

A1 Regional geographies

A1.1 There is no single appropriate regional breakdown of a country, in terms of either sizes or definitions of regions. The breakdowns used, however, have tended to develop from historical, political and administrative structures. In this way, the UK first naturally breaks down into the four countries. In turn, the most obvious and recognisable ‘regionalisation’ of the UK at the next level of detail might be regarded as counties, and so on to electoral ward. From the statistical/economic perspective the interest has been in a breakdown of the UK into the three countries and nine main regions.

A1.2 Since 1996, the nine regions have been based on the English ‘Government Office Regions’. In 1998, the status of this regionalisation was re-inforced by its adoption as part of the ‘Nomenclature of Units for Territorial Statistics Classification’ (NUTS) of the regions of the European Union as a whole. The nine English regions and the three countries are now known as NUTS 1 and are the basic building blocks of most regional macroeconomic statistics for the UK and wider studies. In addition, the NUTS classification involves a detailed hierarchy down to smaller regions and areas. NUTS 2 divides the UK into 37 areas, used primarily for the allocation of EU structural funds; NUTS 3 divides the UK into 133 areas that are principally individual counties, unitary authorities or groups thereof; NUTS 4 (now officially LAU1) is then close to local authorities (about 440); and NUTS 5 (now officially LAU2) to about 11,000 electoral wards. The hierarchy is nested, e.g. each NUTS 1 region breaks down into NUTS 2 regions. Example maps of the classifications are shown on the following three pages.

A1.3 The development of regional policy has endorsed and further re-inforced this structure of sub-national disaggregation. For example, the RDAs created in 1999 relate to the English NUTS 1 regions. NUTS, therefore, goes a long way to meeting the demand for a distinct and recognisable geographical breakdown of the UK. However, three further points should be noted:

- while increasingly adopted, the NUTS structure is not universal. In particular the administrative structures of the key public policy areas of health, education and the police operate according to alternate geographical definitions;
- NUTS has not in the past and may not necessarily in the future be fixed, although the adoption of the NUTS EU Regulation now sets down fairly strict rules for when a change to the NUTS structure may be justified. For statistical purposes, constant geographical definitions are invaluable, particularly for comparisons over time; and
- the demand to define regions may actually serve to emphasise latent inequalities within a nation. There will be an inevitable tendency to group together like populations where geographical and administrative boundaries are less clear-cut (particularly perhaps at finer levels of detail).

