables and focuses on three methods as examples; record swapping, random rounding and cell perturbation. Sections 4 and 5 provide an evaluation of these methods demonstrating the approach that will be adopted to decide on the SDC method(s) that will be used to protect 2011 Census outputs.

2. Approach

2.1. Development and agreement of UK SDC Policy for 2011 Census Outputs

In November 2006 the UK SDC Policy position for the 2011 Census was agreed by the Registrars General of Scotland, England and Wales and Northern Ireland. The Registrars General have agreed to aim for a common UK SDC methodology for 2011 Census outputs to achieve harmonisation. The SDC Policy position is based on the principle of protecting confidentiality set out in the National Statistics Code of Practice. The Registrars General concluded that the Code of Practice statement can be met in relation to census outputs if no statistics are produced that allow the identification of an individual (or information about an individual) with a high degree of confidence. The Registrars General consider that, as long as there has been systematic perturbation of the data, the guarantee in the Code of Practice would be met. It has therefore been agreed that small counts (0's, 1's, and 2's) could be included in publicly disseminated Census tables provided that

a) uncertainty as to whether the small cell is a true value has been systematically created; and

b) creating that uncertainty does not significantly damage the data.

The exact threshold of uncertainty required has not been decided. The Registrars General will make this judgement at a later stage in the context of results from methodological research into the balance of protection afforded, and damage caused by various SDC methods. The decision to allow small cells in publicly disseminated tables means that both pre-tabular methods and post-tabular methods or combinations of the two can be considered for 2011. The Registrars General have expressed a preference for pre-tabular methods, provided there is not undue damage to the data.

The UK SDC policy also highlighted the following points;

- Aim is to make as much of the census tabular outputs as possible publicly accessible. However, if certain tabular outputs are seriously compromised by SDC then these could be released under other access arrangements (e.g. licence or safe setting) where data access restrictions allow less stringent levels of SDC to apply in order to protect data utility.
- It is considered that attribute disclosure is the key disclosure risk, because identification reveals no new information to the user. Attribute disclosure involves a user discovering something new from the census data that was not previously known to them.
- Consistency and additivity across tabular output is a priority for users and these will be given a high priority when assessing the utility of SDC methods.

- Methods will be chosen which afford an acceptable level of protection and preserve the highest level of utility of outputs.
- Clear explanations will be given to users and expert audiences on the protection afforded by the SDC strategy and other steps applied which protect confidentiality.
- SDC methods for all types of census output will be assessed concurrently because of their interdependencies.
- Users will be updated and consulted during the research period.
- An Independent review will be conducted by the UK Census Design and Methodology Advisory Committee.

2.2. Governance for the 2011 Census SDC Strategy

A UK SDC working group has been formally set up to steer work, provide advice and quality review work associated with developing the SDC methodology for the 2011 Census. The working group consists of representatives from all three UK Census Offices to ensure a harmonised approach to the development of the 2011 Census UK SDC Strategy that is in line with the agreed policy.

A Disclosure Control Subgroup of the UK Census Design and Methodology Advisory Committee (UKCDMAC) is also currently being set up. This subgroup will be responsible for providing advice on methodological issues and will act as a formal quality review panel for the SDC workpackage prior to seeking methodological agreement from UKCDMAC. The UK SDC strategy will be signed off by the UK Census Committee (UKCC).

2.3. 2011 Census UK Statistical Disclosure Control (SDC) Work plan

A work plan for the methodological research phase of the 2011 Census UK SDC strategy has been developed. The plan addresses (pre-defined) tabular outputs, microdata samples and flexible user defined tabular outputs whilst taking into account the impact of interactions between these types of output. An outline of the agreed approach for developing the SDC strategy follows;

The initial stage of methodological research involved conducting a review of SDC in a census context. This review will facilitate the development of the SDC strategy for the 2011 Census by drawing together;

i) research conducted prior to implementing SDC in the 2001 Census
ii) reasoning behind SDC decisions for the 2001 Census
iii) evaluations of SDC methods used in the 2001 Census
iv) lessons learnt from 2001
v) international approaches to SDC
vi) work already conducted for the 2011 Census

Following this a high level review is being conducted to address the advantages and disadvantages of a wide range of SDC techniques for protecting all types of 2011 Census outputs and the issues concerned with implementation and the interactions between the outputs. Examples are provided in Section 4. Using this high level review, a preliminary list of SDC techniques which should be explored further will be drawn up. SDC methods not on this short-list will be discounted from further research.

