

Changing industrial classification to SIC(2007) at the UK Office for National Statistics

Paul A. Smith, Statistical Sciences Research Institute, University of Southampton, Highfield, Southampton, SO17 1BJ, UK. p.a.smith@soton.ac.uk

Gareth G. James, Office for National Statistics, Cardiff Road, Newport, NP10 8XG, UK. gareth.james@ons.gov.uk

Abstract

As part of the changes to industrial classifications following the United Nations' revision to the International Standard Industrial Classification, ISIC rev. 4, the UK moved to its version of a new classification between 2007 and 2011. We describe the processes involved in changing an industrial classification, including model-based adjustment methods and changes to survey designs and operations. We discuss the quality of the approaches used for different time periods in the same series, and the ways in which consistent time series are produced for users of economic statistics. We provide some general evaluation of the changeover, and guidance on the best approaches to follow when updating classifications.

Key words: conversion matrix, overlap control, quality assessment, classification change

1. Introduction

Changes in standard classifications, such as that used for economic activity, usually occur at reasonably regular, though infrequent intervals. They are required to ensure classifications remain up-to-date and relevant, but when they occur consume large amounts of resource and can create discontinuities in time series. Therefore they need to be carefully managed (MacDonald 1995). A change in the classification system has effects in many of the stages of the statistical production process (as codified by, for example, the Generic Statistical Business Process Model, UNECE 2013). In this paper we consider the most recent classification change implemented by the UK's Office for National Statistics (ONS) as a model to demonstrate some techniques and derive some general guidance.

Industrial classifications are hierarchical. They are harmonised internationally down to a particular level of detail, with

- a world framework, the International Standard Industrial Classification (ISIC, the current version is Rev. 4), produced by the United Nations;
- a regional implementation. In the European Union (EU), Eurostat (the statistical office of the EU) currently uses NACE¹ Rev. 2 (a 4-digit hierarchical classification), which is consistent with ISIC Rev. 4 at the 2-digit level; and
- national implementations, such as the UK's Standard Industrial Classification (SIC). Within the EU, Member States have the option to enhance NACE for their own purposes, as long as the national classification corresponds with the harmonised classification to at least a given level of detail (usually 4-digit industry for NACE). The national implementations are also in an appropriate native language with support for dictionaries and look-ups to make the classification usable (see for example Beekman 1992).

Harmonised classifications within the UK have been an important way to make statistics from different government departments comparable since the first UK SIC in 1947, and (with a few exceptions) have

¹ Nomenclature statistique des Activités économiques dans la Communauté Européenne

been used consistently throughout the UK's statistical system. The UK's current classification is known as SIC(2007) (the '2007' denoting the year of publication). The classification forms the basis of statistical outputs, and provides a framework for combining estimates from different surveys in derived statistics such as the National Accounts. The classification codes also provide one dimension for stratification in business survey designs.

Changes in the classification in the UK have occurred about every 10-15 years (see Smith and Penneck 2009 for an overview of industrial classifications used in UK statistics since 1907). The most recent large changes in the UK's classification were from SIC(80) to SIC(92) – the implementation occurring in the mid-1990s – and from SIC(2003) to SIC(2007). (The change from SIC(92) to SIC(2003) was not large, and only involved minor changes at the most detailed level.)

The most recent change proved particularly challenging, as the new classification (SIC(2007)) contained more detail than its predecessors, and included a number of industry restructures. Examples include the separate identification of Repair and Maintenance, a new section on Water Supply and related activities, a new section on Information and Communication, the move of Retail Sale of Automotive Fuel from Motor Trades to Retail, and the move of some publishing and printing activities from Manufacturing to Services. Table 1 shows the changing numbers of categories at different levels of the classification. The detail of the 5-digit level was agreed in the UK following a series of user consultations called 'Operation 2007' (Hughes 2008) run by a cross-government group (including ONS). Sub-divisions of the 4-digit codes were agreed only where there was user demand, and where it would be practical and meaningful to distinguish between the proposed sub-categories. It can be seen that fewer disaggregations of 4-digit codes to 5-digit codes were accepted under SIC(2007) than under SIC(2003), probably reflecting the greater detail already present in the new 4-digit codes, but also reflecting a reluctance to create 5-digit codes unless there was a strong case that they were both necessary and practicable. That cross-government group also co-ordinated the implementation of SIC(2007) across the Government Statistical Service in the UK following the timetable set by Eurostat for all EU Member States.

Table 1: Comparison of detail (number of categories) between SIC(2003) and SIC(2007):

	SIC(2003)	SIC(2007)
Section (letter)	17	21
Division (2-digit)	62	88
Group (3-digit)	225	272
Class (4-digit)	514	615
Sub-class (5-digit)	699	728

The change in industrial classification was also aligned with a change to the European product classification to CPA 2008 (Classification of Products by Activity), and this particularly affected the National Accounts, which use both product and industry estimates in producing balanced measures of the national economy. The scale of these changes called for review of most aspects of the methodology used in business surveys in ONS to allow the production of estimates on the new classification.

Planning for the project and work to get agreement on the classification breakdowns using the fifth digits started in the early 2000s, so that the new classification, with its index and accompanying notes

for coding could be published in January 2007. Thereafter, the work to implement the new classification in surveys and outputs really began; the timing of changes is outlined in Fig. 1. All organisations on the Inter-Departmental Business Register (IDBR) – ONS’s sampling frame for most business surveys – were dual-coded during 2007, so that by January 2008 all had a SIC(2003) and a SIC(2007) code. The annual, structural surveys collecting information about activity in 2008 were the first to be sampled on the new classification, with selections taking place from the IDBR towards the end of 2008. The first outputs on the new classification, from short-period (monthly and quarterly) surveys, were reported to Eurostat from the start of 2009. The old classification was still used for sample selection in monthly and quarterly surveys until the start of 2010, however, and inputs to National Accounts were retained on the SIC(2003) basis until the publication of the Blue Book (the UK’s main National Accounts publication) in September 2011, after which time all reporting of economic output from ONS has used the new classification.

Timetable of implementation of SIC(2007) at ONS:

- Pre-2007 Agreement of fifth-digits in SIC codes
- 2007 Addition of SIC(2007) codes to the IDBR, available from January 2008
- Late 2008 Selection of first annual survey samples using SIC(2007) stratification
- Jan/Q1 2009 Delivery of short-period statistics to Eurostat on SIC(2007)
- Jan/Q1 2010 Short-period surveys stratified by SIC(2007); outputs still required in National Accounts on SIC(2003)
- Sep 2011 National Accounts move to SIC(2007) with publication of the Blue Book

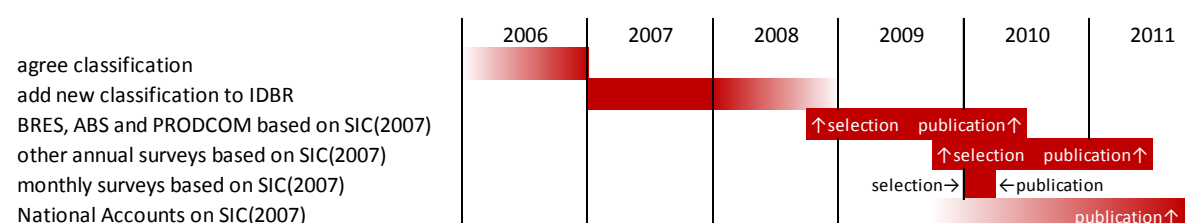


Fig. 1: Diagrammatic representation of transition to SIC(2007) in ONS. Only the first instance of each survey with design and sample selection updated to the new classification is represented. Fading shows continuing work leading up to or following an implementation. IDBR = Inter-Departmental Business Register, BRES = Business Register and Employment Survey, ABS = Annual Business Survey and PRODCOM = Products of the European Community survey.

The prolonged timetable has created some interesting challenges, particularly around the publication of two simultaneous sets of outputs. In some cases, publication has been retrospective, with backcasting (see section 4.1) used to estimate how historical outputs would have looked on the new classification. In other cases, publication was (almost) concurrent, with more recent periods being reported on both classifications.

In this paper we review the implementation of SIC(2007), and the developments and changes required in the methodology of surveys, particularly with respect to consistency of estimates across time. Section 2 covers the changes to the IDBR, and section 3 describes changes to the sample designs of business surveys largely (but not wholly) to accommodate the new classification. Section 4 covers different methods of estimating on the new classification and includes comments on some of the

effects on the National Accounts. Section 5 discusses the impacts on quality of estimates using the standard European dimensions of output quality (Eurostat 2015a), and the paper concludes with some general comments about the implementation, and includes some lessons learned and suggestions for future implementations.