Figure A1.1: NUTS 1 Statistical Regions of the United Kingdom

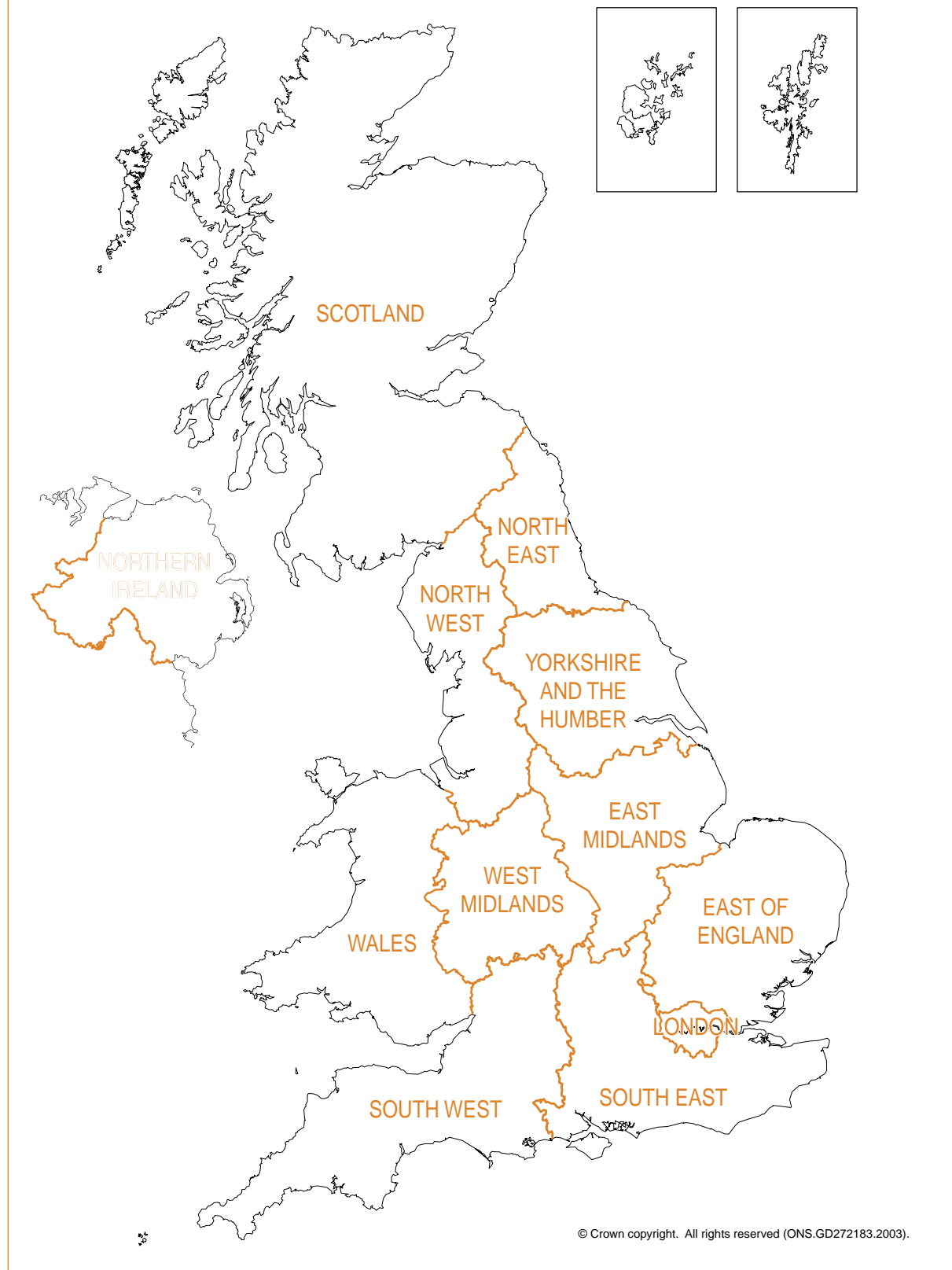
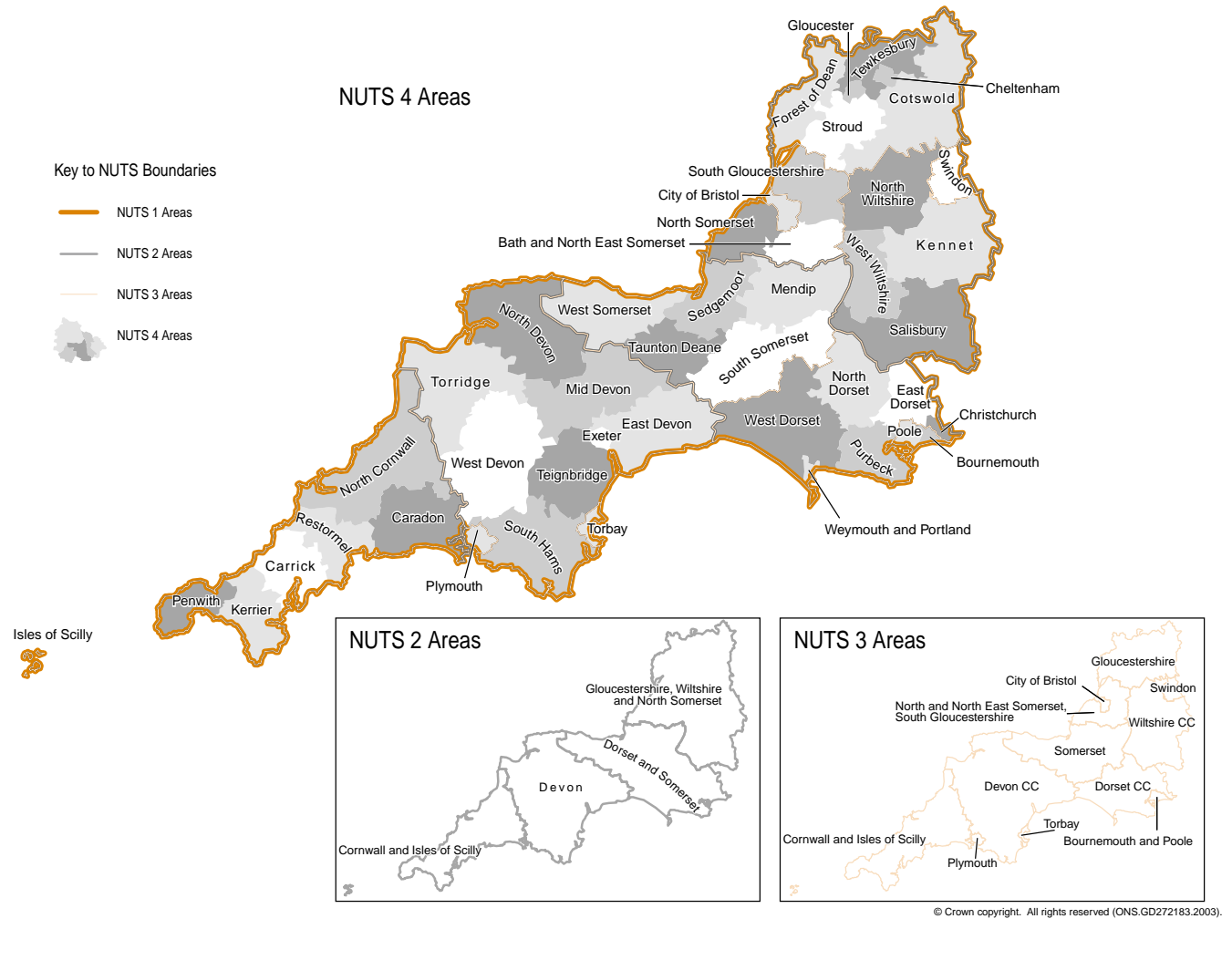