The short-listed disclosure control methods will then be evaluated using a disclosure risk - data utility framework (Shlomo and Young 2006). This quantitative evaluation will follow the approach used by Shlomo (2006) and will be used to identify the recommended SDC method(s) for the 2011 Census for all types of outputs although the focus will be on tabular outputs. Examples of this evaluation are provided in Section 5.

An additional stage of research will be timetabled to further develop methods for safeguarding microdata to ensure sufficient protection from disclosure.

It is vital that the development of the 2011 Census UK SDC strategy takes account of the interdependencies which exist with the 2011 Downstream

Processing Schedule and work to design 2011 Census Geography and Outputs. Findings from user consultations will be incorporated into the evaluation process and users will be consulted and updated with research findings and decisions as appropriate.

At this early stage, the final sign-off of the UK SDC strategy has been timetabled for July/August 2009.

3. SDC methods

This paper focuses on SDC methods for protecting census tabular outputs rather than microdata samples although the dependencies between the methods used to protect different outputs will be recognised in the evaluation stage. SDC methods for census tables implemented at Statistical Agencies include both pre-tabular and post-tabular methods or combinations of both. Pre-tabular methods are implemented on the microdata prior to the tabulation and typically include forms of record swapping between a pair of households matching on some control variables (Willenborg and de Waal, 2001). This method has been used for protecting census tables at the US Census Bureau (in 1991 random record swapping was used whereas targeted record swapping was used in 2001) and for the 2001 UK Census. Record swapping can be generalized into a pre-tabular method called PRAM (the Post-Randomization Method) (Gouweleeuw, Kooiman, Willenborg and De Wolf, 1998). This method adds 'noise' to categorical variables by changing values of categories for a small number of records according to a prescribed probability matrix and a stochastic process based on the outcome of a random multinomial draw. Another pre-tabular SDC method is over-imputation. This involves randomly deleting variables in existing records and imputing the variables using the Edit and Imputation System already in use during census downstream processing.

Post-tabular methods are implemented on the entries of the tables after they are computed and typically take the form of random rounding, either on the small cells of the tables or on all entries of the tables. Small cell adjustments (rounding) have been carried out on the census tables by the Australian Bureau of Statistics (ABS) in 2001 and the UK ONS, and full random rounding has been carried out by Statistics Canada and Statistics New Zealand. Controlled rounding is a procedure that uses linear programming techniques to round entries up or down and ensures that all rounded entries add up to the rounded totals. It is available in the SDC software package, Tau Argus, (Hundepool, 2002), however, at present the controlled rounding option is not able to cope with the size, scope and magnitude of census tabular outputs. Other post-tabular methods include cell suppression or some form of random perturbation on the internal cells of the census tables. Cell suppression is not primarily used in the census context because of the large number of tables that need to be consistently suppressed. The ABS have developed a cell perturbation method for their 2006 Census that is designed to potentially alter every cell in every table by a small amount, remove all small cells, always randomise the same table in exactly the same way and ensure additivity.

In addition to the methods described above disclosure risk can also be managed by restricting the design/complexity of the tables, setting geographical thresholds or implementing rules that determine the sparsity of tables, e.g. minimum average cell size.

This paper focuses on record swapping, random rounding and the ABS cell perturbation method in order to demonstrate the evaluation that will be undertaken for SDC methods for 2011 Census.

3.1. Record Swapping

Record swapping involves exchanging geographical variables between randomly selected pairs of households within the Census data. In order to minimise bias pairs of households are determined which match on some control variables, such as a large geographical area and age-sex distribution of the households. Record swapping can be targeted to high-risk households ensuring that households most at risk of disclosure are likely to be swapped.