2. The business register

The IDBR is maintained by ONS, and is the main sampling frame used for official business surveys in the UK. It contains information on about 2.7m sites (known as Local Units (LU) – shops, factories, offices, and so on), which are grouped into about 2.1m Reporting Units (RUs), which are the sampling units used in most ONS business surveys (within this article we will use ‘business’ synonymously with RU for this reason). One or more RUs form an Enterprise, which is the smallest autonomous business structure on the IDBR. In the majority of cases, an Enterprise has one RU and one LU, and the three units are indistinguishable; even large enterprises are normally one RU, although a relatively small number are split into multiple RUs to facilitate data collection (for more information see Smith, Pont & Jones (2003)). The distribution of Enterprise sizes is very skewed, and large proportions of activity and employment take place in Enterprises composed of multiple LUs.

The IDBR contains a range of variables for each type of unit, the values of which are obtained from various data sources, with a hierarchy specifying which sources are preferred. One of the most important variables ascribed to the unit is the SIC code, which represents its principal economic activity, and rules exist to determine the SIC code for Enterprises and RUs based on the codes of the LUs (for a simple summary see Smith 2013, section 5.2.1.2 and Box 5.1). A stable SIC code is stored for sampling purposes, and this code is used throughout a calendar year for monthly surveys, for example. Then reclassification effects can be saved-up and handled together at a fixed point, usually the year-end. A ‘live’ classification variable stores the most up-to-date SIC code. Updating the IDBR includes copying the current live codes to the stable codes once a year. Additional variables were added to the IDBR in preparation for the change to SIC(2007), so that (stable and live) codes for both SIC(2003) and SIC(2007) could be stored for each unit.

One source of information about new businesses on the IDBR is Value Added Tax (VAT) records from HM Revenue & Customs (HMRC). When a business is formed, it must describe its activity when it registers for VAT, and these descriptions are coded by HMRC using automated coding software (called ACTR – Automatic Coding by Text Recognition). The software and coding dictionary are harmonised across government departments so that there is consistency in the way in which classification codes are assigned. Both the original descriptions and the assigned codes are passed to the IDBR and stored, so the descriptions are available for recoding to deal with future classification updates. For businesses already on the IDBR, another source of information is ONS’s annual Business Register & Employment Survey (BRES), which updates register information and collects employment variables; the register-updating part has previously been administered separately. Not all businesses are surveyed in BRES, but all large businesses are included every year, and medium-sized ones once every third or fourth year. The smallest businesses are included in BRES only with small sampling fractions, and primarily for estimation of employment, so classification information for such businesses on the IDBR is derived mainly from the administrative data sources. Businesses selected for BRES are asked to supply a written description of the principal economic activity at each site (LU), and these are then coded to the SIC using the ACTR automated coder (Williams 2006). (Previous descriptions of economic activity are prefilled on the questionnaire for existing sites to make the respondents’ task less onerous – only changes need be notified.)

To facilitate coding to the new classification, a knowledge base for the new classification was developed by constructing a list of full descriptions of economic activity. These were edited by a classification specialist for consistency, and used as the basis for the automated coder to code business descriptions to SIC(2007). This enabled dual-coding of most businesses on the IDBR (1.7m of the 2.7m local units, for which original business descriptions were available and usable). For the remaining units – those without a description, or where the description was insufficient – SIC(2007) codes were assigned probabilistically, based upon the distributions observed in those units that had been dual-coded. Of course, any SIC(2003) codes that mapped entirely to just one SIC(2007) did not need to go through this process. Based on reclassifications of businesses, 462 of the 5-digit SIC(2003) codes matched exactly to a 5-digit SIC(2007) code; 229 matched to multiple codes, and the 8 remaining codes did not have any businesses classified to them. Of SIC(2007) 5-digit codes, 503 matched one code, 205 matched multiple codes, and 20 had no businesses. Some checking and manual intervention was necessary to avoid problems arising from the descriptions (and to feed these back as improvements to the automated coder), but the IDBR was fully dual-coded before January 2008. However, feedback on codes from businesses meant that further cleaning went on, and some further change in the register was quite evident, particularly in the early months of 2008. There were some noticeable changes in the proportion of SIC(2003) industries contributing to SIC(2007) industries early in the year (Fig. 2), but by mid-2008 these proportions had largely stabilised. The same sort of pattern is shown for local units in aggregate in Anon. (2012).

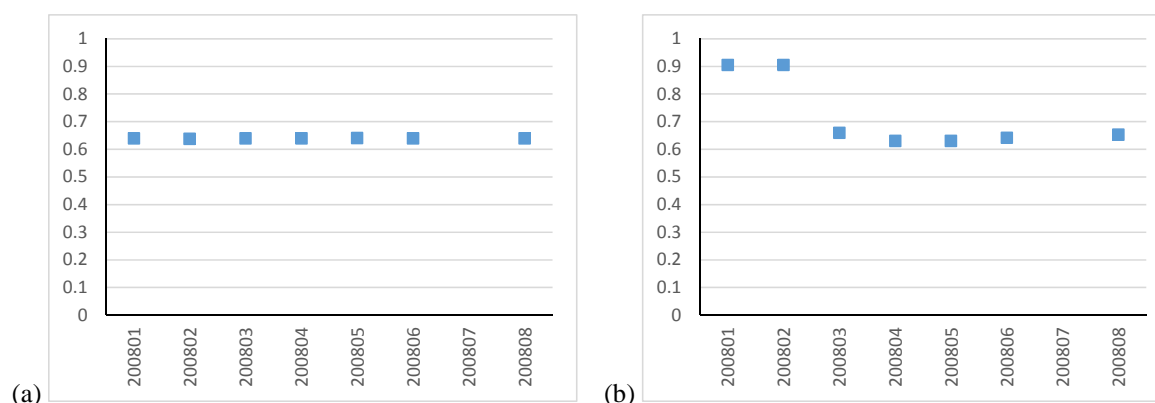


Fig. 2: Proportions of undeclared turnover from an original SIC(2003)-based classification code contributing to a new SIC(2007) code for the early months of 2008, based on dual-coded microdata from the IDBR. Here the classifications are aggregates of industries used for UK Supply and Use Tables (SUT); the identities of the classifications are withheld for disclosure control reasons. Note that there is no data point for July 2008. (a) A typical example with little variation; (b) some industries had noticeable changes in early 2008 before stabilising.

The new SIC(2007) codes were used for sample selection purposes (as stratification variables) for the first time in the autumn of 2008, when samples for three annual surveys were selected: BRES, the Annual Business Survey (ABS) and PRODCOM (PRODucts of the European COMMunity; the UK implementation is headed 'UK manufacturers sales by product').

3. Survey redesign

The change in SIC meant that the business surveys needed to be redesigned. Almost all of ONS's business surveys are stratified by a cross-classification of industry (SIC) and size, with size usually defined as employment-based sizebands. Therefore changing the surveys' sample designs to the new codes, together with re-allocating the (fixed) sample size was the minimum requirement, but in many cases the opportunity was taken to introduce other improvements at the same time.

An example of wider survey redesign was the introduction of the Monthly Business Survey (MBS) at ONS, of which an overview is given here, and a more detailed account can be found in Taylor *et al.* (2011). MBS was introduced as a replacement for a number of short-period surveys, each covering different parts of the economy (namely production, and parts of the services sector). Improvements to the surveys had been planned for some time, but the need to introduce the new classification made this work a higher priority. (Here we see one of the incidental effects of changing a classification, that it often prompts implementation of further changes) The new design covered much more of the economy as a single survey: (construction and agriculture were not included initially, although construction surveys have since been added to the MBS family). A combined survey has advantages when businesses change classification, particularly for moves between the production and services sectors. Such a difference is not now apparent to survey respondents from the questionnaires, which previously might have seen the same business receive a 'Manufacturing' questionnaire on one survey and a 'Services' questionnaire on another. Thus, a combined approach removes the source of some queries from businesses, and it also facilitates easier transfer of data between processing systems. In addition to changes in stratification and sample allocation (detailed further, below), improvements were made to the coverage, questionnaire design, and editing and imputation methodology, with the aim of introducing greater standardisation and therefore more coherence. This approach followed general practice, which often sees several updates to a survey implemented simultaneously. This approach can be convenient for users, as they need to deal with only one discontinuity, but on the other hand, some users want to know how much of any discontinuity is attributable to each development, and these may not be separately estimable when changes are introduced together.