Figure A1.2: NUTS 1 and 2 Statistical Regions of the United Kingdom



Figure A1.3: South West NUTS 2, 3 and 4 Areas



A2 Extending the ABI to meet regional requirements

A2.1 This annex was produced by the Office for National Statistics to illustrate some of the options relating to regional stratification of the ABI and enhancing the sample size. The resource and compliance cost estimates included in support of the options set out below are sensitive to the exact requirements for data at a regional level. Further detailed work will be needed when there is a final specification of the requirement to firm up these estimates.

Summary

A2.2 The key points are:

- the present sample design of the ABI stratifies by country and by detailed 4-digit industry. Increasing the number of ‘regional’ strata implies reducing the number of industry strata and, without an increase in the overall sample size, would provide less industry detail at regional level than currently at national level. The detailed industry estimates for the UK would also be of worse quality;
- we would need to select around 570,000 businesses to produce regional estimates of the same quality as the present UK estimates. Compliance costs would go up by £20–25 million from a base of £7.3 million for the ABI; and
- selecting 225,000 businesses rather than the full 570,000 is an intermediate option that widens the confidence intervals by a factor of around 1.3 compared with the full option. This can be done for less than half the costs but still represents an increase in compliance costs of around £7 million.

Introduction

A2.3 This Annex discusses issues relating to regional stratification of the ABI, enhancing the sample size and improving register proving of local units. The estimates of resources are broad brush estimates since optimising the ABI sample design to meet both UK and sub-national users’ needs, including those of quality, requires some considerable methodological effort. In particular, the resource estimates depend critically on what variables are required at what level of geography and the various form types that are used to achieve the required level of accuracy.

Stratification: an introduction

A2.4 The number of NUTS and SIC codes at the various levels are shown in the table A2.1 below.

Table A2.1: ABI cell structure

NUTS (GB) level	Number of regions	SIC level	Number of industries
Countries	3	Section	17
1	11	Division (2 digit)	62
2	36	Group (3 digit)	225
3	128	Class (4 digit)	517

A2.5 The ABI is currently stratified at 4-digit SIC by country and by 6 employment sizebands, i.e. $517 \times 3 \times 6 = 9,306$ cells across the whole economy from the above table. But ABI2 does not cover the whole economy and just over 6,000 cells are actually used for ABI2 stratification. Up to 2001, Northern Ireland had been a fourth regional category but from 2002 there is a separate Northern Ireland ABI run by the NI Department of Enterprise, Trade and Industry.