Record swapping can also be modified to take into account imputation rates, i.e. by only swapping those records with no imputation. In a census context, geography variables are often swapped between households because this results in less edit failures due to the assumption that other census variables are independent of geography. Swapping geographical variables also means that at higher geographical levels and within control strata marginal distributions are preserved.

For this analysis, random record swapping was carried out for a 10 per cent swapping rate. The control variables that were used to determine the pairs of households were the number of persons in the household according to sex and three broad age groups and a "hard-to-count" index of the household based on the 1991 UK Census enumeration. The record swapping was carried out within a large geographical area (Local Authority (LA)) and households were swapped in and out of small geographical areas (Output Areas (OA)). In addition, a targeted record swapping was carried out by defining an additional control variable based on a "flag" for the household that had at least one person in a small cell in one of the census tables under evaluation. Note that on average, about 0.15% of the households selected for swapping were not swapped because no paired record was found for them. In general, those records would have to be swapped outside the large geographical area (LA) but this was not carried out in this analysis.

Table 1: The Main Advantages and Disadvantages of Record Swapping.

Advantages	Disadvantages
○ Consistent totals between tables ○ Tables are additive ○ Protection offered to both tabular outputs and microdata, (further protection may be required for microdata) ○ Marginal distributions preserved at a higher geographical level and within control strata ○ Some protection against disclosure by differencing ○ Given household characteristics, other census variables are likely to be independent of geography therefore less bias will occur ○ Control variables (variables upon which swapped records must match) can be determined according to requirements ○ Swapping rates are flexible ○ Conditional independence assumption means record swapping will not necessarily result in inconsistent and illogical records i.e. less edit failures ○ Flexible table generation is possible provided the disclosure risk resulting from record swapping is acceptable. ○ Only needs to be applied once to the base data (microdata).	○ Effects of perturbation hidden and cannot be measured or accounted for in statistical analysis, i.e. users cannot be provided with a measure of whether a number in a table is the true value ○ Table not visibly perturbed – clear explanations needed to ensure transparency ○ Public perception that no disclosure control has been applied ○ Geographic fields such as workplace are not swapped hence origin-destination tables not protected ○ Method introduces bias

There are also additional advantages and disadvantages associated with both random and targeted record swapping. Random record swapping maintains a higher data utility compared with targeted record swapping at the same swap rate, however, targeted record swapping provides a greater level of protection against disclosure since it targets the risky records. Targeted record swapping results in a greater distortion to tabular distributions (particularly the joint distributions) compared to random record swapping since perturbation is carried out on uniques and outliers rather than at random.

3.2. Random Rounding

The most common post-tabular methods of SDC for census tables are variations of rounding (Shlomo, 2006). Here we focus on full unbiased random rounding to base 3. Let *x* be a cell entry in the table and let $Floor(x)$ be the largest multiple *k* of the base *b* such that $bk < x$ for an entry *x*. In addition, define $res(x) = x - Floor(x)$ .For an unbiased rounding procedure, *x* is rounded up to $(Floor(x) + b)$ with probability $res(x)/b$ and rounded down to $Floor(x)$ with probability $(1-(res(x)/b))$. If *x* is already a multiple of *b*, it remains unchanged. The expected value of the rounded entry is the original entry since:

$$(x - Floor(x)) \times (1 - \frac{res(x)}{b}) + (x - (Floor(x) + b)) \times \frac{res(x)}{b} = 0 .$$

For this analysis the random rounding was semi-controlled in that the overall total for the table was preserved by controlling the rounding process. Marginal totals however were not controlled and were rounded separately from the internal cells and therefore tables are not additive.