The first task in designing the new MBS was to decide upon appropriate SIC(2007) codes to define the industry strata. This was not a straightforward task, as different levels of the hierarchy and different groupings of codes can be used, as was the case across its predecessor surveys. Consultations were carried out with the main customer of the surveys, National Accounts, so that the strata would be consistent with their requirements, which would be based around the industrial groupings used in the supply and use tables (SUTs). In 2009, those were as-yet undefined under SIC(2007), which required some work to be conducted quickly to establish the likely groupings to be used in National Accounts from 2011 onwards when SIC(2007) would be introduced. The results of this exercise were used as the basis for the MBS stratification. (Further details of the 114 input-output groups, which would be defined for use in the SUTs and balancing in the National Accounts under the new classification can be found in Drew & Dunn 2011.) Many of the SIC(2007) SUT industry groups were formed around the 2- or 3-digit SIC code level, and meant that SIC(2007)-stratification of the MBS could be at a more aggregated level than in the previous designs based on SIC(2003) – only around 150 industries instead of 300 (Taylor *et al.* 2011). A similar approach, with the stratification detail being driven by the principal output requirements, was taken when redesigning other surveys. Use of a broader stratification gave more flexibility to produce an efficient sample allocation without increasing the fixed sample size. However, it should be noted that this also increased some estimator variances (particularly for detailed industry classifications that no longer formed strata), as domain estimation was needed to get the full SIC(2007) industry detail, and also for the SIC(2003)-based estimates that were still required by National Accounts until their move to SIC(2007)).

With industry stratum boundaries defined, a review was conducted as to whether the sizebands, which separate the industry strata into sampling strata, should be retained or updated. Some of this work was necessary, as various combinations of sizebands were used across the predecessor surveys, but the opportunity was taken to conduct a wider review. In many cases, updated sizeband boundaries were introduced to improve the design (even in industries which mapped one-to-one from SIC(2003) to SIC(2007)), although the number of sizebands was kept the same (changing this would have necessitated changes to the processing systems that could not have been easily accommodated at the same time).

Samples for business surveys at ONS are usually allocated using principles of Neyman allocation (Cochran 1977, p.98-99), although adapted to account for various practical considerations and precision requirements on some lower-level outputs. The process requires information on the number of businesses in each sampling stratum (which is obtained from the IDBR) and a measure of the variability in the responses, given the estimator used within each sampling stratum. The standard approach for estimating the population variance is to use information from the responses in a particular stratum in a previous period:

$$\hat{v}_h = \frac{1}{n_h - 1} \sum_{i \in h} (y_i - \hat{y}_i)^2 \quad (1)$$

where y_i is the previous period response for business i , n_h is the number of responses in stratum h in that period, and \hat{y}_i is an estimate of the mean for unit i appropriate for the estimation model in use in the survey (eg the overall mean \bar{y} for expansion estimation, or βx_i for ratio estimation with auxiliary variable x); the same basic approach applies when h represents old strata or when it represents the new strata, though with the change in classifications, past responses for use in the new SIC(2007) strata sometimes come from businesses in several different SIC(2003)-based strata, and these needed to be weighted appropriately. Standard options for a weighted variance include:

$$\hat{v}_g = \frac{1}{\left[\sum_{i \in g} d_i \right] - 1} \sum_{i \in g} d_i (y_i - \tilde{y})^2 \quad (2)$$

and

$$\hat{v}_g = \frac{1}{(n_g - 1)\bar{d}} \sum_{i \in g} d_i (y_i - \tilde{y})^2 \quad (3)$$

where d_i is the inverse of the sampling probability for unit i , $\bar{d} = \frac{1}{n_g} \sum_{i \in g} d_i$, g indexes the new (SIC(2007)

× sizeband) strata, and $\tilde{y} = \sum_{i \in g} d_i \hat{y}_i / \sum_{i \in g} d_i$ is a weighted estimate of the mean of the y appropriate for

the estimation model used in the survey. Only (3) collapses to (1) when $d_i = k \forall i$. Even what may appear a relatively simple case, for example of two old (SIC(2003)) industry strata being mapped entirely into one new (SIC(2007)) industry stratum may require this treatment if different employment sizebands or sampling fractions were used in the old industries or between the old and new.

A similar pooling, but based on an assumption that variances within a group of strata are equal, is presented in van den Brakel (2010, eqn. (30)); this might be a better approach where sample sizes in group g are small, which may make variance estimates unstable. In ONS business surveys, groups g based on SIC(2007) were of sufficient size to use the approach in (2). For monthly surveys, variances were calculated separately for each month over a year, and the average was used in allocation (Taylor *et al.* 2011). In some cases, a new SIC(2007) sampling stratum had no previous information to use, since the SIC(2003) equivalent industry was out-of-scope. In these cases, an allocation was made based upon variances derived from IDBR turnover data ; the resulting allocation was checked and adjusted in cases where it looked implausible. Thus, the final allocations were based upon a number of procedures and assumptions, and were reviewed when real data had been collected, resulting in some minor adjustments.

ONS uses a Permanent Random Number (PRN) system to coordinate its sampling from the IDBR (Ohlsson 1995, Smith *et al.* 2003). Each business on the IDBR is allocated a random number in the range [0, 1) which it retains thereafter. When a sample is drawn, the businesses within a stratum are ordered by PRN (along the 'PRN line') and the sample consists of a consecutive group of businesses in

this ordering, with rotation achieved by adding businesses with the next largest PRNs and dropping businesses with the smallest PRNs. The PRN line is considered to ‘wrap around’ from 1 back to 0.

The sample selection procedures employed specify how long a business should expect to be retained in a survey, how many periods should elapse before it is selected in the same survey again, and for the smallest businesses, a maximum of one survey at any given time. The change to industry stratification using the new classifications meant that currently sampled businesses from several different strata in the old design formed the initial sample in a new stratum. If there was complete freedom to choose which businesses to include in the sample, then it would be possible to maximise the overlap appropriately using optimisation methods (see Mach *et al.* 2006, Johnson *et al.* 2012 and Schiopu-Kratina *et al.* 2014 for some examples of suitable methods); similarly, if particular surveys were restricted to parts of the PRN line (as in the Swedish system SAMU, Lindblom 2003), the problem would be simplified. But in ONS there are no restrictions to parts of the line for particular surveys, and the sample selection is managed by the PRN rotation system on the IDBR, so it was important to retain the existing system.

By using the PRNs for businesses in each stratum in the design using the new classification, we automatically achieve the required selection probabilities (each PRN sample is a simple random sample of the stratum population, using the properties in Ohlsson 1995). We have a choice of which PRN to use as the starting value for the new rotation, and the best start is the one which maximises the overlap, but also obeys rules on time in survey and survey holidays. The best starting PRN was found by evaluating a penalty function for each possible PRN start (because the PRNs in any particular register are fixed, there are N_h possible starts in stratum h). The PRN start corresponding with the minimum value of this penalty was chosen as the start for the sample on the new classification. The penalty is

$$\sum_{t=1}^r \sum_{i \in h} (L_{it} + B_{it} + D_{it})$$

where r is the rotation period for the survey, t sums over a full rotation of the survey, the three components of the penalty function were defined as:

$$L_{it} = \begin{cases} r - u_{it} & \text{if } i \in S_{ht} \text{ and } 0 < u_{it} < r \\ 0 & \text{otherwise} \end{cases}$$

$$B_{it} = \begin{cases} r & \text{if } i \in S_{ht} \text{ and } s_{it} > r \\ 0 & \text{otherwise} \end{cases},$$

$$D_{it} = \begin{cases} r - s_{it} & \text{if } i \notin S_{ht} \text{ and } 0 < s_{it} < r \\ 0 & \text{otherwise} \end{cases}$$

S_{ht} is the set of sample units defined by the PRNs at period t , s_{it} is the number of consecutive periods before period t for which unit i was selected in this survey (= 0 if it was not selected in period $t-1$) and u_{it} is the number of consecutive periods before period t for which unit i was not selected in this survey (= 0 if it was selected in period $t-1$). L penalises units which have an insufficient holiday between periods in the survey (taking a holiday of length at least r as the target), B penalises units which spend more than r consecutive periods in the sample (and more heavily than the expected $s_{it} - r$ in order to penalise breaches to published expectations for the length of time in particular surveys), and D penalises units which are dropped from the survey early, that is before they have spent r periods in the sample. In practice D tends to dominate the penalty function, but this is acceptable because the minimum of $\sum \sum D_{it}$ occurs where most previously sampled units (which have not reached r periods in sample) are being re-included in the sample.