A2.6 Sample stratification is important as a means of optimising the accuracy of a survey as a whole. It also serves to control the sample size allocated to particular sub-categories (e.g. regions) and hence the accuracy for those sub-categories. The separate categories for Scotland and Wales (and previously Northern Ireland) in the ABI stratification were introduced so that the sample sizes for these regions could be separately controlled. The English regions are not categorised separately. Thus, the sample within England is allocated to optimise estimates for England as a whole and the individual regions receive whatever sample is necessary to achieve that. If separate control of sample size and accuracy is required for the English regions then regional stratification within England would need to be introduced.

A2.7 For a given survey sample size, the number of categories in the sample stratification is constrained by the need to have an adequate sample in each stratum. Ideally we need at least 20 responding businesses in each of the cells which are sampled as opposed to completely enumerated. As noted above, the present ABI design has over 6,000 cells and around 74,000 businesses are selected. This looks inadequate and to introduce a more detailed regional stratification in ABI would require either a larger sample size overall or a reduction in detail for one of the other stratification dimensions (or a combination of both these approaches). Since the 'size' dimension is crucial to the accuracy of the survey, any reduction in detail would need to come from the SIC industry dimension. At present the industry stratification for England is at 4-digit SIC level. This could be reduced to the 2- or 3-digit levels or to some combination of 2/3/4 digit levels (according to the size of the industry in question) as is done for Scotland (referred to below as the 'Scottish option'). There are also complex issues of how to 'regionalise' large companies that straddle many regions.

Sample size and accuracy

Option 1: Maintaining the current sample size but improving regional representativeness **A2.8** It has been estimated that the current sample of around 74,000 provides 95 per cent confidence limits of ± 3.0 per cent for total GVA and ± 0.8 per cent for total employment at a GB level. Improving the regional estimates with the present sample size would necessarily involve regional stratification, almost certainly at the NUTS 1 level. This would give eleven geographic areas rather than the present three. The sample size in each region would not be large enough or representative to allow the current industry detail.

A2.9 Given the known user requirements for detailed industry level estimates at both 4- and 5-digit levels of the SIC, a move towards more detailed regional stratification at the expense of industry detail would require a change in the methodology for industry estimates in the ABI, which would now be required at a more detailed industrial level than provided by the stratification. Small industries would have small or even zero sample sizes and a method of synthetic estimation would need to be developed to provide estimates for such industries. This could be similar to the method currently used for the production of small area estimates in the ABI at present. Some methodological development would however be essential in order to meet the needs of both industry and regional customers.

A2.10 This option would not incur any additional on-going costs, but there would be considerable development time and costs that would be needed to take this forward.

Option 2: Producing regional results at NUTS1 level with the current UK level of precision **A2.11** To obtain a precise estimate of the sample size and costs in producing regional results with the current UK level of precision is a complex process involving considerable methodological effort, but a very rough estimate can be obtained as follows. Roughly a third of the ABI sample size is effectively a census of larger businesses and so numbers need not be increased. The remaining two thirds is a sample survey spread over the 11 NUTS 1 regions in GB. To obtain approximately the same accuracy from this at the regional level as is currently achieved at the GB level would require that its sample size be multiplied by around 11. (A feature of sample surveys is that, broadly, the accuracy of the estimates is determined by the sample size and, for large populations, the sampling fraction plays no part. Therefore, broadly you need the same size sample to produce results for the West Midlands as you do for the UK.) Overall this implies an increase in the total ABI sample size by a factor of about 7.7, to around 570,000.

A2.12 A simple adoption of the present approach of using long and short forms in the same proportions would also increase running costs and compliance costs by a factor of 7.7. However the increase in costs could be mitigated by using a less detailed form type for the additional sample. The short forms contain the key variables required but omit some of the more detailed breakdowns. Assuming that the forms sent to the additional sample incur only half the cost of average ABI forms sent to small and medium sized businesses the above proposal would increase compliance costs by a factor of about 4.3. Data collection costs would go up more, by a factor of around 6, because some costs are associated with the number of forms and some with the number of data items. This option would be very expensive in terms of ongoing costs and would ultimately require the same development resource as for option 1. However, with this option, some improvements can be delivered without any changes to the underlying methods and systems, simply by increasing the sample. This is technically feasible from the 2004 survey, but the financial implications are large, as would be the political impact of sending an ABI form to 1 in 4 businesses each year.