Table 2: The Main Advantages and Disadvantages of Semi Controlled Random Rounding (controlling for overall total)

Advantages	Disadvantages
○ Consistent totals between tables if semi-controlled ○ Removes all 1s and 2s from tables, hence removes cases of perceived disclosure as well as actual disclosure ○ Introduces ambiguity into the zeros which exist in the table ○ Provides good protection against disclosure by differencing (although not 100% guarantee) ○ Can be accounted for in statistical analysis - users able to take into account effect of rounding in their analysis and can easily determine between which values the true value must lie ○ Feasible, since rounding is already available in Supercross software ○ Already in use in many National Statistics Institutes; e.g. New Zealand and Canada ○ Easily understood ○ Simple to implement ○ Clear and transparent to users	○ Tables are not additive ○ Rounds all cells, including safe cells ○ Can be unpicked in a proportion of cases due to internal cells of the tables and the margins being rounded independently. Risk of unpicking less when rounding to base 5 which involves greater information loss ○ Does not provide protection to microdata ○ Requires complex auditing to ensure protection ○ If users combine tables of small geographic areas, the total figures will not equal the figures given for larger areas ○ Where tables are combined to create user-defined areas, cells could be significantly altered by the rounding process ○ If flexible table generation is made available users could be provided with different versions of the same table according to the random numbers used to round the cell entries. Hence risk of disclosure may be increased as a result. ○ Time-consuming as must be applied to every table

3.3. ABS Cell Perturbation

For the protection of their 2006 Census outputs, the ABS has conducted research into a new cell perturbation algorithm (Fraser and Wooton 2006). In the past they have released static tables of data however flexible table generation will be used for 2006. This will enable users to design and populate their own tables. The new perturbation algorithm is designed to protect these tables by potentially altering every cell in every table by a small amount. In doing so it adds sufficient 'noise' to each cell so that by differencing, users would end up with more noise than real data. The algorithm always randomises the same table in exactly the same way. It also preserves higher level totals between tables with common geographies.

The SDC algorithm involves two stages; the first adds the perturbations to the cell values and the second stage restores additivity to the table.

Perturbation Stage

1. A value m is predetermined defining the range of the perturbation distribution.
2. Each record in the microdata is assigned an rkey or record key. The rkey is a value drawn at random from the discrete uniform distribution [0, m-1].
3. Each table is then considered independently. The rkeys relating to the records making up each cell in the table are combined to give a cell key or ckey as follows: ckey = $\mathrm{mod}\left|\sum(rkey), m\right|$

The use of the mod (remainder) function means that the distribution of the ckeys is also a discrete uniform on [0, m-1].
4. A look-up table is defined with original cell values on the rows and ckeys on the columns. Thus the lookup table will have m-1 columns and the maximum cell value in the original table will correspond to the number of rows.
5. The look-up table provides the perturbation value relating to each cell determined by the original cell value (row) and the ckey (column).
6. This perturbation is then added to the original cell value.

Additivity Stage

After the perturbation stage, the same cell in different tables is consistent (has the same perturbation added). However the tables do not add up. Additivity is restored using an iterative algorithm which visits single and pairs of cells adding -1, 0, +1 at each iteration stopping when all rows and columns add up.

It does this at the same time as minimising the overall difference between the additive and original table.

The look-up table can be designed according to the specification of the statistical agency. For example the first row of the look-up table could be specified with all zeros which means all original cell values of zero have zero perturbation added, moreover, it would also be possible to design the look-up table such that all ones and twos are removed from tabular output (as they plan to do for the ABS 2006 Census). In fact, the look-up table could also be designed to mimic the effects of other SDC procedures such as random rounding.

For this analysis only the perturbation stage of the ABS method has been implemented since the code for the additivity stage is not currently available. The following look-up table was used:

Original Cell value	**Perturbation to be drawn from the following distribution (using the cell key)**
0	Remain as zeros
1	Normal distribution with mean 0 and variance 2 truncated at -1 and +5
2	Normal distribution with mean 0 and variance 2 truncated at -2 and +5
3	Normal distribution with mean 0 and variance 2 truncated at -3 and +5
4	Normal distribution with mean 0 and variance 2 truncated at -4 and +5
5+	Normal distribution with mean 0 and variance 2 truncated at -5 and +5

The ABS cell perturbation method is a slightly more informed post-tabular method of disclosure control compared to random rounding since it utilises microdata information during the perturbation stage.