This algorithm was applied stratum-by-stratum and survey-by-survey, so did not coordinate samples across surveys. Nevertheless, it provided a way to maintain a large overlap within the existing system,

although the resulting overlap was often smaller than would have been achieved had the design remained unchanged. This resulted in larger variances of estimates of period to period change across the transition. A number of further initiatives were introduced to try to maximise response and realise the largest possible overlap, including giving advance notice to businesses of any changes in the questionnaire, and increasing resources for response-chasing in these periods.

4. Creation of new outputs

For the most part, survey outputs at ONS have been produced on the same classification as the stratification. However, with a change in classification, outputs have been required on both classifications. Careful consideration has been needed to ensure this is achieved sensibly.

Two broad approaches have been used: backcasting and dual-running. Backcasting is a *macro*-method, because it uses only aggregate statistics as the basis for a model, which is used to produce estimates on a different classification. Dual-running is a *micro*-method, because it uses the microdata, usually dual-coded, as the basis for estimation on both classifications.

The principal use of backcasting has been in the production of historical time series on SIC(2007), formerly available only under SIC(2003). For many series, backcasting has been used to give a reclassified series starting in the 1990s; this corresponds with survey redesigns in the UK (for example, the Annual Business Inquiry (now the Annual Business Survey) was introduced in 1997 (Smith *et al.* 2003), so many annual series start from this date. This provides a sufficiently long run of reclassified estimates for many uses. Econometric models, however, may require very long spans of data, so in some cases longer series have been developed, though the method of backcasting should be taken into account when developing and using such models.

Dual-running, the micro-method, may be thought of as the production of two sets of estimates (on SIC(2003) and SIC(2007)) with respect to the same reference period, at approximately the same time. It has been used for current or recent periods where dual-coded register data are available. This method was used both to produce reclassified estimates for SIC(2003)-stratified surveys (for example, the short-period surveys in 2009 were sampled using SIC(2003), but SIC(2007) outputs were required by Eurostat), and for SIC(2007)-stratified surveys (for example the short-period surveys in 2010, which were stratified and processed under SIC(2007), but when SIC(2003) outputs were also required for National Accounts until the publication of the 2011 Blue Book). See van den Brakel (2010) for a discussion of stratification and estimation methods for these two micro-method scenarios, and section 4.2 below.

The two methods (macro / backcasting, and micro / dual-running) were used at ONS to produce different parts of the same time series. Backcasting is used for estimates and outputs referring to periods longer ago, and dual-running for more recent or current periods. Therefore, the time series produced on SIC(2007) have distinct sections to them:

- oldest periods estimated directly on the old classification and adjusted by backcasting;
- more recent periods with data produced by dual-running from surveys designed on SIC(2003);
- the most recent periods, based on surveys redesigned on SIC(2007).

These sections are linked together where appropriate to avoid discontinuities (see de la Fuente Moreno (2014) for information on linking and splicing methods). Part of such a series for the Retail Sales Index is shown in Fig.3. We now examine more closely the macro and micro methods used.

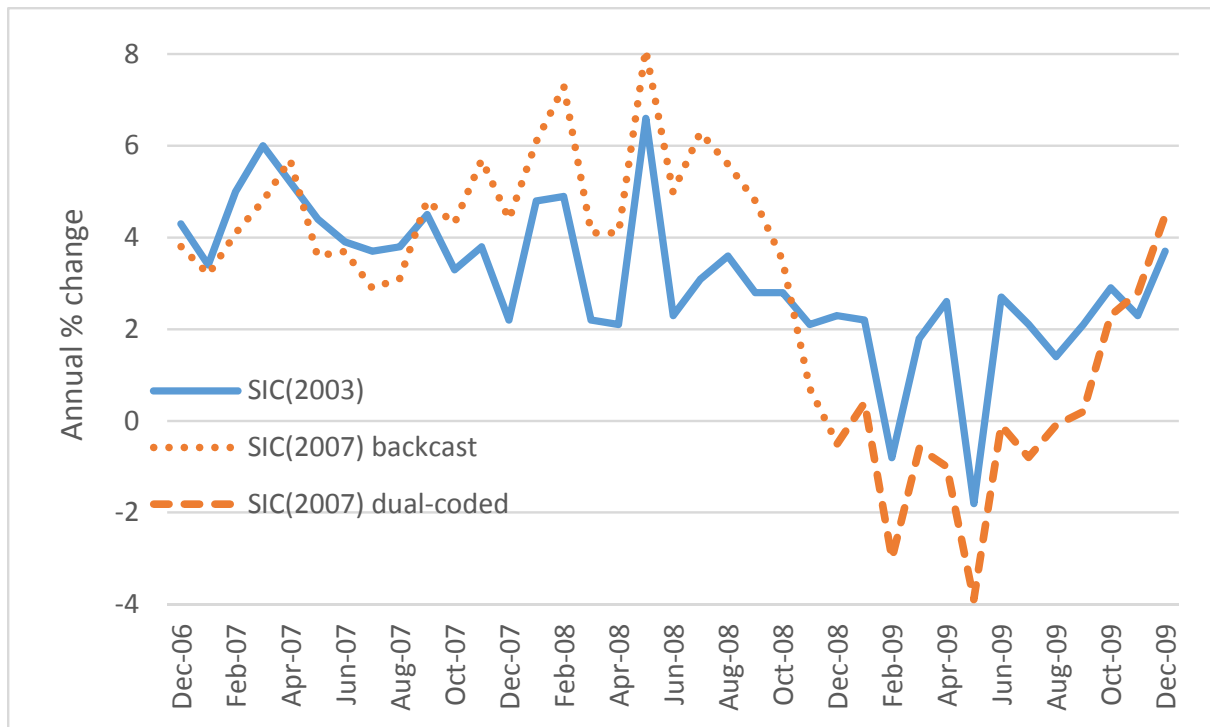


Fig. 3: An example reclassified series composed of linked parts – yearly movements in Retail Sales value (ie not deflated), seasonally adjusted, Dec 2006-Dec 2009 (redrawn from McLaren 2009, Fig. 5). The definition of retailing expanded in SIC(2007) to include automotive fuels (previously collected in another survey, as it was part of Motor Trades), and the pattern in the fuel series accounts for the very different patterns of movement in the last 6 months of 2008. The backcast portion of the SIC(2007) series contains an estimate for automotive fuels, derived from the survey which previously included it, converted to the new classification. The whole of the period shown is based on a stratification using SIC(2003) – the change to SIC(2007) stratification came in Jan 2010.

4.1 Backcasting

Backcasting, the process used to derive most historical estimates using SIC(2007) is a macro method, which is basically dependent on fitting a suitable model to the available aggregate data. The model is assumed to hold over a long time period. Such methods are applied only to the industry-level (that is, aggregate) estimates, and do not involve the micro data in any direct way other than (sometimes) to fit the model. Conversion matrices (or concordances) are the most frequent method for codifying these models for reclassifications; such methods are commonly used by National Statistical Institutes (eg Bayard and Klimek 2004, Russell *et al.* 2004, Yuskavage 2007) when handling a change in classifications. Indeed, this approach was used by ONS at the last major classification change, to SIC(92) in the mid-1990s. The big advantage of such methods lies in their efficiency: there is no need to retrieve and recode historical microdata (which may no longer exist, or may be impossible to recode), and the quality of the resulting estimates is usually acceptable. They do however rely on assumptions about the stability and appropriateness of the model, particularly where they are used for conversion over a long period. Where stability cannot be assumed it may be possible to use historic data to develop a sequence of time-varying conversion matrices (Yuskavage 2007). In the change in ONS, conversion matrices from 2008 were used consistently for years before 2008, using the assumption of stability. This was largely because of a lack of resources for recoding earlier versions of the IDBR.

Conversion matrices do not deal with classification changes which incorporate or remove whole areas of activity (structural zeroes), as there is then no information to which to apply conversion factors to

obtain an estimate. A further disadvantage of the application of fixed conversion matrices is that historical re-classifications of individual large businesses will be missed (since the same, recent proportions are used to split and re-aggregate estimates for all historical periods), which may lead to unrealistic results. This emphasises the need for manual examination of individual cases and converted series to identify when the assumptions about model stability have broken down. In these cases a suitable adjustment can be made to the converted series.