Option 3: An intermediate option **A2.13** A similar calculation to the above, assuming a sample size increase by a factor of 3 (to 225,000 – with the sampled portion rising by a factor of 4), would increase compliance costs by a factor of 2 and data collection costs by a factor of 2½. The reduction in accuracy at regional level provided by this option as compared with option 2 is very complex to estimate. However, again, a rough approximation is possible. Assuming that the census of large firms has about 50 per cent coverage of activity, the loss of accuracy lies only in the 50 per cent of activity covered by the sampled part of the population. The additional inaccuracy for this part of the population is affected approximately by the square root of the ratio of the sample sizes, i.e. 1½. Overall the confidence intervals would widen by a factor of about 1.3.

A2.14 This option is also expensive although obviously not as much as option 2. It would still add around £7 million or so to compliance costs and has a similar timescale for implementation.

Effect of stratifying at NUTS 2/3 level and the ‘Scottish option’

A2.15 As has been indicated above, an increase in the detail of the regional stratification requires a reduction in the detail of the industry stratification. The most pressing requirement appears to be separate stratification of the English regions at the NUTS 1 level. However there are also some important EU funding issues dependent on NUTS 2 and NUTS 3 level data. Not all regions at these levels would be of equal interest, so one possibility might be to adopt a mixed level of regional stratification at NUTS 1/2/3 levels similar to the ‘Scottish option’ for industry stratification, whereby they are not interested in a single level of industrial classification (e.g. 3 digit) but rather have an approach that identifies the particular need for the sector. This could be combined with an increase in overall sample size to maintain reasonable accuracy in both the industry and regional dimensions.

Non-sampling errors

A2.16 All of the above assumes that the only source of error is the result of sample size limitations. In fact there are two other significant error sources which need to be considered. The first is the quality of the business register. Greater expenditure on the Annual Register Inquiry would improve register quality, though quantification of benefit in relation to cost is very difficult.

A2.17 The second important issue relates to the reporting units used for the inquiry that are at the national level. For large businesses a modelling process has to be carried out to apportion data collected at this level to a level where it is useful for sub-national estimates. Work is ongoing to improve the model currently in use but the modelling process will always leave a residual and unquantifiable source of error in sub-national estimates.

A3 Organisations consulted and submissions received

A3.1 The Review has benefited greatly from the experience and views of the organisations and individuals we have met or who submitted written contributions. We would like to express our gratitude to all those that we have seen or who have submitted views.

A3.2 The Review Team met with the following bodies and individuals:

Association of Regional Observatories
Bank of England
British Chambers of Commerce
Business Strategies from Experian
Cambridge Econometrics
Confederation of British Industry
Croner Reward
Department for Culture, Media and Sport
Department for Education and Skills
Department for Environment, Food and Rural Affairs
Department of Trade and Industry
Department for Transport
Department for Work and Pensions
east midlands development agency
Fraser of Allander Institute
Government Office Regional Directors
GLA Economics
HM Treasury
Institute for Fiscal Studies
London Development Agency
National Institute for Economic and Social Research
Northern Ireland Economic Research Centre
Northern Ireland Department of Enterprise, Trade and Investment
Northern Ireland Statistics Research Agency
Office for National Statistics
Office of the Deputy Prime Minister
One North East
Oxford Economic Forecasting
Regional Development Agency Chairs
Jacob Ryten
Scottish Executive
Statistics Commission
Trades Union Congress

Welsh Assembly Government
Welsh Economy Research Unit
David Wroe

A3.3 In addition, we received written submissions from those listed below. The submissions are available on the internet at www.hm-treasury.gov.uk/allstopp.

Association of Regional Observatories
Countryside Agency
Department of Trade and Industry
GLA Economics
London Development Agency
Manchester Chamber of Commerce
Northern Ireland Department of Enterprise, Trade and Investment
Regional Development Agencies
Royal Statistical Society
Scottish Executive
South West of England Regional Development Agency
Welsh Assembly Government
West Midlands New Economics Group