Table 3: The Main Advantages and Disadvantages of the ABS Cell Perturbation method

Advantages	Disadvantages
○ Tables are consistent ○ Provides protection for flexible tables ○ Depending on the design of the look-up table, the method can perturb distributions that are approximately unbiased with small variances ○ Efficient - allegedly has a quick run time ○ Able to produce perturbations for large high dimensional hierarchical or cross classified tables ○ Protects against differencing ○ Method is extremely flexible; look-up table can be specifically designed to suit needs and different look-up tables could potentially be used for different tables. Moreover the look-up table could be designed to mimic random rounding for example.	○ Tables not additive (additivity module is not applied here) ○ Once additivity is applied, consistency is lost ○ Needs to be applied to each table separately ○ Public perception that no disclosure control has been applied (unless incorporated into look-up table) ○ No protection for microdata ○ Method less transparent than others e.g. rounding ○ Depends on the appropriate choice of look-up table which may not be suitable for all tables (i.e. sparse) ○ Statistical effects are highly dependent on the choice of look-up table

4. Short-listing SDC methods for quantitative evaluation using a Disclosure Risk - Data Utility Framework

The quantitative risk-utility framework being used to evaluate the SDC methods (see Section 5) for the 2011 Census is not sufficient on its own. Many SDC methods have qualities which cannot be accounted for quantitatively and thus qualitative advantages and disadvantages of SDC methods must also be addressed. These in combination with the results from the risk-utility assessment will inform the recommended approach to SDC in 2011. The qualitative characteristics that will be considered include:

- overall practicality and feasibility of implementation
- interaction between different types of output

- user acceptance of chosen methods (whether additivity and/or consistency are achieved)
- whether the methods are suitable for flexible table generation
- whether the methods offer protection to microdata as well as tabular data
- protection afforded in terms of perceived disclosure

The advantages and disadvantages of each method are therefore being addressed and compared to discount SDC methods which will not suit the disclosure control requirements for 2011 Census outputs. Discounted methods will be excluded from further consideration and a short-list of SDC methods will be assessed using the risk-utility framework. The final short-list of SDC methods is due to be agreed in October 2007. Example results evaluating the risk and utility of record swapping, random rounding and the ABS cell perturbation method on tabular outputs are presented in Section 5.

The advantages and disadvantages of the three SDC methods focused on in this paper were addressed in Section 3. All methods are feasible for the 2011 Census but all have their own limitations. The pre-tabular method of record swapping has the advantage that it can be used to protect microdata as well, whereas random rounding and the ABS cell perturbation method protect only tabular outputs. The ABS cell perturbation method and record swapping provide better protection against disclosure by differencing which will be important if flexible user defined tabular outputs are to be made available in 2011. Record swapping ensures that marginal distributions are preserved at a higher geographical level and within control strata and results in additive and consistent tables. However, random record swapping can result in a high proportion of risky cells left unaltered. Targeting the risky cells reduces the risk of disclosure at the same record swapping level but causes greater distortion to tabular distributions. The ABS method (when applying both stages) results in additive tables however tables representing the same population subgroups may not end up with consistent totals. Random rounding has the opposite effect of maintaining consistent totals between tables but destroying additivity within the table. The method of random rounding is the most transparent of the three methods discussed here and it can be easily accounted for in statistical analyses however it does not provide as high a level of protection against differencing as cell perturbation and record swapping.

5. Quantitative Analysis of Proposed SDC methods

As described above once agreed the short-list of SDC methods will be evaluated quantitatively focusing on an assessment of the statistical impact of the method on data utility and disclosure risk. A software package (Shlomo and Young, 2006) developed to calculate a variety of information loss metrics (by comparing the protected data with the original pre-disclosure controlled data) will be used for this analysis. Here we present a selection of the information loss measures and one risk measure described in Shlomo and Young, 2006 and use them to compare the three SDC methods for two example tables. It should be noted that these are preliminary results and are included as an illustration of the analysis that will be undertaken. A more thorough analysis investigating further methods using a wide range of tables, varying parameters (e.g. swapping rates, rounding base, look-up table), and further disclosure risk and information loss measures will be required for the final analysis.