The foundation of conversion matrices is a dual-coded business register or census, from which cross-tabulations show the proportions of businesses that map from a SIC(2003) code (or group of codes) to a SIC(2007) code (or group of codes), and vice versa. In ONS, such matrices were produced from extracts of the IDBR, based on the most detailed (5-digit) level of classification, with proportions reflecting business sizes (in terms of turnover and employment). These matrices were then used to apportion SIC(2003) aggregate estimates to SIC(2007) codes, which were then re-summed to derive SIC(2007) aggregate estimates for historical periods.

Conversion matrix options: The choice of which conversion matrices to use was decided on a case-by-case basis for the various ONS outputs, with compromises being inevitable; consistency was kept wherever possible:

- **Timing:** For most of the short-period surveys, conversion matrices were used to create SIC(2007)-based estimates for reference periods up to the end of 2008, with micro methods (see below) being used from the start of 2009. In order to estimate and link out any discontinuities, conversion matrix-based estimates were also produced for 2009, and the matrices themselves were produced with IDBR extracts taken at that time. Matrices were somewhat volatile through early-2008, caused by the new SIC(2007) codes settling (Fig. 2), but became much more stable later in the year. Early conversions were therefore of lower quality, but were redone when the stabilised matrices were available. Further register updates were applied from the 2008 structural surveys, increasing quality and stability further.
- **Base variable:** Most ONS business survey estimates are based on ratio estimation (Smith *et al.* 2003), and outputs that use register employment as an auxiliary variable in estimation (generally those with labour-market related outputs) used conversion matrices based on employment size, whereas those that use turnover as an auxiliary variable (generally related to business output) used conversion matrices based on turnover size. Note that this approach can reduce the consistency in estimates of variables derived from both types of data, such as productivity; the inconsistency was accepted in ONS because of the benefits of using conversion matrices appropriate to the variables, but other decisions would be possible. Matrices based on counts of businesses were generally not used, but would be appropriate if trying to backcast series of estimates of numbers of businesses.
- **Statistical Units:** The matrices could be calculated separately using different types of units as their basis; the usual survey practice was used as a guide as to which one to use. Most ONS business surveys are based on responses from RUs (see section 2), and these used conversion matrices calculated from RU information. Labour market estimates are based on additional information from LUs, and these series used conversion matrices calculated from LU information.
- **Level of conversion:** Although the matrices are usually calculated at the 5-digit classification level, alternative levels of aggregation can be used. Naturally, the level of the series (estimates) to be converted will largely inform the decision, but coding error would be more likely at lower levels (the detail may be incorrect within the appropriate broad industry group). Examination of the conversions at detailed level in ONS suggested that the timing of extracts (Fig. 2) from the IDBR had a larger effect on coding error than the choice of the level at which the conversion matrix was applied.

- In all cases, each derived series was carefully checked and this allowed manual interventions to be made where appropriate.

The settling of the SIC(2007) codes on the IDBR also had some notable effects on the conversion matrices. Many of the initial codes were imputed from old SIC(2003) codes, as there was no business description available, and these cases could only follow the table of 'official' correspondences. Since then, many of the imputed codes have been replaced with directly-coded SIC(2007) codes, which has seen the number of non-zero correspondences in the cross-classified tables increase (sometimes codes derived from business descriptions fall outside the 'official' correspondences). As an example, in the matrices for conversion from 4-digit to 3-digit codes (with a few exceptions), the number of non-empty cells approximately doubled from January to May 2008. (Of course, some of the changes in SIC(2007) codes may reflect actual changes in activity after the SIC(2003) code was assigned, or errors in the original assignment of SIC(2003) codes, as well as errors in the SIC(2007) code).

There are also choices for the stage of processing at which conversion should occur. ONS's Index of Production (IoP), for example, is largely based around estimates of turnover, which are then deflated using price indices before being seasonally adjusted. Thus, for the IoP, there were three reasonable choices regarding conversion: (1) conversion of all inputs (turnover estimates and price indices); (2) conversion after deflation; or (3) conversion after seasonal adjustment. Investigations revealed the differences between deflating before conversion and conversion followed by deflation seem to be generally small. The largest effects were where the weights of components deflated differently change over time. Consultation with experts suggests that converting after deflation might make more sense economically, but there are practical considerations, such as co-ordination of the timing of the reclassification of the deflator (a price index) and the data to which it is applied, which may be more important than the choice of stage to convert. A reviewer suggested that another reason for converting the values before deflation is that the conversion matrix was developed using these values. Relationships among the variables in these matrices are more likely to remain stable than the relationships between variables before and after deflation. The need for a history for deflators on the new classification too suggested that conversion of deflators first would give the best coherence between the various outputs of the reclassification. A considerable amount of checking was required to ensure that the historical price indices were coherent and credible.

Applying the conversion matrix to seasonally adjusted series would give linear combinations of the seasonally adjusted input series, and these outputs should in general appear seasonally adjusted. This seems to work in practice, although there remains a concern that a linear combination of the errors in the seasonal adjustment decomposition may show some residual seasonality. In line with the principles in the European Statistical System Guidelines on Seasonal Adjustment (Eurostat 2015b sec. 3.4) we therefore prefer direct seasonal adjustment of the new series. This allows for different seasonal patterns in the new component series. Therefore we prefer conversion of all inputs (including the deflators) to the new classification, and for seasonal adjustment to take place after deflation.

Conversion matrices were made available for users to do their own conversions. As well as the turnover- and employment-based versions, a conversion matrix using the number of units was produced and published (ONS 2010); each published matrix had entries rounded, where appropriate, to reduce the risk of disclosure.

A further use of conversion matrices, noted here for completeness, can lie in the dual-coding of micro data. ONS found a demand for this on historical social survey datasets, which are made available for use by approved researchers. The actual application was on Occupational Classification codes, but the principal for Industrial Classification codes would be the same. The data sets contained an old classification code for each case (row), for which users wanted to assign a new code. Naturally, without

a frame (or the ability to link to one) and no means to recontact the respondent, there is no way of gathering the required information precisely.

However, a new classification code can be assigned probabilistically using proportions from a conversion matrix. As an example, if an old code maps to three new codes in the ratio 60% : 30% : 10%, then a Uniform(0,1) random number can be generated for the case, and a new code generated according to which of the intervals (0.0,0.6], (0.6, 0.9] or (0.9, 1.0] the random number lies in. Although this does not guarantee that the code for any particular case is correct, at an aggregate level distributions should be reasonable. The method can be refined further, as required, for example for longitudinal or panel data, judicious use of case IDs as random number generator seeds can be used to ensure consistency of coding over time.

To this end, ONS made available conversion matrices and program code in the more popular statistical programming languages to researchers to dual-code their own data sets in a controlled way.

4.2 Microdata methods

The foundation of micro methods lies in the survey responses themselves. Each sampled unit has two codes (SIC(2003) and SIC(2007)) – one of which will have been used for stratification, and the returns are re-aggregated to form estimates on the other classification. There are different methods that may be used for this (domain estimation, poststratification, and so on), each with its own advantages and disadvantages. Four alternative approaches are considered by van den Brakel (2010). The method most commonly used for ONS outputs, a domain estimation approach, is described here.

We first note that estimates for most ONS business surveys are derived using the ratio estimator. For an estimated total of a variable y , \hat{t}_y , this may be written as follows (Särndal *et al.* 1992, eqn. (6.5.9)):

$$\hat{t}_y = \sum_{i \in S} d_i g_i y_i$$

where:

d_i is the design weight (the inverse of the selection probability), and reflects the stratification of the survey

g_i is a calibration weight, in general determined by defining calibration groups and an appropriate estimator (from a very wide class). The specific case of the (within-stratum) ratio estimator gives

$$g_i = X_h / \sum_{i \in S_h} d_i x_i \text{ for } i \text{ in stratum } h.$$

y_i is the (sometimes Winsorised) survey response, or imputed value in the case of non-response, for the variable y . i indexes the businesses, and s is the selected sample. It was decided to maintain the winsorisation from the original classification for resource and consistency reasons. But another option would be to re-assess outliers on the poststratification on the new classification.

The choice of the most appropriate calibration groups for outputs was discussed at length, and the Government Statistical Service Methodology Advisory Committee consulted². Under stratification by only one SIC, the calibration groups are usually just the sampling strata, or groups of sampling strata. However, with two classifications being used for simultaneous outputs, the choice is not clear. Any change of the groups (for example to cross-classification of the SICs) would lead to a break in time series, and also present a risk of small sample sizes in each group. Therefore the decision was made to calibrate the SIC(2003) estimates within SIC(2003)-defined groups, and SIC(2007) estimates within SIC(2007)-defined groups, with the compromise of the two sets of outputs being calibrated differently.