5.1. Data

The effects of the SDC methods will be considered for two tables at two different levels of geography, Output Area (OA) and ward level for an Estimation Area in England relating to Southampton, Eastleigh and Test Valley. Geography is represented as rows in the table and the other variables span the columns. Table 4 describes the structure of the two tables.

Table 4: Example tables

	Variables and Number of Categories	Number of Persons in the Table	Number of Internal Cells	Average Cell Size	Number of Zeros	Number of Small Cells
Table A	Religion (9) Age-Sex (6) OA (1,487)	437,744	80,298	5.45	47,433 (59.1%)	10,137 (12.6%)
Table B	Economic Activity (9) Sex (2) Long term illness (2) Ward (70)	317,064	25,250	125.82	427 (16.94%)	226 (8.97%)

These two tables were selected to study whether the methods have varying effects over different levels of geography.

5.2. Risk and Utility Measures

Disclosure Risk

Let R_i represent the record *i*, I the indicator function having a value of 1 if true and 0 if false, C_1 the set of cells with a value of 1, C_2 the set of cells with a value of 2, $|C_1 \cup C_2|$ the number of cells with a value of 1 or 2. The disclosure risk measure can be interpreted as the percentage of records in small cells that have not been perturbed:

$$DR = \frac{\sum_{i \in C_1 \cup C_2} I(R_i \ not \ \ perturbed\, or \ \ imputed)}{|C_1 \cup C_2|}$$

Distance Metrics on Internal Cells of the Tables

Distance metrics are used to measure distortion to distributions. A distance metric is calculated for each row in the table and then the overall average across all of the rows is taken as the information loss measure. This format is used since the rows in census tables generally represent a geographical area whereas the columns define the categories of a specific table, such as sex$\times$age group$\times$economic activity. When comparing the average distance metric across rows, we need to take into account the level of dispersion as expressed by the standard error (confidence interval).

Let D^k represent a row *k* of table *D* and let $D^k(c)$ be the cell frequency *c* in the table. Let n_r be the number of rows in the table. *Pert* refers to the disclosure-protected table and *orig* to the original table. The distance metrics are:

Hellinger's Distance:

$$HD(D_{pert}, D_{orig}) = \frac{1}{n_r}\sum_{k=1}^{n_r}\sqrt{\sum_{c\in k}\frac{1}{2}\left(\sqrt{D_{pert}^k(c)} - \sqrt{D_{orig}^k(c)}\right)^2}$$

Relative Absolute Distance:

$$RAD(D_{pert}, D_{orig}) = \frac{1}{n_r}\sum_{k=1}^{n_r}\sum_{c\in k}\frac{|D_{pert}^k(c) - D_{orig}^k(c)|}{D_{orig}^k(c)}$$

Average Absolute Distance per Cell:

$$AAD(D_{pert}, D_{orig}) = \frac{1}{n_r}\sum_{k=1}^{n_r}\frac{\sum_{c\in k}|D_{pert}^k(c) - D_{orig}^k(c)|}{n_k} \qquad \text{where} \qquad n_k = \sum_c I(c\in k)$$

the number of cells in the k^{th} row.

The standard errors are calculated as follows (for example, the *AAD* metric):

$$\frac{1}{n_r - 1}\sum_{k=1}^{n_r}\left(AAD(D_{pert}^k, D_{orig}^k) - AAD(D_{pert}, D_{orig})\right)^2$$

where $AAD(D^k_{pert}, D^k_{orig}) = \frac{\sum_{c \in k} | D^k_{pert}(c) - D^k_{orig}(c) |}{n_k}$

These distance metrics can also be calculated for sub-totals and totals of the tables.