² See <http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/method-quality/advisory-committee/2008-2011/15th-meeting/index.html> (accessed 27 Oct 2016)

Thus, for example, during periods in which the surveys were stratified by SIC(2003), the two sets of estimates would be compiled as follows:

$$\hat{t}_y^{03} = \sum_{i \in S} d_i^{03} g_i^{03} y_i$$

$$\hat{t}_y^{07} = \sum_{i \in S} d_i^{03} g_i^{07} y_i$$

(with natural notation) and when stratified by SIC(2007), the estimates would be given by

$$\hat{t}_y^{03} = \sum_{i \in S} d_i^{07} g_i^{03} y_i$$

$$\hat{t}_y^{07} = \sum_{i \in S} d_i^{07} g_i^{07} y_i$$

It was felt that having consistency within time periods was better than having consistency between the SIC(2003) and SIC(2007) outputs. This is a composition of the methods in sections 3 and 5 of van den Brakel (2010), with a need for a way to link the estimates from the two approaches at the transition point.

4.3 National Accounts

The main conversions for National Accounts purposes were based on SUTs. These provide the basis for calculation of Gross Value Added (GVA) weights to allow other elements of the National Accounts to be weighted together appropriately for the new classification. The industrial groupings to be used for SUTs based on the new classification were agreed quite early so that they could be used as a basis for the revised stratifications in surveys (see section 3). The change to a new product classification, CPA 2008, occurred at the same time, which had an effect on the way groups were put together. A consultation provided evidence which was used to form a 114-group classification consistent with SIC(2007) and CPA 2008 which encompassed international reporting requirements and provided some additional detail for users within the UK. These groups form the main processing level in the construction of the National Accounts, and component processes have also moved to this structure, introducing greater harmonisation into National Accounts compilation. We focus on the industrial classification changes rather than the product classification ones in this section.

Supply-side estimates in the SUTs up to 2006 were converted by the application of the conversion matrices as used for the surveys. The later conversion of National Accounts allowed a stable set of matrices to be used throughout the accounts, mostly derived from the final IDBR reclassification, but with specific information for industries not covered by the IDBR derived from a range of different sources and used to complete the conversion. Since the basis of National Accounts is monetary, the turnover-based versions of the conversion matrices were used. Such a conversion is closest to the concepts used in measuring output and intermediate consumption. The IDBR does not adequately cover financial industries and non-market producers, and in these cases turnover is often not a suitable target for measuring output. So other sources were used to derive conversion factors here (many of which were largely unaffected by the change in classification, and so could be dealt with by one-to-one conversions), so that all the elements of the economy were covered. Once the full matrix was available, it was used to convert the industry totals (on one margin of the SUTs) to give totals on the SIC(2007) classification.

SUTs contain considerable detail on products, some of which is derived from the EU's harmonised product survey, PRODCOM, and the parallel update to the EU's product classification system (to CPA 2008) allowed product patterns in SIC(2003) industries to be converted to CPA 2008. Then the product distributions were applied to the converted industry totals by associating SIC(2007) industries with one or more SIC(2003) industries (Drew and Dunn 2011). This led to some inconsistencies between the industry totals and the product totals, which required some further balancing interventions. The

SUT is a matrix, and its major entries are generally on the diagonal (showing the principal products of the various industries). In cases where manual interventions were required there was therefore a working assumption that the diagonal entries should be maintained as far as possible. Whilst aggregate values should conceptually remain unchanged through such a conversion process, there were some minor revisions to totals in the SUTs up to 2006 in moving to the SIC(2007) classification. The largest of these was 0.05% of the total, so the old and new series are indistinguishable graphically. The overall effect of the classification change on GDP is included in the differences shown in Fig. 4; it is confounded with some other changes introduced at the same time, so it is not as clear as having the classification change separately identified, but it gives an impression of the reclassification effect.

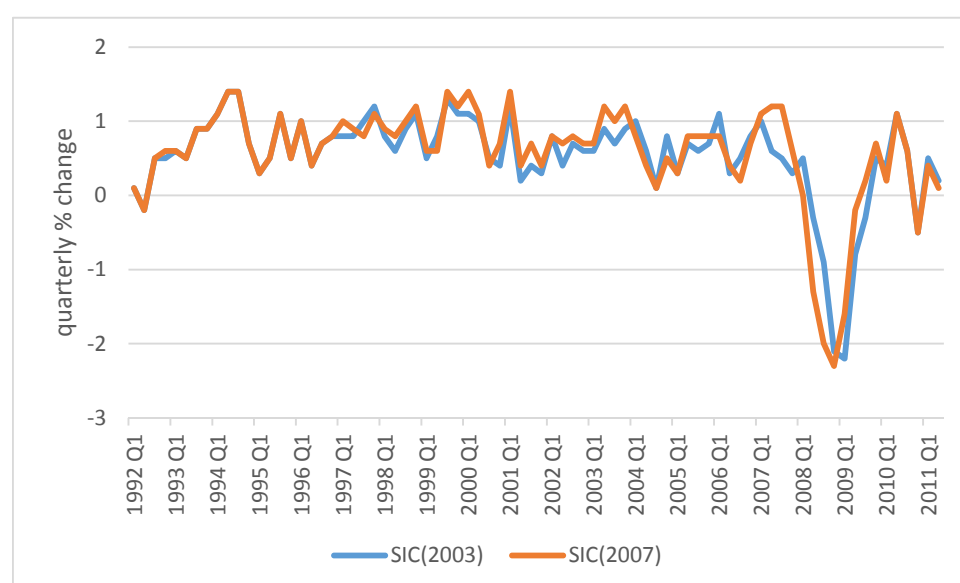


Fig. 4: Time series of quarterly changes (%) in UK GDP from before and after the implementation of the reclassification to SIC(2007). The differences are due to reclassification and some additional changes made at the same time, so only part of the difference is the reclassification effect.

The SUTs for 2007 to 2009 could be based on data classified according to SIC(2007). The 2007 estimates from the ABS, the main survey source for SUTs, were reworked using the SIC(2007) classification, and from 2008 onwards the ABS was designed on SIC(2007) and could be used directly. The 2008 and 2009 ABS estimates were initially converted, and used with existing monthly movements to provide an interpolated series of within-year changes. Such changes are important and for interpretation need to be seasonally adjusted. In order for seasonal adjustment of this derived monthly series on the new classification to be sufficiently stable, there was a need for a longer time series, so additional years of ABS data were also converted and monthly paths interpolated.

To ensure that estimates were consistent over time, any statistically important differences between the three parts of the series – estimates processed by conversion matrices, the reworked data in 2007 and the newly reclassified data for 2008 – were taken into account. For the ABS estimates, this was “done through a linking process where statistically [important] differences are blended into the historical estimates over time. This ensures that revisions to the levels of historical published data are minimised and the economic context of historical data is preserved as closely as possible” (Drew and Dunn 2011). Yuskavage (2007) describes in detail a similar approach with series composed of three parts when the US Bureau of Economic Analysis introduced the North American Industry Classification System (NAICS); in their case the middle period was based on a succession of conversion matrices calculated over ten years by recoding the underlying industry data.

Seasonal adjustment of all time series forming the SUTs was also needed, and following the same principles as discussed above (section 4.1), was undertaken after the conversions and any adjustments had been made.

5. Considerations in choosing a conversion method

5.1 Disclosure issues associated with changing classifications

Reclassification potentially causes difficulties for statistical disclosure control, because small differences in classifications can result in ‘slivers’, domains which contain very few observations but which can be estimated quite accurately by differencing estimates on the two classifications (Hundepool *et al.* 2012, sec. 5.2.3). This can apply to the results of the reclassification, or to the conversion matrix itself. For example, Anon. (no date) were able only to calculate a restricted range of ABS estimates for 2008 on the old SIC2003 classification because of disclosure issues, which are described as ‘especially complex’ and not detailed (a common practice to increase disclosure protection). Similarly, Bayard and Klimek (2004) could not use some cells of the conversion matrix from the US Census Bureau because they failed disclosure tests. In many cases, the effect of converting value series and deflators separately will result in added protection from disclosure for the final outputs, as long care is taken with any release of the conversion matrix itself.

5.2 Quality of conversions

Consistency

The application of the model-based conversion matrices to part of a time series, with other parts of the time series covered by direct estimation from surveys designed on different classification systems generates some fairly standard problems in consistency of surveys, and general guidance on producing consistent time series of estimates in this case is given by van den Brakel *et al.* (2008). A model-based approach to this involves the Kalman filter, and Bollineni-Balabay *et al.* (2016) give a complex example covering multiple changes, which demonstrates the efficacy of the procedure in more complex cases than reclassification.

The use of conversion matrices generally builds in some consistency of estimates. As long as the matrix covers all of the measured activity, and the new classification does not measure some activity not previously covered, the conversion approach guarantees that the total in the economy is unaltered. Reprocessing the microdata does not have the same property, and this results in changes to estimated totals.

In the UK almost all of the National Accounts were converted by application of the conversion matrices based on turnover, which were derived from the business register. Conversion of other variables (such as fixed capital formation, FCF) using this matrix gives consistency, but is not necessarily the best for FCF. A conversion matrix could be constructed for FCF, but since this is not a variable on the IDBR it would need to be constructed from a dual-classified survey source, and would therefore be subject to sampling error. It is an interesting question whether a matrix tailored for a specific variable but with sampling error is better than a non-tailored matrix (which relies on the relationship between the target variable and the variable in the conversion matrix) based on the whole population.

Historical consistency

It is interesting to consider what the effect of successive reclassifications is on historical data. The UK has had six national classifications since the first harmonised classification was introduced in 1948. It is not known for certain whether adjustments were made in the early transitions to new classifications, but certainly the updates to SIC(80), SIC(92) and SIC(2007) have involved the application of conversion matrices, and the SIC(92) and SIC(2007) conversions required adjustments to the National Accounts. This means that the oldest information has had a series of model-based adjustments applied to keep it consistent with modern classifications, so that users can use long time series in their economic models. Soroka *et al.* (2006) consider the effects of such changes, and conclude that the final adjusted data no longer bear a close resemblance to the original estimates (both in the level and in some cases the pattern) – historical accuracy is downplayed in favour of long-term historical consistency.

Timing consistency

Because of the different periodicities of structural and short period surveys, samples are often selected for different periods at approximately the same time. When classifications change, this can mean that the timing of the change in classification is different for different types of surveys (see for example Walker 1993). The same effect was felt for the change to SIC(2007), when the annual surveys changed to sampling on the new classification in 2008, for which samples were selected at the end of 2008, and the monthly and quarterly surveys changed from January/Q1 2010. This leads to consistency issues between annual and short period surveys for estimates for 2008 and 2009. These are ameliorated through conversions to and from the new classification.

Accuracy

Ultimately the decision to change classifications is taken to improve the accuracy of the outputs based on the classification – making it relevant to current industrial organisation and emerging activities. Making classifications current necessarily makes them less relevant for historical statistics, and care must be taken in long-term backcasting not to introduce clearly nonsensical patterns (such as production of high-definition televisions in the 1980s).

At the time a classification is introduced there are some more-or-less short-term effects on accuracy of the resulting series, which are a consequence of the process of changing. There is the time taken for the new classification to stabilise (as businesses are progressively contacted through surveys and administrative processes and errors in assigning them to new classification codes are gradually corrected); the effect of the errors in the classification codes is generally to increase the variance and not to institute a bias (imputed classifications must be done in a stochastic manner where there is no other information in order to ensure this property).

The sample design also undergoes a transition, and has several possible effects. Variability in new classifications will introduce lack of homogeneity into new stratifications until the classification matures, which will increase the variance of estimates. The effect of changing to rotational sampling in new strata with a maximised, but smaller than steady state, overlap is to reduce the beneficial effect from covariances between periods, and therefore to increase the variance of estimates of change.

So during the transition period there are several effects which all act to increase the variability of survey estimates. Therefore one strategy for maintaining quality during reclassification might be to temporarily increase sample sizes to compensate for the increase in variability, and this approach was taken in a limited number of surveys in the Netherlands in 1993 (Beekman 1992). In the medium term the move to a more relevant classification should marginally increase homogeneity in strata and therefore reduce variability.

Coherence and comparability

The debate over the use of single, harmonised classification systems in official statistics has long been won, but there are still issues of historical coherence, particularly when different parts of series are converted by different methods. This requires attention to the joins between methods, which need to be adjusted so as to give a coherent time series with the best estimates on the new classification.

5.3 Communication

The implementation of a new classification system is an important event for users of statistics because of the impacts on their own inferences, and it is therefore important that the plans and procedures being used to update the classifications are communicated to them. There is a case for providing some preliminary information on the effects of reclassification (as was done for the Retail Sales Index in the UK (McLaren 2009)), as a way of helping users by giving them early warning of the impacts of the change in classification.

Once the material changes have been made, it is important for the metadata which describe these changes to be recorded and readily available to users to enable them to use this information to support their own analysis. Some of the metadata is directly published, such as the description of the new classification, its codebook and any indexes. But other information, such as how particular portions of a time series have been converted and the methods by which these pieces have been spliced together coherently can be less easily discovered, and should also be made explicitly available.

The effect of classification changes is not generally well presented in graphical presentations. It has become more commonplace (at least in the UK) to label graphs with significant events that help users to understand the evolution of time series. At least at the time of issuing reclassified estimates, it would be beneficial to have visualisations which highlight where the changes in methods have taken place within a time series, in a similar way to Fig. 3.

6. Conclusions

It is necessary to update classification systems occasionally so that they remain relevant by accounting for changes and developments in the goods and services which form the economy; this is probably even more important now as innovative products are frequently introduced to the market, and as the modern economy starts to take on new forms. As a consequence there will be changes to statistical series, and these will reduce the quality of estimates for a period while the new classification is introduced and settles down, in the sense that statistical units move from temporary codes at the instant of transition to 'correct' codes with time.

This process affects many stages of the statistical production process (Generic Statistical Business Process Model, UNECE 2013), but will most usually be felt as a period where the variability of estimates is increased somewhat. It is necessary to work through the detail of a reclassification procedure – there are many steps in processing, and all of them need to be consistent with the new classification, otherwise there is a big risk to have estimates which are not coherent. Similarly, there is a large job of quality assurance of the outputs from the various methods for converting series onto a new classification – a need to check the credibility of, for example, macro methods applied across historical classifications, when the model assumption of consistency through time breaks down. In general, conversions should be checked to identify anomalies such as missing series, sizeable changes in turnover ratios, spikes which are attributable to a single business rather than an industry class as a whole and so on. Time series consistency is usually a strong requirement from users, though it can

have the effect of making estimates very different (in level and sometimes in evolution) on new classifications from their original estimates on the original classification (Soroka *et al.* 2006).

The temporary effects on the quality of statistical outputs at the time of the classification update are necessary in order to avoid more widespread effects on quality as a result of the classification itself becoming out of date. In the medium term there are benefits. The new industrial structure should better measure new industries, and enable rapidly growing sectors to be separately identified so that they can be monitored.

We conclude by offering some considerations on the process of implementing a new industrial classification in a national statistical institute, which may be of use by other organisations undertaking such a change:

- With limited resources, it will be impossible to do everything. A better approach is to decide on areas which are most important, and to focus on these, spending most time ensuring the quality of backcasts of these estimates. Even so, it will be unlikely that all user requirements can be met, and it must be accepted that the quality of some converted series will be lower than that of the original estimates, possibly substantially so.
- The various options available in conversion, with no unique, obviously ‘right’ approach, can lead to difficult decisions. Where the quality resulting from any one approach cannot be demonstrated as being appreciably better than others, it may be prudent to choose approaches that can be easily explained and justified.
- Discontinuities will be introduced in some time series. These, and differences in treatment of different periods in series should be clearly identified for users (for example Fig. 3). Explanations of these differences should be presented to users. (Obviously when fitting seasonal models, discontinuities should be estimated from the evolution of the series, and then removed.)
- Good and clear communication is essential. Document the decisions made and the methods employed; these will be invaluable next time.
- Retain information on the effects of changes in classifications (separately from other changes if possible), and on experimental work to investigate alternative approaches and their effect on outputs. We had hoped to include more examples in this paper to illustrate different effects, but the detailed comparisons were not available. Gathering these details will provide extra evidence with which to evaluate approaches and confirm (or otherwise) the validity of our recommendations.
- The period of implementation is likely to be long. Plan ahead, and get input from those areas moving last (for example, National Accounts) at the start, to inform re-design work at the beginning.
- Take the opportunity to review and improve other aspects of data collection and the production of statistics when implementing a change in classification; the additional cost of doing this at the same time may be quite low.