Variance of Cell Counts

An information loss measure can be calculated to measure the impact on the variance of the estimates. The variance of the counts is examined across the rows before and after the SDC methods as follows:

For each row *k*, we calculate: $V(D^k_{orig}) = \frac{1}{n_k - 1} \sum_{c \in k} (D^k_{orig}(c) - \overline{D}^k_{orig})^2$ where $\overline{D}^k_{orig} = \frac{\sum_{c \in k} D^k_{orig}(c)}{n_k}$ and $n_k = \sum_c I(c \in k)$ the number of cells in the k^{th} row. Next we calculate the ratio for each row:

$$VR(D^k_{pert}, D^k_{orig}) = \frac{V(D^k_{pert})}{V(D^k_{orig})}$$

Change to Rank Orderings

Changes to the underlying ordering of cell counts (impact on rank correlation) within the table can be studied. The original counts are sorted according to their size and deciles (10 equal groupings) $v^{orig}(c)$ are defined. This is repeated for the perturbed cell counts which are sorted according to both their size and the original order in order to maintain consistency for the tied variables. Deciles $v^{pert}(c)$ are then defined for the perturbed variable after the sort. The information loss measure is the percent of cells that have changed deciles. The measure is calculated across different categories in the table e.g. table columns, and then an overall average is the final measure:

$$RC = \frac{100 \times \sum_{c \in k} I(v_k^{orig} \neq v_k^{pert})}{n_k}$$

where I is the indicator function and is 1 if the statement is true and 0 otherwise, k is a column in the table and n_k is the number of cells in that column.

5.3. Results

Table 5: Disclosure risk measures

Probability that a record in a small cell has not been perturbed	Benchmark Totals Random Rounding	10% Random Swap	10% Targeted Swap	ABS Cell Perturbation
Risk (ward)	0	0.619	0.509	0.142
Risk (OA)	0	0.651	0.506	0.188

Table 5 displays the risk measure for the different SDC methods for the two tables. Since the risk measure focuses on small cells the risk is 0 for random rounding. This does not mean that risk has been entirely eliminated since there are other risk measures that can and will be considered. The risk is far smaller for the ABS method in comparison to record swapping because there is a higher probability that a small cell would receive a non-zero perturbation. The targeted swap focuses on perturbing small cells and hence the risk is less than for the random swapping method.

Table 6: Distance metrics, OA level

OA – Distance Metrics	Benchmark Totals Random Rounding	10% Random Swap	10% Targeted Swap	ABS Cell Perturbation
Hellingers' Distance	1.6616 (0.0246)	1.2875 (0.0249)	1.6027 (0.0265)	1.7388 (0.0228)
Relative Absolute Distance	8.8507 (0.2262)	4.2542 (0.1149)	5.2674 (0.1289)	11.2215 (0.2767)
Absolute Average Distance	0.4016 (0.0061)	0.4870 (0.0100)	0.5275 (0.0093)	0.6217 (0.0088)

Table 7: Distance metrics, ward level

Ward – Distance Metrics	Benchmark Totals Random Rounding	10% Random Swap	10% Targeted Swap	ABS Cell Perturbation
Hellingers' Distance	1.095 (0.1193)	1.2389 (0.1211)	1.3811 (0.1177)	1.2167 (0.1113)
Relative Absolute Distance	4.2778 (0.636)	3.4715 (0.4775)	4.0725 (0.5093)	6.1881 (1.0078)
Absolute Average Distance	0.7301 (0.0667)	3.6897 (0.6579)	3.7048 (0.5732)	1.300 (0.1235)

At the OA level the ABS method performs the worst for all three distance metrics because there is a high probability that small cells are perturbed using our specified look-up table (see section 3.3) and at OA level the table is particularly sparse (see table 4). In all cases the targeted swap distorts the distributions in the table more than the random swap as expected. The best method (in terms of distortions to distributions) is either the random rounding or the random swap in this case, but the results in general would depend on the table and the distance metric considered. The standard errors for each measure are displayed in brackets.