Acknowledgements

Many people in ONS were involved in implementing SIC(2007), and it would not have been possible to write this article without their efforts. We are very grateful to the Associate Editor and three anonymous referees for their comments, which have led to a much improved article. Some of the final considerations derive from a visit to ONS of the late David Binder. The views expressed in this article are those of the authors and do not necessarily represent those of the ONS.

References

- Anon. no date. *SIC 2003 comparison of main industrial groupings in 2007 and 2008 by key variables*. Newport: ONS. Available at <http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guid-method/method-quality/specific/business-and-energy/annual-business-survey/annual-business-inquiry-industry-groupings.pdf> (accessed 5 April 2016).
- Anon. 2012 *Supply-Use Tables and Standard Industrial Classification (2007) conversion*. The Scottish Government. Available at www.gov.scot/Resource/0042/00421501.doc (accessed 28 October 2015).
- Bayard, K.N. & S.D. Klimek. 2004. "Creating a historical bridge for manufacturing between the Standard Industrial Classification system and the North American Industry Classification System". In *Proceedings of the Business and Economic Statistics Section: American Statistical Association*, August 2003. 478–84.
- Beekman, M.M. 1992. "Development and implementation of a new standard industrial classification." *Netherlands Official Statistics* 7: 18-26.
- Bollineni-Balabay, O., J. van den Brakel & F. Palm. 2016. "Multivariate state-space approach to variance reduction in series with level and variance breaks due to survey redesigns." *Journal of the Royal Statistical Society, Series A* 179: 377-402. DOI: 10.1111/rssa.12117.
- Cochran, W.G. 1977. *Sampling techniques*. New York: Wiley.
- de la Fuente Moreno, Á. 2014. "A "mixed" splicing procedure for economic time series". *Estadística Española* 56: 107-121.
- Drew, S. & M. Dunn. 2011. *Blue Book 2011: Reclassification of the UK supply and use tables*. Office for National Statistics. Newport: ONS. Available at <http://www.ons.gov.uk/ons/rel/input-output/input-output-supply-and-use-tables/reclassification-of-the-uk-supply-and-use-tables/reclassification-of-the-uk-supply-and-use-tables-pdf.pdf> (accessed 15 July 2015).
- Eurostat. 2015a. *ESS handbook for quality reports, 2014 edition*. Luxembourg: Publications Office of the European Union.
- Eurostat. 2015b. *ESS guidelines on seasonal adjustment, 2015 edition*. Luxembourg: Publications Office of the European Union.
- Hughes, J.C. 2008. "SIC 2007: implementation in ONS." *Economic and Labour Market Review* 2: 41-44.
- Hundepool, A., J. Domingo-Ferrer, L. Franconi, S. Giessing, E. Schulte Nordholt, K. Spicer and P.-P. de Wolf. 2012. *Statistical disclosure control*. Chichester: Wiley. DOI: 10.1002/9781118348239.
- Johnson, W., S. Paben, & J. Schilp. 2012. The use of sample overlap methods in the consumer price index area redesign. In *Proceedings of the fourth International Conference on Establishment Surveys (ICES-IV)*, 12-14 June 2012. Available at www.amstat.org/meetings/ices/2012/papers/301813.pdf. (Accessed 27 October 2016.)
- Lindblom, A. 2003. *SAMU 4 The system for coordination of frame populations and samples from the Business Register at Statistics Sweden. Background Facts on Economic Statistics* 2003: 3. Stockholm: Statistics Sweden.
- Mach, L., P.T. Reiss and I. Şchiopu-Kratina. 2006. "Optimizing the expected overlap of survey samples via the northwest corner rule." *Journal of the American Statistical Association* 101: 1671-1679. <http://dx.doi.org/10.1198/016214506000000320>
- MacDonald, B. 1995. "Implementing a Standard Industrial Classification (SIC) System Revision". In *Business Survey Methods*, edited by B. Cox, D. Binder, B. Chinnappa, A. Christianson, M. Colledge and P. Kott, 115-129. New York: Wiley. DOI: 10.1002/9781118150504.ch7
- McLaren, C. 2009. *Classification changes in retail sales*. Newport: ONS. Available at: www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/retail-sales/classification-changes-in-retail-sales.pdf. (accessed 10 November 2015.)

- Ohlsson, E. 1995. Co-ordination of samples using permanent random numbers. In *Business Survey Methods*, edited by B. Cox, D. Binder, B. Chinnappa, A. Christianson, M. Colledge and P. Kott, 153-169. New York: Wiley. DOI: 10.1002/9781118150504.ch9
- ONS. 2009 *UK Standard Industrial Classification of Economic Activities 2007 (SIC 2007) - Structure and explanatory notes*. Basingstoke: Palgrave MacMillan. Available at www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/standard-industrial-classification/sic2007---explanatory-notes.pdf. (accessed 5 Jan 2015.)
- ONS. 2010. *Weighted tables with percentages UK SIC 03 - UK SIC 07 and UK SIC 07 - UK SIC 03*. Newport: ONS. Available at <http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/standard-industrial-classification/index.html>. (accessed 5 Apr 2016.)
- Russell, M., P. Takac & L. Usher. 2004. "Industry productivity trends under the North American Industry Classification system." *Monthly Labor Review* November 2004: 31-42.
- Särndal, C.-E., B. Swensson and J. Wretman. 1992. *Model-assisted survey sampling*. New York: Springer.
- Schiopu-Kratina, I., J.M. Fillion, L. Mach and P.T. Reiss. 2014. "Maximizing the conditional overlap in business surveys." *Journal of Statistical Planning and Inference* 149: 98-115. <http://dx.doi.org/10.1016/j.jspi.2014.02.002>.
- Smith, P. 2013. "Sampling and estimation for business surveys." In *Designing and conducting business surveys* by G. Snijders, G. Haraldsen, J. Jones and D.K. Willimack, 165-218. Hoboken, New Jersey: Wiley. DOI: 10.1002/9781118447895.ch05.
- Smith, P. & S. Penneck. 2009. *100 years of the Census of Production in the UK*. GSS Methodology Series No 38. Newport: ONS. Available at <https://www.ons.gov.uk/ons/guide-method/method-quality/specific/gss-methodology-series/gss-methodology-series--38--100-years-of-the-census-of-production-in-the-uk.pdf>. (accessed 27 October 2016).
- Smith, P., M. Pont and T. Jones. 2003. "Developments in business survey methodology in the Office for National Statistics, 1994-2000" (with discussion). *Journal of the Royal Statistical Society, Series D* 52: 257-295. DOI: 10.1111/1467-9884.03571.
- Soroka, S.N., C. Wlezien and I. McLean. 2006. "Public expenditure in the UK: how measures matter." *Journal of the Royal Statistical Society: Series A* 169: 255-271. DOI: 10.1111/j.1467-985X.2006.00397.x.
- Taylor, C., G. James and P. Pring. 2011. "The development of the Monthly Business Survey." *Economic and Labour Market Review* 5: 95-103.
- UNECE. 2013. *Generic Statistical Business Process Model GSBPM (Version 5.0)*. Geneva: UNECE. Available at <http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0> (accessed 10 November 2015).
- van den Brakel, J. 2010. "Sampling and estimation techniques for the implementation of new classification systems: the change-over from NACE Rev. 1.1 to NACE Rev. 2 in business surveys." *Survey Research Methods* 4: 103-119. DOI: <http://dx.doi.org/10.18148/srm/2010.v4i2.2354>.
- van den Brakel, J.A., P.A. Smith & S. Compton. 2008. "Quality procedures for survey transitions – experiments, time series and discontinuities." *Survey Research Methods* 2: 123-141. DOI: <http://dx.doi.org/10.18148/srm/2008.v2i3.68>.
- Walker, C. 1993. "Transition to the new standard industrial classification (SIC(92))." *Economic Trends* 472: 88-94.
- Williams, N. 2006. "ACTR/IDBR Test Evaluation Report." *Survey Methodology Bulletin* 57: 33-38.
- Yuskavage, R.E. 2007. "Converting historical industry time series data from SIC to NAICS." The Federal Committee on Statistical Methodology 2007 Research Conference, November 5-7 2007. Available at https://fcsm.sites.usa.gov/files/2014/05/2007FCSM_Yuskavage-V-A.pdf. (accessed 15 July 2015.)