Table 8: Distance metrics, marginal totals

Change in marginal totals (relative difference) BY SUBGROUP	Benchmark Totals Random Rounding	10% Random Swap	10% Targeted Swap	ABS Cell Perturbation
Sex by age-group (over all OAs) (table A)	0.003	0	0	1.078
Religion (over all OAs) (table A)	3.59	0	0	3.468
Sex by long term illness (over all wards) (table B)	0.006	0	0	0.007
Economic activity (over all wards) (table B)	0.009	0	0	0.024

Table 8 shows how the SDC methods impact on the marginal totals rather than the internal cells. The random and targeted record swapping result in no change to the marginal distributions of the tables. This result occurs because by definition record swapping maintains the marginal distributions at levels above Local Authority District and the marginals here represent subgroups of the Estimation Area. The marginal totals representing Estimation Area by religion are affected by the greatest change in relative difference when performing the ABS method and random rounding. This result is likely to be caused by the uneven distribution of marginal counts across religions resulting in a greater number of small cells which are affected to a greater extent by these methods. The other marginal totals considered are affected to a lesser degree because the marginal counts will be more evenly distributed across all variable categories and hence the perturbations applied to the marginal cells are small relative to the marginal count.

Table 9: Variance of cell counts

Average variance ratio over all rows	Benchmark Totals Random Rounding	10% Random Swap	10% Targeted Swap	ABS Cell Perturbation
OA	1.0216	0.9944	0.9985	1.0326
Ward	1.2255	1.0263	1.0079	1.3389

Table 9 indicates that the impact of the SDC methods on the variability of cell counts by row, is not significant since no firm patterns can be seen. At ward level, the effect of the SDC method on the variance appears to be more noticeable with the ABS method and random rounding both increasing the variance. The results from the ABS cell perturbation method are dependent on the look-up table and can vary if the perturbation distributions are changed.

Table 10: Change to Rank Orderings

Cells moved into different percentile (groups of 10)	Benchmark Totals Random Rounding	10% Random Swap	10% Targeted Swap	ABS Cell Perturbation
OA	7%	26%	34%	20%
Ward	0%	2%	3%	2%

This test shows how swapping and to some extent cell perturbation distorts the underlying patterns in the data by changing the rank order of cells. At OA level there is a lot of distortion because more than 70% of cells have values less than 3 whereas at ward level there is greater variation in the cell counts so the SDC methods have less of an impact. Since there is a limit on how much the cell values can change with rounding, the change in ordering is much smaller than with the other three methods.

6. Conclusions

This paper has described the approach that will be adopted to develop the SDC strategy for all 2011 Census outputs. A review of past work (particularly undertaken for 2001) has been conducted and is being used to inform further stages of the project. A high level review of SDC methods has been conducted and will be used to develop a shortlist of methods for further evaluation. Examples from this high level review and a quantitative evaluation (measuring risk and information loss) have been presented for three different SDC methods; record swapping, random rounding and a cell perturbation method. These preliminary results are included as an illustration of the final more detailed evaluation that will be undertaken. It is recognised that developing a 2011 UK SDC strategy which satisfies all user requirements whilst maintaining a high level of data utility is likely to be an unachievable task hence compromises will need to be made. The final recommended approach to SDC for 2011 Census will be informed by both quantitative and qualitative evaluation and the trade-offs between the different methods will need to be communicated to users.

References

Fraser, B. and Wooton, J. (2006) A proposed method for confidentialising tabular output to protect against differencing, Internal report, Data Access and Confidentiality Methodology Unit, ABS.

Gouweleeuw, J., P. Kooiman, L.C.R.J. Willenborg and P.P. De Wolf (1998). Post Randomisation for Statistical Disclosure Control: Theory and Implementation. *Journal of Official Statistics*, 14, pp. 463-478.

Hundepool, A. (2002). The CASC Project. In Domingo-Ferrer, J. (eds.): *Inference Control in Statistical Databases: From Theory to Practice. Lecture Notes in Computer Science*, Vol. 2316. Springer-Verlag.

Shlomo, N. and Young, C. (2006) Statistical Disclosure Control Methods through a Risk - Utility Framework: Proceedings of the Privacy in Statistical Databases CENEX-SDC Project International Conference, Rome, 13-15 Dec 2006.

Shlomo, N. (2006) Review of Statistical Disclosure Control Methods for Census Frequency Tables. *Survey Methodology Bulletin, 57, Office for National Statistics.*

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